Volume 47 Number 2 June 2023 ISSN 0350-5596

Informatica

An International Journal of Computing and Informatics

Editorial Boards

Informatica is a journal primarily covering intelligent systems in the European computer science, informatics and cognitive community; scientific and educational as well as technical, commercial and industrial. Its basic aim is to enhance communications between different European structures on the basis of equal rights and international refereeing. It publishes scientific papers accepted by at least two referees outside the author's country. In addition, it contains information about conferences, opinions, critical examinations of existing publications and news. Finally, major practical achievements and innovations in the computer and information industry are presented through commercial publications as well as through independent evaluations.

Editing and refereeing are distributed. Each editor from the Editorial Board can conduct the refereeing process by appointing two new referees or referees from the Board of Referees or Editorial Board. Referees should not be from the author's country. If new referees are appointed, their names will appear in the list of referees. Each paper bears the name of the editor who appointed the referees. Each editor can propose new members for the Editorial Board or referees. Editors and referees inactive for a longer period can be automatically replaced. Changes in the Editorial Board are confirmed by the Executive Editors.

The coordination necessary is made through the Executive Editors who examine the reviews, sort the accepted articles and maintain appropriate international distribution. The Executive Board is appointed by the Society Informatika. Informatica is partially supported by the Slovenian Ministry of Higher Education, Science and Technology.

Each author is guaranteed to receive the reviews of his article. When accepted, publication in Informatica is guaranteed in less than one year after the Executive Editors receive the corrected version of the article.

Automatic Detection of Stop Words for Texts in the Uzbek Language

Khabibulla Madatov[1], Shukurla Bekchanov[1], Jernej Vičič[2,3]
[1]Urgench state university, 14, Kh. Alimdjan str, Urgench city, 220100, Uzbekistan
[2]University of Primorska, UPFAMNIT, E-mail: jernej.vicic@upr.si
[3]Research Centre of the Slovenian Academy of Sciences and Arts, The Fran Ramovš Institute
E-mail: habi1972@mail.ru, shukurla15@gmail.com, jernej.vicic@upr.si

Keywords: stop word detection, Uzbek language, agglutinative language, algorithm

Received:

Stop words are very important for information retrieval and text analysis investigation. This study aimed to automatically analyze and detect stop words in texts in the Uzbek language. Because of the limited availability of methods for automatic search of stop words of texts in Uzbek we analyzed a newly prepared corpus. The Uzbek language belongs to the family of agglutinative languages. As with all agglutinative languages, we can explain that the detection of stop words in Uzbek texts is a more complex process than in inflected languages: In inflected languages, words such as auxiliary words, articles, prepositions can be included in the stop words group. In agglutinative languages, the meanings of such words are hidden in the text. Therefore, it is not appropriate to apply all known methods of stop words detection in inflected languages directly to agglutinative languages. In this work, the "School corpus" which contains 731156 Uzbek words has been investigated. The bigram method of analysis was applied to the corpus. We proposed the collocation method of detecting stop words of the corpus. We proposed the method of automatically detecting stop words of texts in Uzbek. It is shown that the collocation method is 6 times better than the bigram method.

Povzetek: Razvita je samodejna analiza in odkrivanje posebnih besed v uzbekistanskem jeziku.

1 Introduction

Uzbek language belongs to the Eastern Turkic or Karluk branch of the Turkic language family. External influences include Arabic, Persian and Russian. It belongs to the family of agglutinative languages. As with all agglutinative languages, detection of stop words in Uzbek texts is a more complex process than in inflected languages: in inflected languages, words such as auxiliary words, articles, prepositions form most of the stop words group. In agglutinative languages, the meanings of such words are hidden in the text. Therefore, it is not suitable to apply all known methods of stop words detection in inflected languages directly to agglutinative languages. The experimental results presented in this work that the use of a hybrid method (combining grammatical rules and statistical methods) yields best results in the task of detecting stop words for texts in Uzbek. As a result of this work we compare this method with bigram method (both methods are thoroughly presented in the paper). When someone works on a novel, a story, an article, or a text, this person uses semantic connection of words with artistic decoration in their own language to make it meaningful and interesting. Dealing with sentences, stop words which do not have an independent meaning or have little meaning are often used. As a result, the text size increases. As the volume of information increases, the process of data processing and analysis slows down and as the search space increases, the quality of results (searches) is potentially lowered. In such cases, removing unnecessary words from the text can reduce the amount of information and increase the efficiency of electronic data processing. It is also important to automate the generation of annotations and keywords from large volumes of text. The main purpose of identifying unimportant words is to facilitate automatic text analysis.

2 Related works

Stop words are used in a number of tasks involving language technologies, such as text generation [1] and Turkic languages are no exception, an example of applying stop words to sentiment discovery in Turkish language comments is presented in [2]. Stop words detection methods can be divided into two basic categories:

1. Based on grammar rules,
2. Statistical methods.

In this work, we use both categories for automatic detection of stop words in Uzbek texts.

2.1 Based on grammar rules

The sources mainly provide grammatical rules for finding stop words or a list of stop words for different languages [3],

[4], [5], [6], [7],[8],[9],[10], [11]. The text is grammatically analyzed to identify stop words in Uzbek texts. According to the definition of stop words, words in the Uzbek language that are part of a rhyme, conjunctions, introductory words, adverbs, auxiliary words can be stop words. It is required to automatically separate them from the given text. Due to the lack of syntactic analysis programs in the Uzbek language, using a dictionary, a list of words that are supposed to be stop words will be given. In order to create the list of stop words from the dictionary we investigate and take into account the definition of stop words. In general pronouns, adverbs, connectors, Introductory words can be stop words in Uzbek texts.

2.1.1 Pronouns

Pronoun is a part of speech used instead of a noun, adjective, number. The meaning of pronouns and which word or words they substitute is defined by the context (intra or inter sentence). According to the meaning and grammatical features, the pronoun is divided into generalized - subject (pronouns - nouns: men (I), sen (You), u (he, she, it), kim (who), nima (what), hechkim (nobody), hechnima (nothing), generalized - nominal (pronouns-adjectives: bu (this), shu (this), o'sha (that), qaysi (which), allaqanday (somehow), hechqanday (no), generalized quantitative (pronouns-numerals: qancha (how much), necha (how many), shuncha (so many), o'shancha (so many). Pronouns differ from other parts of speech in polysemy, lack of word formation. Pronouns by meaning and grammatical features are divided into the following types: pronouns of the person –men (I), sen (you), u (he,she,it), biz (we), ular (they) used instead of persons, the proper pronoun - consist of a proper word denoting an object, strengthening its meaning, emphasizing it; indicative pronouns –bu(this), shu (this), o'sha (that), u (he, she, it) , ana (that),etc. indicate an object and its signs; interrogative pronouns indicate the questions askim? (who) nima? (what) qancha? (how many) of the subject, attribute and quantity; definitive-collective pronoun - indicates the generalization, generalization of the subject and its features in relation to hamma (all), bari (all), ba'zan (sometimes), har nima (anything), har qanday (every/any) indefinite pronoun - expresses the denial of meaning in relation to hechkim (noone), hechqanday(no), hechqanaqa (no), hechqaysi (none).

2.1.2 Adverbs

An adverb is one of the independent types of Parts of speech denotes a sign of an action and a state, as well as a sign of a sign. There are the following types of adverb meanings: state adverbs (tez(quick), sekin(slow), piyoda(on foot)); adverbs of a place (uzoqda(far), yaqinda(near), pastda(below); adverbs of time (hozir(now), kecha(yesterday), bugun(today); adverbs of quantity (ancha(much), sal(little), kam(few); adverbs of purpose ataylab(deliberately), jo'rttaga(willingly); adverbs of reason (e.g., noiloj(helplessly), ilojsiz(helplessly), chorasizlikdan(helplessly). All forms, except for the forms of time, place and purpose, according to the most general characteristics, can be attributed to one type and calledas status forms. An adverb as an independent phrase is characterized by the following morphological features:

- has a category of degree: tez (quick), ko'p(much) (oddiy daraja(simple degree)) — tezroq (quicker), ko'proq (more) (comparative degree) — eng tez(the quickest), judako'p(much more)(superlative degree);
- remains unchanged and is often associated with verbs: So'ridaqat-qatduxoba ko'rpachalar ustma-ust to'shalgan edi(The couch was covered with layers of velvet mattress);
- an adverb can also be associated with an adjective and a noun in some places. In such cases, the adverb does not indicate a sign of a sign or a sign of an object, but to the adjective to which it is attached, or to a sign of action understood from a noun: Kecha havo juda sovuq edi. (It was very cold last night).U hozir beqiyos va tasavvur qilib bo'lmas baxtiyor edi;(Now, he was incomparably, unimaginably happy);
- an adverb has suffixes: -cha, -ona, -larcha, -laband etc..
- Adverbs are formed in morphological and syntactic ways: (otlashish hollari bundan mustasno), for each time, including the moment, as in the moment (syntactic method).

By structure, adverbs are divided into simple (kamtarona (modest), vijdonan (conscientious), butunlay(whole)), compound (har dam (always), bir yo'la (together), oz muncha (much), har qachon (always)), paired (kecha-kunduz (day and night), qishin-yozin (winter and summer) and repeated (oz-oz (little by little), tez-tez (often), ko'p-ko'p(many-many). The modal form is considered to be such forms of verbs anchagina (much), juda (very), kam (little), kam-kam (little by little). Adverbs act in a sentence as case, determinant and cut. Adverbs are similar to adjectives in terms of the properties of the expression of a feature, but differ among themselves in grammatical properties: adjectives denote a feature of an object, an adverb that is a sign of an action or state; their function in the sentence, that is, syntactic, is also special.

2.1.3 Connectors

Connectors are auxiliary words that serve to link organized parts of a sentence and simple sentences in the structure of a combined sentence are called connecting words. Auxiliary words that connect two or more fragments of a sentence or sentence are called connectors.

Introductory words. Introductory words-words that are not syntactically related to the sentence. Expresses the speaker's attitude to the expressed thought ("baxtimga"

(fortunately), ”afsuski”(unfortunately)), the general assessment of the thought (”ehtimol”, ”albatta”), to whom it belongs (”menimcha”(to my mind), ”aytishlaricha”(it is said)) or its connection with the preceding thought (”xullas”(so), ”nihoyat”(finally)). Words used in a sentence in the function of an introductory word, expressing the speaker's attitude to the expressed thought, are called modal words. Modal words are not independent words, such as a thing, sign, action, etc., and cannot be part of a sentence. Therefore, they are not syntactically related to the fragments of the sentence: “Demak, ishlasa bo‘ladi. Ehtimol, ketmon bilan yer ag‘darishga ham to‘g‘ri kelar.” (So, it can endure to work. Perhaps, you may even have to roll over with a hoe).

2.1.4 Usage

The rules described in previous subsections were used to detect stop words. The popular explanatory dictionary of the Uzbek language [3] with 80000 words and detected approximately 1100 stop words. These are by definition one-word stop words as they come from the dictionary.

2.1.5 Statistical method

Consider the statistical method of automatic detection of stop words in Uzbek texts. In this method, stop words are found based on the frequency of the word and the frequency of the inverse document Term Frequency – Inverse Document Frequency – TF-IDF [12]. The number of times of word occurrence in a text is defined by Term Frequency – TF. Inverse Document Frequency – IDF is defined as the number of texts (documents) being viewed and the presence of a given word in chosen texts (documents). TF-IDF is one of the popular methods of knowledge discovery. There are such words that are so common in the text, however they are almost insignificant in terms of meaning and conversely, there are words that are rare in the text, but they are very important in terms of the meaning of the text. In order to increase the impact of meaningful words and decrease the frequency of words that do not add up much to the meaning, we multiply TF to IDF. We see the statistical method is used as the basis for finding stop words of many languages. The sources mainly use the TF IDF method to analyze of the word of the text [4], [13], [14],[15],[16],[17],[18],[19],[20],[21],[22]. we see the statistical method is used as the basis for finding stop words of many languages. These sources mainly use the TF IDF method to analyze of the word of the text. Several methods to find stop words for Turkish are given in [18]. Comparing the current work with these sources, we bring scientific novelty of the article.

3 Methodology

Scientific novelty of the article. First, a collocation method is proposed for automatically finding stop words of the Uzbek corpus, consisting of 731156 words and comparing with bigram method its advantage is shown. Second, stop words detecting algorithm is proposed for Uzbek texts.

This section is dedicated to the method of automatic detection of stop words in Uzbek texts. The following procedure was used for detection of stop words.

3.1 Corpus

A corpus named “School corpus” was created using freely available school books such as “Reading book”, “Mother tongue” and “Literature”. The texts were downloaded form Eduportal [1]. Total number of documents is 25. The motivation behind the selection of the texts for the corpus was the following:

- everyone enriches personal language dictionary knowledge during the school period,
- free availability of the texts,
- school textbooks are thoroughly checked for errors,
- marginally big enough selection of documents and length of the documents (taken into account low availability of Uzbek texts in digital form in general).

Some basic data about the corpus:

- name: *School corpus*,
- total number of words: 731156,
- number of unique words: 47165.

The investigation on finding stop words in the Uzbek language has shown that in most stop words that are collocations, each single word is not a stop word when viewed as individual word, but when considered as a collocation word, they become stop words. A few examples that further confirm our claim are presented in the Examples 3.1 and 3.2 where the meaning of the sentences is the same. When viewed as individual words the words *bir* and *martalik* are not stop words, but if they are observed as a collocation, they become stop word(s).

Example 3.1. *Xalqimiz bir martalik shprits vositasida emlanadi.* – (Our people are vaccinated with a disposable syringe.)

Example 3.2. *Xalqimiz shprits vositasida emlanadi* – (Our people are vaccinated with syringes.)

Thus, there is a need to expand the problem of finding stop words which consist of one word. A collocation is considered if there are 2 or more words. Only a two-word collocation is considered in this article and the motivation behind this is that three or more word collocations that act as stop words are not that common, but we still believe that a further work needs to be done in this direction. The proposed methodology does not change for longer collocations.

[1] Eduportal: `https://eduportal.uz/Eduportal/Barchasi/33`

3.2 Bigram method

For the purpose of this article, the following definition will be used: bigrams are pairs of consequent words appearing in the text. Let's consider the use of the bigram method of finding stop words for the corpus. Algorithm 1 presents the implementation of the bigram method.

Algorithm 1: The bigram method

1. Consider total occurrences a_i, a_{i+1} of collocation words in a corpus. Construct a list of unique pairs $UP1$. In our corpus example, the number of such collocations was 731155. Among them 489857 unique pairs.
2. Consider the list $UP1$, for each pair a_i, a_{i+1} from the list take a_i and find the word with the biggest bigram probability in the corpus for the next word a'_{i+1}. There were 90959 (a_i, a'_{i+1}) unique pairs $UP2$.
3. Calculate term frequency (TF) of unique pairs $UP2$, for each document in the corpus. In our example corpus that meant for each of the 25 documents. We denote it as $DjTF(a_i, a'_{i+1}), j = 1.\overline{25}$.
4. $DjTF(a_i, a'_{i+1}) = k_j/h_j$, where h_j is the number of occurrences of the pair words in the document j. k_j is the number of unique pairs in document j.
5. $IDF(a_i, a'_{i+1}) = ln(n/m); n = 25$. m is the number of documents which include unique pairs (a_i, a'_{i+1}), in our example among 25 documents.
6. $W_{ij}(a_i, a'_{i+1}) = \frac{1}{25} \sum_{j=1}^{25} IDF(a_i, a'_{i+1}) * D_jTF(a_i, a'_{i+1})$
7. $W_{ij}(a_i, a'_{i+1})$ – weights of unique pairs .
8. We got 5% of the 90957 unique pairs, which $W_{ij}(a_i, a'_{i+1})$ is close to zero and declare them as stop words.

The Algorithm 1 applied to the "School corpus" produced 4548 pairs of words as collocation stop words. A few examples are presented in Figure 1 bigram.

```
1. chop etildi(published)
2. har bir(each)
3. kitob jamgarmasi(book fund)
4. nima uchun(what for)
5. o'rta talim(secondary education)
6. men ham(me too)
7. bilan birga(along with)
8. yaxshi muqova(good cover)
9. oz vaqtida(It's on time)
10. ham bir(also a)
11. bir necha(a few)
12. barcha varaqlari(all sheets)
13. o'zi ham(himself)
14. bu yerda(here)
15. bo'lib qoldi(has become)
16. u ham(he too)
17. uchun ham(for both)
18. uning bu(its this)
19. butun darslikning(of the whole textbook)
20. yangi darslikning(new textbook)
.....
4529. Velosiped baxtiga(Luckily for the bike)
4530. Vodiy daralariga(To the gorges of the valley)
4531. Voqealarga aralashadi(Interferes with events)
4532. Xarakteri amallari(Character actions)
4533. Xarakterini izohlang(Explain the character)
4534. xonimning uylariga(to the lady's house)
4535. xoqonning hayoti(the life of a hawk)
4536. xotirasini abadiylashtirish(immortalize the memory)
4537. xudoyor davron(godly era)
4538. xushxabar ammo(The good news, however)
4539. yapon arab(Japanese arab)
4540. yasagan qayiqlarni(made boats)
4541. yasalgan fe'llar(made verbs)
4542. yaxshilar ahbob(good fellow)
4543. yig'isi alomatning(crying symptom)
4544. yig'lagan bolasini(crying baby)
4545. yig'och chog'liq(wood chips)
4546. yodlang islom(Remember Islam)
4547. yo'lakda bir(one in the hallway)
4548. yo'llardan biri(one of the ways)
```

Figure 1: Examples selected from the list of all stopwords generated by the bigram Algorithm 1.

3.3 The collocation method's algorithm

The following definition of collocation will be used throughout the article: an occurrence of consecutive words in a corpus. In our case only two-word collocations will be observed. A (two word) collocation and bigram represent essentially the same starting set of word pairs, but bigrams are limited to the most probable pair, collocations take in consideration all pairs.

A collocation is considered for 2 or more words. In this article only a two word collocation is considered. The presented method and derived results are limited to two word collocations for the sake of simplicity, but the method can be abstracted to any length. The method should be used before the single stop word detection method.

To find collocation stop words, we use the following Algorithm 2:

The Algorithm 2 applied to the "School corpus" produced 24490 pairs of words as collocation stop words. A few examples are presented in Figure 2.

3.4 Single word (stop word) detection algorithm

In this section we consider the single word stop words detecting algorithm based on TFIDT(Term frequency and inverse document frequency) of the word. To find single word stop words, we use the following Algorithm 3:

The Algorithm 3 applied to the "School corpus" produced 2358 stop words. A few examples are presented in Figure 4.

3.5 The final stop word detection Algorithm for Uzbek language

In this section we consider the main algorithm of detecting Stop words of text in Uzbek language. We bring this algorithm as in the scheme presented on Figure 3.

Algorithm 2: The collocation method

1. Consider all occurrences of collocations in a corpus. In our case the total number of such collocations was 731155. Among them 489857 collocation words are unique collocation words.
2. $DjTF(a_i, a_{i+1}) = k_j/h_j$, where h_j is the number of occurrences of the pair words in the document j. k_j is the number of unique pairs in document j.
3. $IDF(a_i, a_{i+1}) = ln(n/m); n = 25$. m is the number of documents which include unique pairs, in our example among 25 documents.
4. $Wij(a_i, a_{i+1}) = \frac{1}{25}\sum_{j=1}^{25} IDF(a_i, a_{i+1}) * D_jTF(a_i, a_{i+1})$
5. $Wij(a_i, a_{i+1})$ – denotes weight of a collocation – $(a_i a_{i+1})$.
6. 5 % of all unique collocations had the weigh $W_{ij}(a_i, a_{i+1})$ close to zero and were declared as stop words.

```
1. har bir(each)
2. nima uchun(what for)
3. bir kuni(one day)
4. o'rta talim(secondary education)
5. uchun darslik(textbook for)
6. chop etildi(published)
7. kitob jamg'armasi(book fund)
8. abad ham(forever)
9. abadiy kuchidan(from eternal power)
10. abadiy manziliga(to the eternal address)
11. abadiy muhrlanib(sealed forever)
12. abadligi hamda(   eternity and)
13. Abadul abad badnom(Abadul abad badnom)
14. Abadul abad turajakdur(It will last forever)
15. Abay singari(Abay suchlike)
16. Abbos degan(Abbos named)
17. Abbos qilichi(The sword of Abbas)
18. Abdulaziz qaytib(Abdulaziz returned)
19. Abdulazizga qaradi(He looked at Abdulaziz)
.....
24471. Odamlarni ko'rishadi(They see people)
24472. Odamlarning chehralari(Faces of people)
24473. Odamlarning haqiga(About people)
24474. Odamlarning kamligi(Lack of people)
24475. Odamlarning ko'zidan(From people's eyes)
24476. Odamlarning ko'zini(People's eyes)
24477. Odamlarning nomlarini(The names of the people)
24478. odamlarning og'irini(the weight of people)
24479. odamlarning qaysi(which of the people)
24480. odamlarning va(people and)
24481. odamlarning zilzila(earthquake of people)
24482. odamligi uni(humanity him)
24483. odamligini ham(that he is human)
24484. odamligini ta'minlab(providing humanity)
24485. odamlik qiyofasini(human image)
24486. odamman deb(that I am human)
24487. odamman deganini(I mean man)
24488. odamma saxir(I'm sorry)
24489. odamni ajdodlari(man's ancestors)
24490. odamni ona(mother of man)
```

Figure 2: Examples selected from the list of all stopwords generated by the collocation Algorithm 2.

Algorithm 3: Single word (stop word) detection algorithm

1. $DjTF(a_i) = k_j/h_j$, where h_j is the number of occurrences of the pair words in the document j. k_j is the number of unique pairs in document j.
2. $IDF(a_i) = ln(n/m); n = 25$. m is the number of documents which include unique pairs, in our example among 25 documents.
3. $W_{ij}(a_i) = \frac{1}{25}\sum_{j=1}^{25} IDF(a_i) * D_jTF(a_i)$
4. $W_{ij}(a_i)$ – denotes weight of a word $(a_i$.
5. 5 % of of the 47165 unique words, which $W_{ij}(a_i)$ was close to zero and declared stop words.

Algorithm 4: find and remove Uzbek stop words from text (Corpus)

```
Input(Corpus)
Corpus ←Tokenize(Corpus)
Dictionary ← Extract_From_Dictionary(pronoun,
  modal verb, particle, part of a rhyme, conjunctions,
  introductory words, adverbs, auxiliary words)
;                  // Procedure Check(Corpus)
i←1
while i < len(Corpus) do
    if Corpus(i)∈Dictionary then
        Corpus ← Corpus − Corpus(i)
    i←i+1
/* Procedure Collocation_Two_Words
   (Corpus)                                    */
Corpus ←Tokenize(Corpus)
i←1
while i < len(Corpus) do
    S(i)← token(i)+token(i+1); i←i+1
/* Procedure IDF()                             */
IDF(S(i))←ln(N/n) ;        // N-number of all
 documents; n- number of documents,
 which include S(i)
/* Procedure TFIDF()                           */
j←1
while j < len(Corpus) do
    TF(j)←0
    i←j; while i < len(Corpus) - 1 do
        if S(j)==S(i) then
            TF(j) ← TF(j) + 1
        i←i+1;
    TFIDF(j)←TF(j)*IDF(S(j) if TFIDF(j) close to
    zero then
        Dictionary(j)←S(j);
        i←1; while i < len(Corpus) do
            if Dictionary(j)==Corpus(i) then
                Corpus←Corpus−Dictionary(i);
            i←i+1
    j←j+1
```

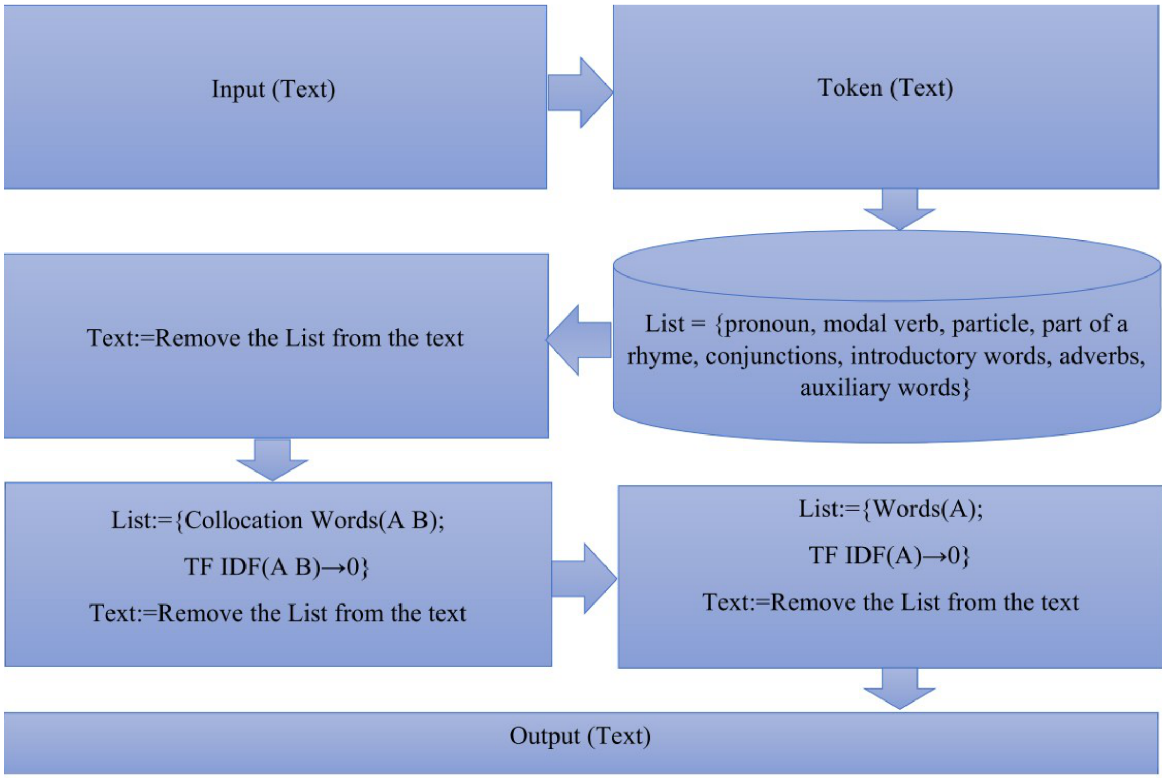


Figure 3: Scheme of the whole process.

```
1. Abdulla(Abdulla)
2. aka(brother)
3. asosida(based on)
4. ayt(say)
5. aytib(telling)
6. aziz(dear)
7. baho(evaluation)
8. bahor(spring)
9. baland(high)
10. beradi(will give)
11. berdi(gave)
12. berib(giving)
13. berilgan(given)
14. bering(read)
15. bichimi(physique)
.....
2344. badiiyatni(art)
2345. bag'ayri(past)
2346. bag'rimdami(in my heart)
2347. baid(height)
2348. balladalar(ballads)
2349. banddin(occupied)
2350. Bandksoy (Bandkushoy)
2351. barchalarining(all of them)
2352. barglarga(to the leaves)
2353. bastai(composer)
2354. baxilga(stingy)
2355. baxtdan(happily)
2356. baytallarga(beetles)
2357. bazmni(party)
2358. begonani(outsider)
```

Figure 4: Examples selected from the list of all stopwords generated by the single word extraction Algorithm 3.

Table 1: Number of stop words created by each presented algorithm.

Algorithm	Number of stop words
Bigram	4548
Collocation	24490
Single word	2358

4 Results

The first phase of the project consisted of creating a solid base for corpus linguistics as there were o readily available corpora for Uzbek language. A corpus named "School corpus" was created with 731156 running words. The algorithms for stop words detection ere applied to the aforementioned corpus and Table 1.

5 Data availability

The presented automatically extracted lists (a list for each described method) are freely available at Zenodo repository [23]: https://doi.org/10.5281/zenodo.6319953

6 Conclusion and further work

The article presents the first attempt at the automatic detection of stop words for Uzbek language. A corpus named

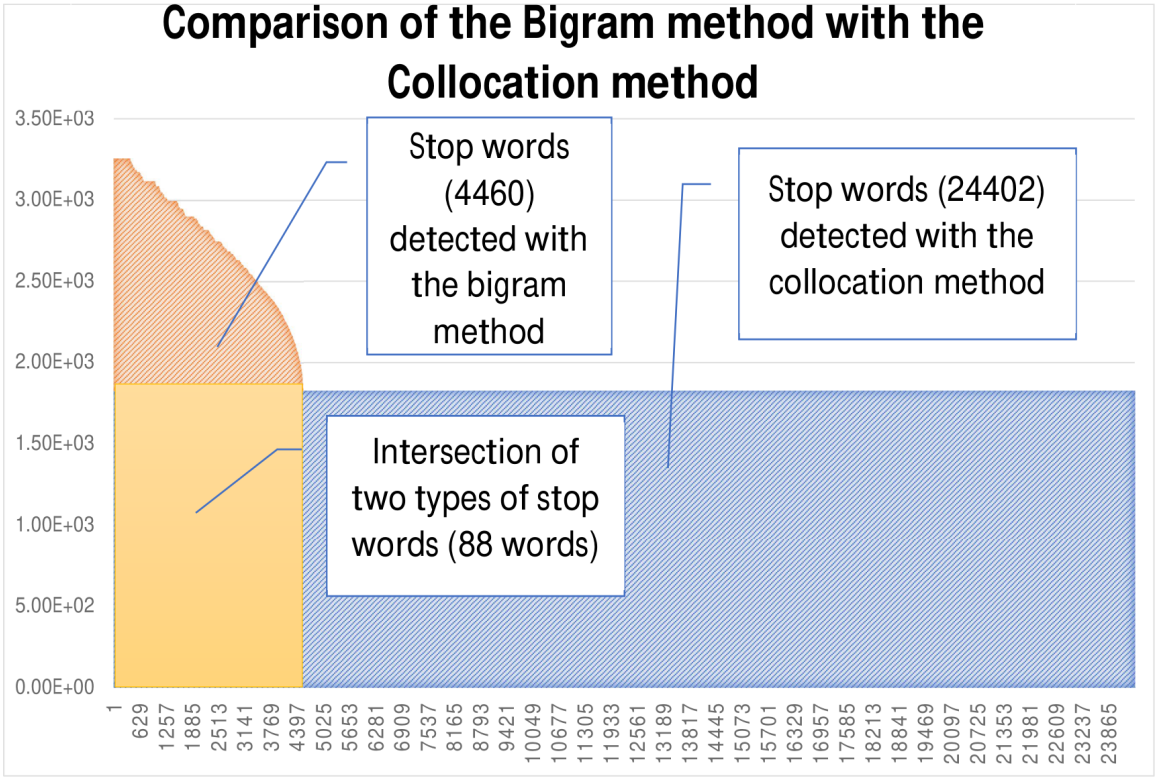


Figure 5: Number of stop words for each algorithm applied to the ”School corpus”.

”School corpus” was created for this purpose, it contains 25 documents and 731155 running words, of which 47165 are unique words. Three methods were applied to the corpus in order to extract (or detect) stop words: a method that extracts single word stop words and two methods that aim at pairs of words, a bigram and collocation method. Each method is described and presented in a form of an algorithm. The methods can be used in a series and the results can be added together to form the final list of stop words.

Taking account the conception of stop words depending on the text every word can be stop words. According to this approach (based on TFIDF). A quick comparison of the methods shows an increase in stop words detection using the collocation method

This research is believed to support other works in Uzbek, not only in the field of automatic stopword detection, but also other related NLP areas [24], such as Uzbek WordNet [25], opinion mining [26], or semantic analysis [27].

References

[1] P. Tomašič, G. Papa, and M. Žnidaršič, “Using a genetic algorithm to produce slogans,” *Informatica*, vol. 39, no. 2, 2015.

[2] R. Yayla and T. T. Bilgin, “Determining of the user attitudes on mobile security programs with machine learning methods,” *Informatica (Slovenia)*, 2021. [Online]. Available: https://doi.org/10.31449/inf.v45i3.3506

[3] S. Matlatipov, X. Madatov, G. Matlatipov, A. O‘razbayev, M. Raximboyev, I. Avezmatov, U. Babajanov, L. Kurbanova, D. Xujamov, and D. Matjumayeva, “”o‘zbek tilining statistik electron lug‘at” exm dasturi uchun guvohnoma,” *Intellektual mulk agentligi*, 2020.

[4] A. W. Pradana and M. Hayaty, “The effect of stemming and removal of stop words on the accuracy of sentiment analysis on indonesian-language texts,” *Game Technology, Information System, Computer Network, Computing, Electronics, and Control Journal*, vol. 4, no. 3, pp. 277–288, 2019. [Online]. Available: https://doi.org/10.22219/kinetik.v4i4.912

[5] R. U. Haque, P. Mehera, M. F. Mridha, and M. A. Hamid, “A complete bengali stop word detection mechanism,” in *Conference Paper · May 2019*. Conference, 2019. [Online]. Available: https://doi.org/10.1109/ICIEV.2019.8858544

[6] R. Rania and D.K.Lobiyal, “Automatic construction of generic stop words list for hindi text,” in *International Conference on Computational Intelligence and Data Science*, vol. 132, International Conference on Computational Intelligence and Data Science. ICCIDS 2018, 2018, pp. 362–370.

[7] P. J. Burns, "Constructing stoplists for historical languages," *Digital Classics Online*, vol. 4, no. 2, 2018. [Online]. Available: https://doi.org/10.11588/dco.2018.2.52124

[8] R. M. Rakholia and J. R. Saini, "A rule-based approach to identify stop words for gujarati language," in *In Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications*, 2017, pp. 797–806.

[9] J. K. Raulji and J. R. Saini, "Generating stopword list for sanskrit language," in *In: 2017 IEEE 7th International Advance Computing Conference*. IEEE 7th, 2017, pp. 799–802.

[10] O. D. Tijani, A. T. Akinwale, S. A. Onashoga, and E. O. Adeleke, "An auto-generated approach of stop words using aggregated analysis," in *In: Proceedings of the 13th International Conference of the Nigeria Computer Society*, 2017, pp. 99–115.

[11] M. Mhatre, D. Phondekar, P. Kadam, A. Chawathe, and K. Ghag, "Dimensionality reduction for sentiment analysis using pre-processing techniques," in *In Proceedings of the IEEE 2017 International Conference on Computing Methodologies and Communication*. ICCMC, 2017, pp. 16–21. [Online]. Available: https://doi.org/10.1109/ICCMC.2017.8282676

[12] C. Sammut and G. I. Webb, Eds., *TF–IDF*. Boston, MA: Springer US, 2010, pp. 986–987. [Online]. Available: https://doi.org/10.1007/978-0-387-30164-8_832

[13] Y. Wang, K. Kim, B. Lee, and H. Y. Youn, "Word clustering based on pos feature for efficient twitter sentiment analysis," *Human-centric Comput*, vol. 8, no. 17, pp. 1–25, 2019. [Online]. Available: https://doi.org/10.1186/s13673-018-0140-y

[14] N. Ousirimaneechai and S. Sinthupinyo, "Extraction of trend keywords and stop words from thai facebook pages using character n-grams," *International Journal of Machine Learning and Computing*, vol. 8, no. 6, 2018.

[15] C. Slamet, A. R. Atmadja, D. S. Maylawati, R. S. Lestari, W. Dharmalaksana, and M. A. Ramdhani, "Automated text summarization for indonesian article using vector space model model," in *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 288, no. 1, Conference. IOP, 2018. [Online]. Available: https://doi.org/10.1088/1757-899X/288/1/012037

[16] G. Li and J. Li, "Research on sentiment classification for tang poetry based on tf-idf and fp-growth," in *Proceedings of 2018 IEEE 3rd Advanced Information Technology*, Electronic and Automation Control Conference. IAEAC, 2018, pp. 630–634. [Online]. Available: https://doi.org/10.1109/IAEAC.2018.8577715

[17] H. M. Zin, N. Mustapha, M. A. A. Murad, and N. M. Sharef, "The effects of pre-processing strategies in sentiment analysis of online movie reviews," in *AIP Conf. Proc.*, vol. 1891, no. 1. AIP Conf., 2017, pp. 1–7. [Online]. Available: https://doi.org/10.1063/1.5005422

[18] S. K. Metin and B. Karaog'lan, "Stop word detection as a binary classification problem," *Anadolu University Journal of Science and Technology A- Applied Sciences and Engineering*, vol. 18, no. 2, pp. 346–359, 2017. [Online]. Available: https://doi.org/10.18038/aubtda.322136

[19] J. K. Raulji and J. R. Saini, "Generating stop word list for sanskrit language," in *In Advance Computing Conference IEEE 7th International*. IEEE, 2017, pp. 799–802.

[20] S. J. R. Rakholia R. M., "A rule-based approach to identify stop words for gujarati language," in *Suresh Chandra Satapathy Vikrant Bhateja Siba K.*, 2017.

[21] R. M. Rakholia and J. R. Saini, "Information retrieval for gujarati language using cosine similarity based vector space model," in *Theory and Applications*. Springer_Singapore, 2017, pp. 1–9.

[22] X.Madatov and S. Matlatipov, "Kosinus o'xshahshlik va uning o'zbek tili matnlariga tatbiqi haqida," *O'zMU xabarlari*, vol. 2, no. 1, 2016.

[23] K. Madatov, S. Bekchanov, and J. Vičič, "Lists of uzbek stopwords (1.1) [data set]," Zenodo. [Online]. Available: \url{https://doi.org/10.5281/zenodo.6319953}

[24] K. Madatov, S. Bekchanov, and J. Vičič, "Dataset of stopwords extracted from uzbek texts," *Data in Brief*, vol. 43, p. 108351, 2022.

[25] K. A. Madatov, D. Khujamov, and B. Boltayev, "Creating of the uzbek wordnet based on turkish wordnet," in *AIP Conference Proceedings*, vol. 2432, no. 1. AIP Publishing LLC, 2022, p. 060009.

[26] S. Matlatipov, H. Rahimboeva, J. Rajabov, and E. Kuriyozov, "Uzbek sentiment analysis based on local restaurant reviews," *arXiv preprint arXiv:2205.15930*, 2022.

[27] U. Salaev, E. Kuriyozov, and C. Gómez-Rodríguez, "Simreluz: Similarity and relatedness scores as a semantic evaluation dataset for uzbek language," *arXiv preprint arXiv:2205.06072*, 2022.

EU Smart Cities: Towards a New Framework of Urban Digital Transformation

Miljana Shulajkovska*, Gjorgji Noveski, Maj Smerkol, Jure Grabnar, Erik Dovgan and Matjaž Gams
Department of Intelligent Systems, Jožef Stefan Institute
Jamova cesta 39, 1000 Ljubljana, Slovenia
E-mail: miljana.sulajkovska@ijs.si, g.noveski@yahoo.com, maj.smerkol@ijs.si, jure.grabnar@ijs.si, erik.dovgan@gmail.com, matjaz.gams@ijs.si
*Corresponding author

Keywords: smart city, mobility policy, decision making, machine learning

Received: May 30, 2023

The URBANITE H2020 project aims to address urban mobility challenges caused by growth and new transportation methods. It develops a decision support system for policymakers, incorporating simulation, evaluation of key performance indicators, a recommendation/decision support system, and machine learning capabilities. The system helps identify and improve key performance indicators, proposes effective policies, and enhances urban digital transformation for sustainable and efficient mobility.

Povzetek: Podan je pregled novih storitev, razvitih v H2020 projektu pametnih mest Urbanite.

1 Introduction

Rapid urbanisation and population growth [1] pose significant challenges for modern cities. Smart cities (SC) have emerged as a solution for sustainable development, leveraging technology and data to enhance citizens' quality of life [2, 3, 4, 5]. In the context of the European Union (EU), the concept of SC has been a focal point of urban development and digital transformation initiatives. Several studies and research papers have explored various aspects of EU smart cities and their journey towards a new framework of urban digital transformation. Neirotti et al. analysed European SC by exploring their potential for innovation and sustainability [6]. Kitchin et al. examined the enabling and success factors in the development of SC in Europe [7]. Hollands conducted a systematic analysis of SC initiatives in Europe, highlighting diverse approaches and strategies [8]. Deakin et al. explored the role of policy in shaping smart urban futures in Europe [9]. These studies shed light on technology, governance, citizen engagement, and policy frameworks in the transformation of EU SC. Four cities were selected for this study: Bilbao, Amsterdam, Helsinki, and Messina. Each city is actively addressing specific transportation challenges. Bilbao, in Spain's Basque Country, has implemented measures to reduce pollution and congestion by closing city centre streets to private vehicles. Amsterdam, the capital of the Netherlands, focuses on cyclist safety and promoting a cyclist-friendly environment. Helsinki, the capital of Finland, plans to construct a tunnel near the port to enhance mobility and reduce congestion. Messina, in Italy, aims to improve its public transport network by introducing new lines for better accessibility and connectivity. These four cities serve as valuable case studies, illustrating different approaches and initiatives in urban digital transformation. By analysing the experiences and strategies of Bilbao, Amsterdam, Helsinki, and Messina, valuable insights can be gained towards developing a new framework for urban digital transformation in the context of EU SC. In that context, Urbanite strives to create more liveable, inclusive, and resilient cities that leverage technology and innovation to address urban challenges, improve sustainability, and enhance the quality of life for citizens. By fostering collaboration and knowledge exchange, the project aims to accelerate the transformation of European cities into smart and future-ready urban centres.

In this paper, we propose a novel approach within the Urbanite project, addressing the specific challenges faced by each city through the utilisation of multiple modules. These modules include a simulation tool, subjective key performance indicators (KPIs) tailored to each city, a recommendation engine, and machine learning (ML) techniques. In the following sections, we provide a concise overview of the general schema and discuss each module individually, highlighting their functionalities and contributions to the overall framework.

2 Urbanite architecture

In this section, we introduce a new framework developed within the Urbanite project [10], aimed at implementing SC solutions throughout Europe. Urbanite aims to enhance the quality of life for urban residents by leveraging innovative technologies and sustainable practices. The project brings together multiple stakeholders, including municipalities, research institutions, and industry partners, to collaborate on creating smarter and more efficient cities. One of the key objectives of Urbanite is to foster the integration of various SC components, such as smart mobility, en-

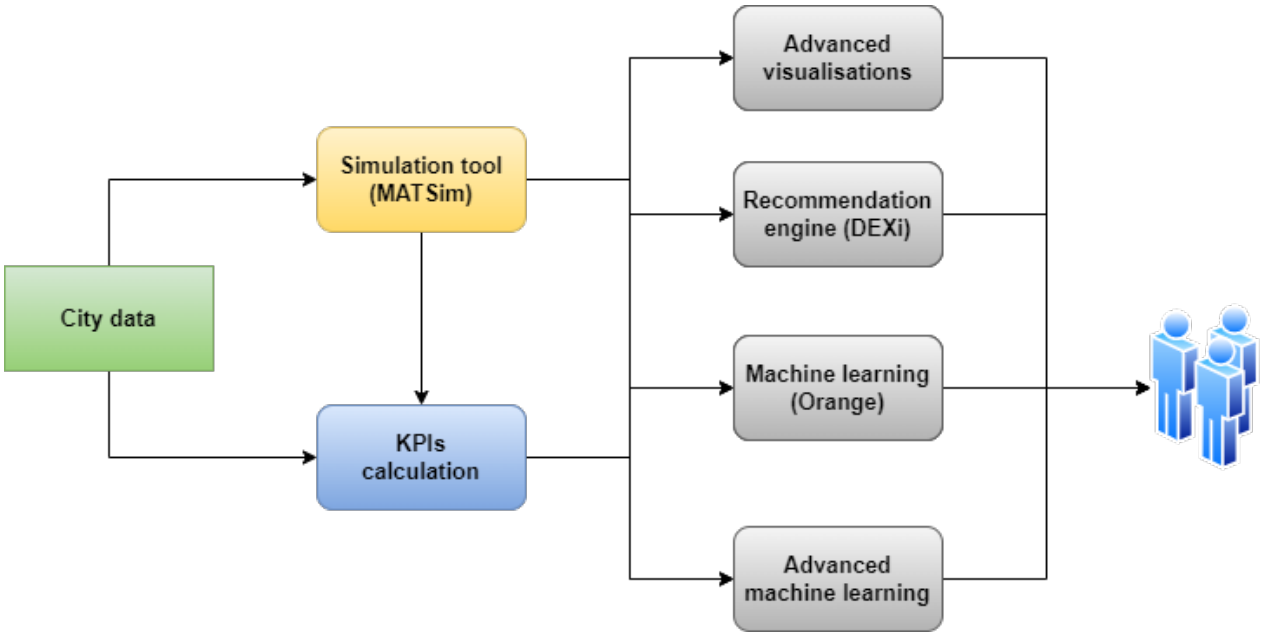


Figure 1: General schema of the software framework.

ergy management, and digital infrastructure. By harnessing data and technology, the project seeks to optimise urban services and resources, improve environmental sustainability, and enhance the overall urban experience. Urbanite promotes the concept of citizen-centric SC, where the needs and well-being of residents are at the core of urban development. It emphasises citizen engagement and participation in decision-making processes, encouraging the active involvement of communities in shaping the future of their cities. The project also focuses on promoting cross-sector collaboration and knowledge sharing among cities. Through pilot initiatives and best practice exchanges, Urbanite aims to facilitate the replication and scalability of successful SC solutions across different urban environments in Europe. A lot of research has been done on Urbanite presented in the following papers [11, 12, 13, 14].

The proposed software framework adheres to a general schema, as depicted in Figure 1. It begins by fetching data from a data platform and employing a microscopic traffic simulator to simulate a variety of scenarios. Following the completion of the simulations, specific KPIs defined by users are computed. These simulations, alongside the KPIs, are subsequently utilised by a range of modules, encompassing advanced visualisations, a recommendation system, and ML modules. Collectively, these modules provide policy recommendations and aid decision-makers in making well-informed choices.

The usage of microscopic traffic simulations has gained prominence as a cost-effective approach for testing, implementing, and evaluating mobility policies and urban changes, circumventing the expenses associated with real-world experiments. The simulator relies on city-related data such as population statistics, network maps, and public transit schedules to operate effectively. Once executed, the simulator enables the calculation of city-specific subjective KPIs related to factors such as air pollution, congestion, cyclist safety, and more. The resulting simulation output, coupled with the calculated KPIs, is then leveraged by other models integrated within the framework.

The recommendation system, implemented with the Dexi tool [15], compares two scenarios and selects the preferable option based on subjective preferences, such as lower CO_2 or NOx emissions.

The ML module, implemented using Orange [16], serves the purpose of evaluating the quality of mobility policies through microscopic traffic simulations. The user-friendly nature of Orange makes it accessible to users without a programming background.

Additionally, an advanced version of the ML module is utilised to propose mobility policies based on a previously simulated finite set of scenarios, further enhancing the framework's capabilities.

Overall, the proposed software framework incorporates multiple modules and techniques to facilitate the testing, evaluation, and recommendation of mobility policies, ultimately contributing to more informed decision-making processes in urban planning.

3 Simulation

The novelty of our study revolves around four distinct cities, each with its own unique set of demands and challenges. In order to address these challenges, we used MTASim (Multi-Agent Transport Simulation) simulation tool. In addition to MTASim, we evaluated several other

state-of-the-art simulation methods to address the challenges faced by the four cities in our study. These methods included SUMO (Simulation of Urban Mobility) and PT VISsum.

SUMO is a widely used microscopic traffic simulation tool capable of simulating large-scale transportation networks. It offers detailed modeling of individual vehicles, their interactions, and traffic dynamics. SUMO considers factors such as lane-changing, traffic lights, and road infrastructure to provide realistic simulations of urban traffic scenarios.

PT VISsum, on the other hand, focuses on public transport simulation. It enables the modeling of various aspects of public transport systems, such as schedules, routes, and passenger behavior. PT VISsum allows for the evaluation of public transport performance and the analysis of potential improvements in terms of efficiency, reliability, and passenger satisfaction.

With the successful identification of MTASim as the optimal approach, we proceeded to apply it to each of the four cities under study. By implementing MTASim, we aimed to tailor the solution to the specific demands and characteristics of each city, taking into account their individual requirements and objectives.

MATSim is a powerful simulation framework designed to model complex transportation systems. It employs an agent-based approach, simulating the behaviour and interactions of individual travellers within a network. MATSim operates by simulating the daily activities of each traveller, including their commuting patterns, mode choice decisions, and route selections. By capturing the heterogeneity of traveller behaviour, MATSim enables a detailed understanding of transportation dynamics and their implications for urban mobility. The simulation process begins with an initial demand, which is then simulated in the mobsim module and evaluated in the scoring module. The scoring module assesses transportation options and scenarios based on criteria like travel time, cost, environmental impact, and user preferences. Through iterative iterations, the simulation dynamically adapts and optimises system performance, responding to changing conditions and policy interventions via the replanning module. This cyclic process is illustrated in Figure 2.

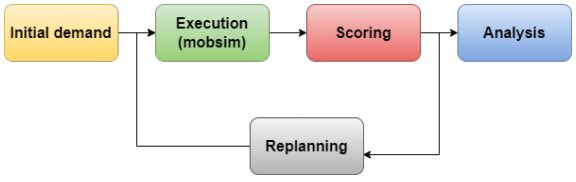


Figure 2: MATSim cycle.

To run the simulator, several input files are required. First, the network is generated from OpenStreetMap (OSM) data [17]. This network serves as the foundation for the simulation, capturing the road and transportation infrastructure of the studied area.

Next, travel plans are generated to simulate individual behaviour within the network. These travel plans dictate the movements and activities of simulated travellers, allowing the simulator to capture their interactions and decision-making processes.

In addition to the network and travel plans, other files are needed, including public transport schedules, descriptions of vehicles, and a configuration file that acts as a bridge between the user and the simulation tool. The configuration file allows users to fine-tune various parameters of the simulator according to their specific requirements and objectives.

Once the simulation is completed, several output files are generated. The most important of these is the event file, which contains a detailed description of people's movements and activities within the network. This file serves as a valuable resource for analysing and evaluating the simulated scenarios. Utilizing the event file, KPIs can be calculated to assess the efficiency, effectiveness, and other relevant metrics of the simulated transportation system.

4 Key performance indicators

4.1 Bike infrastructure

The KPI for bike infrastructure measures the extent and quality of the infrastructure available to support bicycle transportation. This includes factors such as the number of bike lanes, bike parking facilities, and the quality of road surfaces. The information taken into account is freely available from OSM. Based on the reported properties of the road a number of points is assigned to each road segment. Higher numbers are better, where 0 is a motorway (inappropriate and illegal to bike) to 10 (a bike-only road).

4.2 Bike speed limit

The KPI for bike speed limit refers to the maximum speed limit for bicycles on specific roads or bike lanes. This KPI is important for ensuring the safety of cyclists and other road users and promoting sustainable mobility by encouraging more people to cycle. The information taken into account is freely available from OSM. Based on the speed limit, each road segment is assigned a different number of points from 0 to 10.

4.3 Bikeability

The KPI for bikeability is a comprehensive metric that assesses the overall quality of the cycling environment. This KPI takes into account the bike infrastructure KPI and the bike speed limit KPI.

4.4 Bike intensity

The KPI for bike intensity measures the volume of bike traffic on a specific road or bike lane. This KPI is essential for understanding the usage and popularity of cycling as a mode of transportation and can help identify areas where improvements are needed to support increased bike traffic. The KPI is calculated by counting simulated bikes moving on each road segment.

4.5 Bike congestion

The KPI for bike congestion measures the level of traffic congestion experienced by cyclists on specific roads or bike lanes. This KPI is important for understanding the quality of the cycling experience and identifying areas where infrastructure improvements or traffic management strategies may be necessary to reduce congestion and improve safety for cyclists. The bike congestion KPI is calculated by first calculating the traffic flow on bikeable road segments and detecting low speeds with a high volume of bikes.

4.6 Share of bikes

This KPI measures the proportion of trips made by bicycle of all trips made. This allows the user to see where there are bikes and cars competing for the road surface, which can be dangerous, as well as identify areas where cycling should be encouraged either via infrastructure improvements, informing the public or other interventions. This KPI is complementary to the share of cars and the share of public transport.

4.7 Share of cars

This KPI measures the proportion of trips made by cars in a given area. It provides insights into the prevalence and effectiveness of car use as a mode of transportation, which can have significant impacts on urban mobility, air quality, and congestion. This KPI is complementary to the share of bikes and the share of public transport.

4.8 Share of public transport

This KPI measures the proportion of trips made by public transport vehicles, such as buses, trains, and trams, in a given area. It provides insights into the prevalence and effectiveness of public transport as a mode of transportation, which can have significant impacts on urban mobility, accessibility, and air quality. This KPI is complementary to the share of bikes and share of cars.

4.9 Acoustic pollution

This KPI measures the level of noise pollution in a given area, which can have significant impacts on public health, quality of life, and urban mobility. High levels of noise pollution can contribute to stress, sleep disturbance, and hearing loss. The acoustic pollution calculation is based on the simulated vehicle movements and geometry of buildings along the roads.

4.10 CO_2, PM_{10}, NOx

These KPIs measure the levels of carbon dioxide, particulate matter, and nitrogen oxides in a given area, which can have significant impacts on air quality, public health, and climate change. High levels of these pollutants can contribute to respiratory problems, cardiovascular disease, and other health issues. The amounts of air pollutants emitted are calculated based on the simulated vehicle movements and emission factors from the Handbook of Emission Factors (HBEFA).

4.11 Average pedestrian trip time

This KPI measures the average time it takes for pedestrians to complete a trip in a given area. It provides insights into the accessibility and quality of the pedestrian infrastructure, which can have significant impacts on urban mobility, safety, and public health. Due to limitations of the traffic simulation used the pedestrian trips do not take into account the infrastructure, only the distance between the source and destination. Therefore, this KPI is an estimation and not an exact value.

4.12 Congestions and bottlenecks

Congestions and bottlenecks are key performance indicators that help evaluate the efficiency of the urban mobility system. High levels of congestion can result in increased travel times, decreased accessibility, and reduced economic productivity. By monitoring and analysing the levels of congestion and bottlenecks, decision-makers can identify areas where traffic management interventions, such as lane restrictions or public transportation improvements, may be necessary. The congestions and bottlenecks KPI is implemented by calculating the traffic flow on each road segment and identifying segments with high volume but low speed.

4.13 Harbour area traffic flow

Harbour area traffic flow is a critical KPI in evaluating the efficiency of cargo transportation in urban areas. High levels of traffic flow can result in congestion and bottlenecks in the harbour area, leading to increased travel times and reduced economic productivity. This KPI is implemented by adding virtual traffic sensors to the relevant road segments of the simulation.

4.14 Public transport usage

This KPI measures the number of passengers using public transport services in a specific period. It is a critical metric for urban mobility decision-makers as it provides

insights into the demand for public transport services and helps identify potential opportunities to improve service quality and coverage to meet the needs of the public.

4.15 Average speed of public transport

This KPI represents the average speed of public transport vehicles in a given area or route. It provides insight into the efficiency and reliability of public transport services, as well as the effectiveness of traffic management policies. Improving the average speed of public transport can reduce travel time and encourage more people to use public transport.

4.16 Number of bike trips

This KPI measures the number of trips made by bicycles in a specific period. It is a critical metric for urban mobility decision-makers who aim to promote sustainable and healthy transportation alternatives. By encouraging more people to use bicycles, cities can reduce traffic congestion, improve air quality, and promote physical activity.

5 Decision Support System

Dexi is a decision tool that assists individuals and organisations in making informed choices by leveraging data and analytics. It is designed to simplify complex decision-making processes and provide actionable insights.

The tool offers a user-friendly interface that allows users to define decision criteria, set up models, and conduct scenario analyses. Dexi provides visualisation features to present the results in intuitive and understandable formats, such as charts, graphs, and dashboards. This helps users grasp the implications of different options and assess the potential outcomes of their decisions.

Dexi supports both strategic and operational decision-making processes across different domains. It can be applied to various use cases, such as financial planning, risk assessment, marketing campaign optimisation, supply chain management, and resource allocation. By leveraging advanced analytics, Dexi empowers users to make data-driven decisions that align with their goals and objectives.

Overall, Dexi is a versatile decision tool that combines data integration, analysis, and visualisation capabilities. It supports evidence-based decision-making, empowers users with actionable insights, and enhances efficiency in decision processes across various domains.

In the system, we utilise the output of Dexi in two ways. The first one is creating a textual suggestion which will inform the decision maker about which mobility policy is better in subjective terms with regards to another policy. The second way is visually by the usage of a chart.

6 Machine Learning

The purpose of the ML module in the Urbanite framework is to estimate the quality of a proposed policy without previously simulating it. The main concept centres around employing a single simulation run as a training example. Various groups of parameters associated with the simulation's input and output serve as the features, while the KPIs represent the target variables. To illustrate this approach, the city of Bilbao is selected as a practical case study. Our analysis focuses on the potential impact of closing Moyua Square in the city centre and altering the number of cyclists on air pollution, particularly by estimating CO_2 emissions. Multiple ML algorithms are tested, and the findings indicate that closing the main square in the city centre and promoting cycling has a positive effect on reducing CO_2 emissions.

To implement the idea in a user-friendly manner, Orange was used. Orange is a powerful machine learning (ML) tool developed by the research group at the Faculty of Computer and Information Science (FRI) and Jozef Stefan Institute (JSI) in Slovenia. It is open-source software that provides a user-friendly interface for data analysis, visualisation, and ML modelling. Orange is designed to make ML accessible to users without extensive programming knowledge. It offers a visual programming environment where users can create ML workflows by connecting pre-built components called "widgets." These widgets represent various data processing, analysis, and modelling techniques, allowing users to construct complex ML pipelines intuitively. With Orange, users can perform a wide range of tasks, including data preprocessing, feature selection, clustering, classification, regression, text mining, and more. It supports data visualisation and integration with other popular ML libraries and tools. The open-source nature of Orange encourages community involvement and contributions. Users can access the source code, contribute to its development, and create custom widgets tailored to their specific needs. The tool has an active user community, which provides support, shares resources, and promotes collaboration.

In this context, in Figure 3, the outcomes of the implemented policy using Orange visualisation widgets are illustrated. It depicts the correlation between the number of cyclists and the level of co_2 emissions. The x-axis represents the number of cyclists near the square, while the y-axis represents the number of cyclists in the centre. The varying colours indicate the number of co_2 emissions as the target variable, with the baseline scenario marked by an orange circle. The figure demonstrates that closing the main square to private traffic and decreasing the number of private vehicles in its vicinity leads to a reduction in co_2 emissions.

7 Advanced Machine Learning

Unlike the standard ML module, the advanced ML module leverages more sophisticated tools to tackle intricate problems. The key novelty of this module lies in its utilisation of multiclass-multioutput ML models, enabling the simul-

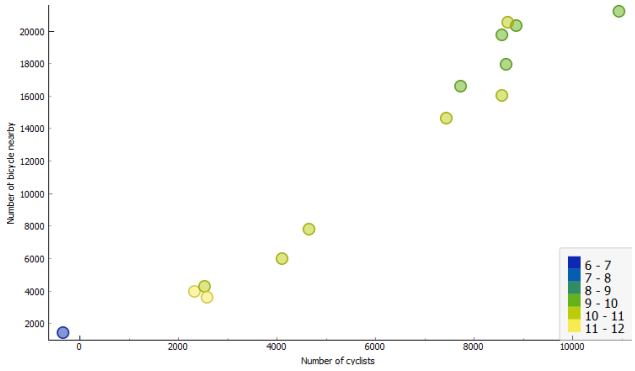


Figure 3: CO_2 emissions in the Moyua square (baseline scenario is marked with an orange circle).

taneous prediction of multiple outcomes using a diverse set of input variables. The primary goal of the ML module is to assist decision-makers in defining potential city scenarios and utility functions, allowing the ML model to identify policies that best align with given constraints and preferences. Notably, the module offers significant improvements in policy testing speed, with performance gains of several orders of magnitude. The system underwent testing in Bilbao's Moyua area, successfully achieving a predefined reduction in emissions and other KPIs. Furthermore, it provided valuable insights into optimal policies for closing specific districts to private traffic and determining the ideal timing for these closures based on data from simulated and learned scenarios.

The complexity of the problem lies in predicting multiple target variables which are discrete and continuous. The policy we want to predict is related to the start hour and duration of closing and what part of the city centre to close. Therefore the problem was split into classification (area of closure) and regression (start hour and duration of closure) tasks. Several ML algorithms that support multiclass-multioutput problems were tested.

Overall, the multiclass-multioutput module in the Urbanite project explains complex relationships between factors like traffic patterns and travel behaviour. It provides insights and recommendations to city planners, aiding informed decisions. This study is the first to address policy testing in a real city using multiclass-multioutput ML. Additionally, the ML module significantly speeds up simulations by several orders of magnitude, transforming time-demanding simulations into nearly interactive ML modules. The ML module reduces simulation time from 3 hours to just 10 seconds per simulation on a PC. While learning the ML module took 23 days for 192 simulations, running a total of 1452 simulations would take approximately 6 months. Experimental results demonstrate a high similarity between ML-simulated city performance and actual simulations.

8 Conclusion

In this paper, we tried to present the main software framework of the Urbanite H2020 project, which aims to address urban mobility challenges in the context of smart cities. The project develops a decision support system that incorporates simulation, evaluation of KPIs, recommendation/decision support, and ML capabilities. The goal is to identify and improve KPIs, propose effective policies, and enhance urban digital transformation for sustainable and efficient mobility.

The paper introduces the concept of smart cities and highlights their significance in addressing challenges posed by rapid urbanisation and population growth. It references several studies and research papers that have explored various aspects of smart cities in Europe, including innovation, sustainability, governance, citizen engagement, and policy frameworks. The Urbanite project builds upon this knowledge to create smarter and more efficient cities that prioritise the well-being of residents.

The paper focuses on four cities: Bilbao, Amsterdam, Helsinki, and Messina, which serve as case studies for understanding different approaches and initiatives in urban digital transformation. Each city faces specific transportation challenges, and the Urbanite project aims to tailor solutions to their unique requirements. By analysing the experiences and strategies of these cities, the project aims to develop a new framework for urban digital transformation in EU smart cities.

The proposed software framework consists of multiple modules, including a simulation tool, subjective KPIs, a recommendation engine, and ML techniques. The simulation tool, based on the MTASim approach, replicates various traffic situations within the network, enabling the evaluation of mobility policies and city changes. Subjective KPIs are calculated to assess factors such as air pollution, congestion, and cyclist safety. The recommendation engine, implemented using Dexi, helps decision-makers choose the most suitable policy based on subjective preferences. ML techniques, implemented using Orange, evaluate policy quality and propose mobility policies based on previously simulated scenarios.

Overall, the Urbanite project contributes to the development of smarter and more sustainable cities by leveraging technology, data, and citizen engagement. The proposed software framework enhances decision-making processes in urban planning, allowing for the testing, evaluation, and recommendation of mobility policies. By addressing the specific challenges faced by each city and fostering collaboration among stakeholders, the project aims to accelerate the transformation of European cities into smart and future-ready urban centres.

Acknowledgement

This work is part of a project that has received funding from the European Union's Horizon 2020 research and innova-

tion programme under grant agreement No. 870338. The authors also acknowledge the financial support from the Slovenian Research Agency (research core funding No. P2-0209).

We express our gratitude to the collaborators involved in the Urbanite project, namely Arianna Villari, Dino Alessi, Eduardo Green, Francesco Martella, Giovanni Parrino, Giuseppe Ciulla, Heli Ponto, Ignacio Olabarrieta, Isabel Matranga, Iñaki Etxaniz, Jorge Garcia, Julia Jansen, Keye Wester, Maitena Ilardia, Maria Fazio, Maria Llambrich, Mario Colosi, Marit Hoefsloot, María José López, Massimo Villari, Nathalie van Loon, Sergio Campos, Tatiana Bartolomé, Thomas van Dijk, Torben Jastrow, Sonia Bilbao. Additionally, we would like to extend our appreciation to the individuals from the Department of Intelligent Systems at Jozef Stefan Institute for their valuable cooperation and support.

References

[1] United Nations Department of Economic and Social Affairs. *World Population Prospects 2019*. New York: United Nations, 2019.

[2] Dameri, R. P. (2013). Searching for smart city definition: A comprehensive proposal. *International Journal of Computers and Technology*, *11*(5), 2544-2551.

[3] Washburn, D., Sindhu, U., Balaouras, S., Dines, R. A., Hayes, N., Nelson, L. E. (2009). Helping CIOs understand "smart city" initiatives. *Growth*, *17*(2), 1-17.

[4] Anthopoulos, L. G. (2015). Understanding the smart city domain: A literature review. In *Transforming City Governments for Successful Smart Cities* (pp. 9-21). Springer.

[5] Bibri, S. E., Krogstie, J. (2017). Smart sustainable cities of the future: An extensive interdisciplinary literature review. *Sustainable Cities and Society*, *31*, 183-212. doi:10.1016/j.scs.2017.02.016

[6] Neirotti, P., De Marco, A., Cagliano, A. C., Mangano, G., Scorrano, F. (2014). European Smart Cities: The Potential for Advancing Innovation and Sustainability. *Journal of Urban Technology*, *21*(1), 3-17. doi:10.1080/10630732.2013.876689

[7] Kitchin, R., Coletta, C., Evans, L., Heaphy, L., MacDonncha, D. (2018). Smart Cities in Europe: Enabling Factors and Success Factors in Smart City Development. *Regional Studies, Regional Science*, *5*(1), 136-138. doi:10.1080/21681376.2017.1390684

[8] Hollands, R. G. (2019). European Smart Cities: A Systematic Analysis of Smart City Initiatives. *Journal of Urban Technology*, *26*(4), 3-27. doi:10.1080/10630732.2018.1488562

[9] Deakin, M., Al Waer, H., and Newton, R. (2020). European Smart Cities: The Role of Policy in Shaping Smart Urban Futures. *Cities*, *98*, 102573. doi:10.1016/j.cities.2019.102573

[10] TECNALIA, *URBANITE website*, 2023. Available at: `https://www.example.com`.

[11] Sulajkovska, M., Smerkol, M., Dovgan. E., Gams M. (2021). Machine Learning-Based Approach for Estimating the Quality of Mobility Policies. *Proceedings of the 24th International Multiconference Information Society URBANITE Workshop 2021*

[12] Smerkol, M, Sulajkovska, M., Dovgan. E., Gams M. (2021). Machine Learning-Based Approach for Estimating the Quality of Mobility Policies. *Proceedings of the 24th International Multiconference Information Society URBANITE Workshop 2021*

[13] Dovgan. E., Smerkol, M, Sulajkovska, M., Gams M. (2021). Supporting Decision-Making in the Urban Mobility Policy Making. *Proceedings of the 24th International Multiconference Information Society URBANITE Workshop 2021*

[14] Sulajkovska, M., Smerkol, M., Gams M. (2022). Urban Mobility Policy Proposal Using Machine-Learning Techniques. *Proceedings of the 25th International Multiconference Information Society*

[15] Bohanec, M. (n.d.). Dexi: A Decision Support System. Retrieved June 26, 2023, from `https://kt.ijs.si/MarkoBohanec/dexi.html`

[16] Orange. (2023). *Orange - Data Mining Fruitful and Fun*. Retrieved June 26, 2023, from `https://orange.biolab.si/`

[17] OpenStreetMap. (2004). *OpenStreetMap*. Retrieved June 26, 2023, from `https://www.openstreetmap.org/`

A Hybrid Modelling Framework for E-Commerce Supply Chain Simulation: Complex Adaptive Systems Perspective

Alejandro Nila Luévano[1], Aida Huerta Barrientos[1], Nicolás Kemper Valverde[2]
[1] Faculty of Engineering, National Autonomous University of Mexico, CDMX, México.
[2] ICAT, National Autonomous University of Mexico, CDMX, México.
E-mail: nilaalejandro@gmail.com, aida.huerta@comunidad.unam.mx, kemper@unam.mx

Keywords: hybrid simulation, complex adaptative system hybrid simulation, supply chain hybrid simulation

Received: July 14, 2022

E-commerce emerged as consequence of electronic transactions developed on 60's, but real boom was observed during 90's along with Internet common use. Complexity sciences approach has several advantages for e-Commerce study. This study addresses the need for modelling and simulation (M&S) of e-commerce supply chain as complex adaptive system (CAS) but with a novel application in the field of hybrid M&S, integrating top-down and bottom-up approaches using synthetic microanalysis, to perform simulation experiments to find natural emergent properties at certain levels as result of the interactions between the constituent parts, so far lacking in the scientific literature. Although previous researchers conducted simulation studies into the e-commerce supply chain as CAS, they all focused on applying agent-based simulation approach only. First, we conduct the literature review on main features of CAS, M&S of CAS as well as the e-commerce supply chain conceptualized as CAS and their modelling and simulation evolution. Second, we present a novel hybrid M&S methodology for integrating top-down and bottom-up approaches using synthetic microanalysis. Then, we applied the methodology to an omnichannel retail business case study. Finally, our concluding remark and future work are drawn. The novel methodology proved to be useful for anticipate business decisions on e-commerce supply chain.

Povzetek: Ta študija s pomočjo hibridne metodologije obravnava potrebo po modeliranju in simulaciji e-poslovne dobavne verige kot kompleksnega prilagodljivega sistema.

1 Introduction

Seminal contributions have been made on Complex Adaptive Systems (CAS) by pioneer researchers as Buckley [1], Holland [2] and Gell-Mann [3]. CAS term was first introduced by Walter Buckley in 1968 to describe a system whose elements or components are directly or indirectly related in a causal network, such that at least some of the components are related to others at any one time, the interrelations may be mutual or unidirectional, linear, non-linear or intermittent, and varying in degrees of causal efficacy or priority. In [4] three characteristics of CAS are proposed: *evolution, aggregate behavior,* and *anticipation.* After that, in [5] one extra characteristic of CAS is added: *hierarchical arrangements of boundaries and signals.* It is important to note that CAS form and use internal models to anticipate the future, basing current actions on expected outcomes, in fact it is the attribute that distinguishes CAS from other kinds of complex systems and makes the emergent behavior of CAS intricate and difficult to understand [6]. Previous studies presented on [7, 8] have almost exclusively focused on one M&S approach to study CAS. However, hybrid modelling which extends the M&S discipline by combining research approaches, methods, techniques, and tools from across disciplines [9] to one or more stages of a simulation study, still is limited. In the literature review we found contributions only in the domain of defense [11].

On the other hand, e-commerce emerged as a consequence of electronic transactions developed on 60's, but real boom was observed during 90's along with Internet common use. Previously, catalogue sales already offered products and services on home delivery, but every transaction must be done by a presential visit or at least using a printed document. Complexity sciences approach has several advantages for e-commerce study, beside the understanding of patterns and behavior of main actors from the point of view of companies, it is also useful for customer understanding and specific patterns that will follow to choose their products or services offered online.

Supply chain has been conceptualized as a CAS by many authors [11 – 17]. The traditional tools used in the study of the supply chain as CAS are the following: System Dynamics (SD), Agent Based Modelling Simulation (ABMS), Dynamic Systems Theory, Observation Data-Based Models, Dynamic Networks, Ordinary Differential equations (ODE), Difference equations and Partial Differential equations (PDE), Cellular Automata, Evolutionary Game Theory (EGT), and Fractional Calculus. In the field of simulation, the implementation of models based on top-down approach, which is used to provide holistic perspective by synthetic microanalysis, the SD is preferred. While implementing models based on bottom-up approach, that is a deductive

perspective, the Discrete-Event Simulation (DES) and ABMS. Combining two or more of the following methods: SD, DES, and ABMS, has experienced near-exponential growth in popularity in past two decades [18]. In the domain of supply chain, but not from the CAS perspective, we found contributions by [19, 20] using two or more simulation approaches. More recently, in [21] DES and ABMS are combined with heuristics to govern train movements destination selection, incorporating an ensemble of simulation runs. In [22] SD, DES and ABMS are applied to an aerospace manufacturer's real case to assess the sustainability performance of alternative supply chain.

This study addresses the need for modelling and simulation of e-commerce supply chain as CAS but with a novel application in the field of hybrid M&S, integrating top-down and bottom-up approaches using synthetic microanalysis, to perform simulation experiments to find natural emergent properties at certain levels as result of the interactions between the constituent parts, so far lacking in the scientific literature. Although previous researchers conducted simulation studies into the e-commerce supply chain as CAS, they all focused on applying agent-based simulation approach only.

Supply chain problem has been addressed with several approaches, that's why in [15] and [17] stands out the need for novel methods that includes tools based on complex systems approach. The dynamic of the market due to growth in internet use and changes of purchasing habits have led the supply chain professionals to find better methods. According to [32] global e- commerce went from 15% of total retail sales in 2019 to 21% and now is estimated 22% so to be focused on interactions, dynamics and emergent patterns of the systems could bring us better solutions in a different point of view. Static supply network designs can help with localization problems but to solve everyday companies' issues is getting more complex and in consequence more specialized to solve.

The rest of the paper is organized as follows: Section 2 describes the literature review on main features of CAS, M&S of CAS as well as the e-commerce supply chain conceptualized as CAS. Section 3 present a novel hybrid M&S methodology for integrating top-down and bottom-up approaches using synthetic microanalysis, of an e-commerce supply chain to perform simulation experiments to find natural emergent properties at certain levels, as result of the interactions between the constituent parts. Section 4 presents an omnichannel retail business case study. Concluding remark and future work are presented in Section 5.

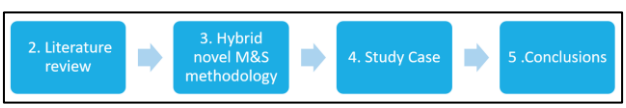


Diagram 1: Paper organization

2 Literature review

2.1 Complex adaptive systems

Complex adaptive systems (CAS) term was first introduced by Walter Buckley in 1968. In his publication named *Society as a complex adaptative system*. Buckley [1] defined a system in general as a complex of elements or components directly or indirectly related in a causal network, such that at least some of the components are related to some others at any one time, the interrelations may be mutual or unidirectional, linear, non-linear, and varying in degrees of causal efficacy or priority. As Buckley [1] added, persistence or continuity of an adaptive system may require, as a necessary condition, change in its structure, the degree of change being a complex function of the internal state of the system, the state of its relevant environment, and the nature of the interchange between the two. From the cybernetic point of view [23], CAS are complex effectors organized and self-regulated to subtract themselves or one of their effects, within certain limits, from contingency, from increased entropy, or from both [24]. The cybernetic perspective of control or self-regulation of adaptive systems emphasizes the crucial role of deviation, seen in both negative and positive aspects: on the negative side, certain kinds of deviations of aspects of the system from its given structural state may be seen as *negative feedback*, while on the positive side, it is necessary the deviation or more generally, *variety* – in providing a pool of potential new transformations of process or structure that the adaptive systems might adopt in responding to goal-mismatch [1]. Another pioneer of CAS research was Prof. John Henry Holland whose seminal work on adaptation in natural and artificial systems [2] leading to the creation of genetic algorithms and eventually the fields of evolutionary computation [4] and learning classifier systems [25].

The models to understand CAS consider the rule-based structure that lets the evolutionary procedures that enable the system to adapt to its surroundings. As noted by [6], most rules can be parsed into simple *condition/action* rules:

IF [condition true]
then execute
[action]

On the other hand, the following are recognized in [3] as the general characteristics of a CAS, mainly related to data and information:

a) Its experience can be thought of as a set of input and output data, with the inputs often including system behavior and the outputs often including effects on the system.
b) The system identifies perceived regularities of certain kinds on the experience, even though sometimes regularities of those kinds are random features misidentified as regularities.

c) The perceived regularities are compressed into schema. Each schema provides, in its own way, some combination of description, prediction, and prescriptions for action.
d) The results obtained by a schema in the real world then feedback to affect its standing with respect to the other schemata with which it is in competition.

2.2 Modelling and simulation of CAS

In recent decades, CAS are at the heart of important contemporary problems, so the research methods and tools need to be modified with an emphasis on the role of computer-based models to increase our understanding of CAS [26]. In this direction, the purpose of modelling and simulation (M&S) of CAS is to find the macro and micro mechanisms behind the *evolution*, *aggregate behavior*, *anticipation and* hierarchical *arrangements of boundaries and signals* of CAS.

The top-down approach is used to provide holistic perspective by synthetic microanalysis, in which experiments are performed to find natural emergent properties and delineate macroscopic phenomena with systemic concepts [27]. Once the system structures have been observed based on top-down perspective and the bottom has been reached, we need to analyze them in terms of the laws on their constituent parts combined with suitable idealization and approximation at the micro level, at this point, the bottom-up approach is a deductive perspective, in which experiments are performed to find natural emergent properties at certain levels as a result of the interactions between the constituent parts [27].

2.3 E-commerce supply chain conceptualized as CAS

In the literature, supply chain has been conceptualized as a CAS by many authors [11, 15, 17, 28-36]. In this direction, supply chain is constituted by large number of agents that interact in a non-linear way, could evolve, learn, and be resilient to their environment. Also, its structure and collaboration mechanism evolve over time [14]. Considering the cybernetics point of view, Figure 1 shows the receptors or sensors, the processor or homeostatic controller and the effectors of a supply chain conceptualized as CAS.

Following the main characteristics of CAS, the constituent parts of a supply chain could be facilities, people, equipment, and communications involved in a supply chain or logistics network. Supply chain networks presents nonlinear behavior on products distribution, where the number of vehicles is not proportional to the number of products because of product volume, weight and even restrictions for destination traveling, that's why behavior would not be modeled with a first order equation, without assuming or forgetting essential system properties.

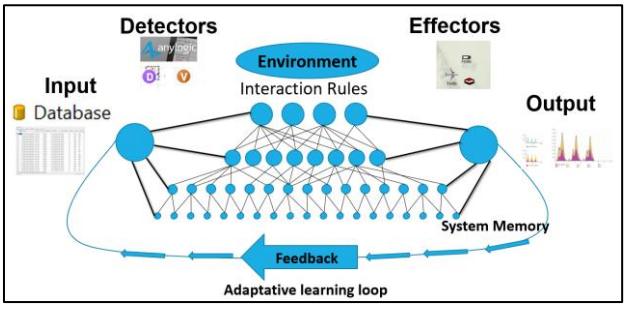


Figure 1: Supply chain conceptualized as CAS.

On supply chain management and in business in general is normal to concentrate investment only in certain separated areas and outcome it is that constituent's effort does not contribute adequately on system desired results. Other important property to address is that complex systems are not controlled at a central level. Most of the time, system constituents do only have communication with nearest or adjacent peers, this communication and interactions trigger emergent behavior at system level. Supply chain systems and logistics systems had been trying for years to create a single communication and management system using ERP (enterprise resource planning), this central software application have proved to help on complexity reduction, but interaction and decisions made by single constituents are always present unleashing emerging patterns that must be taken into consideration. That is why supply chain elements or agents that compose supply chain would not always check every possibility to make every decision. In a better way, every agent evaluates the result of certain action and the repeat successful one with the objective to evolve or improve its complex system.

In order to identify the recent trends in the literature on the application of the CAS approach in the study of e-commerce supply chain, a bibliometric analysis was carried out using the VOSviewer™ software, which is a software tool focused on the distance-based visualization of bibliometric networks. In the visualizations provided by VOSviewer™ software, the distance between two nodes indicates the relatedness of the nodes [38].

We search on SCOPUS database the relevant papers with keywords *e-commerce* AND *complex adaptive systems*. From the results, we observed that in general, CAS approach is related to supply chain and e-commerce by their modeling tools or analysis algorithms. Figure 2 shows the interrelations found in the literature of e-commerce and CAS and agent-based modelling. Figure 3 shows the interrelations between supply chain management and e-commerce. These two subjects are always connected because of their symbiotic nature but it can be highlighted that complexity properties like self-organization along with complexity solutions including evolution models come in to light.

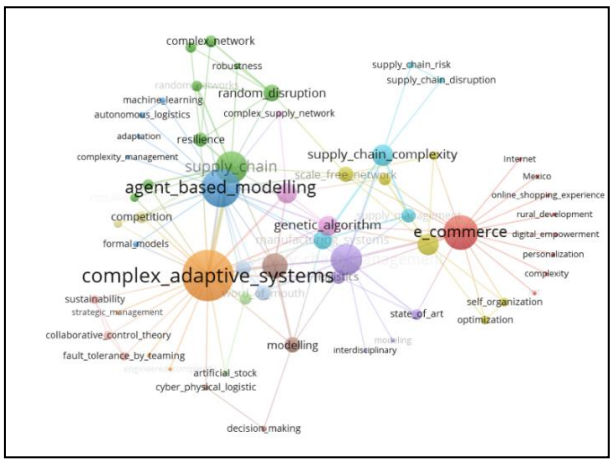


Figure 2: VOSviewer™ software visualization of the keywords network.

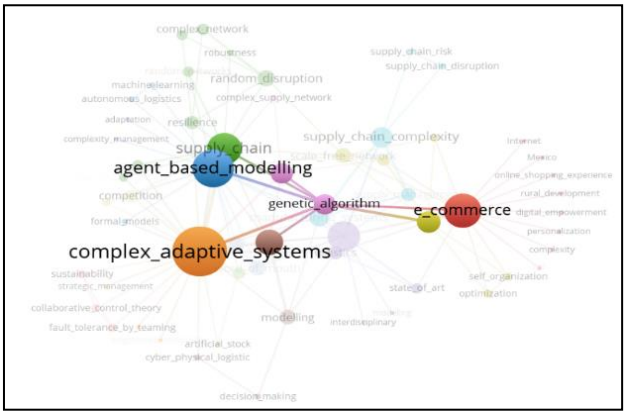


Figure 4: VOSviewer™ software visualization of the keywords network, showing the interactions among agent-based modelling, e-commerce and complex adaptive systems.

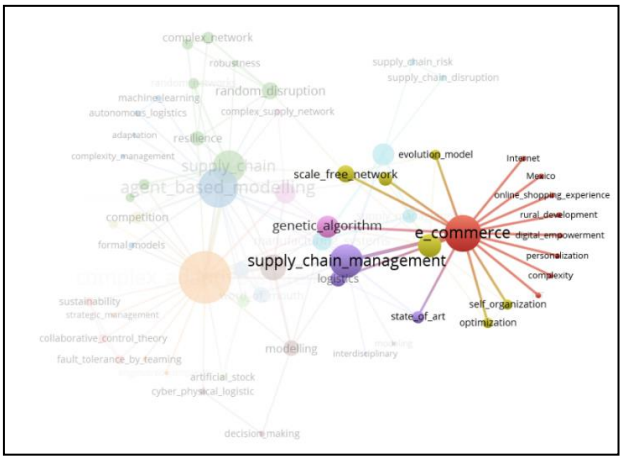


Figure 3: VOSviewer™ software visualization of the keywords network, showing the interactions between e-commerce and supply chain management.

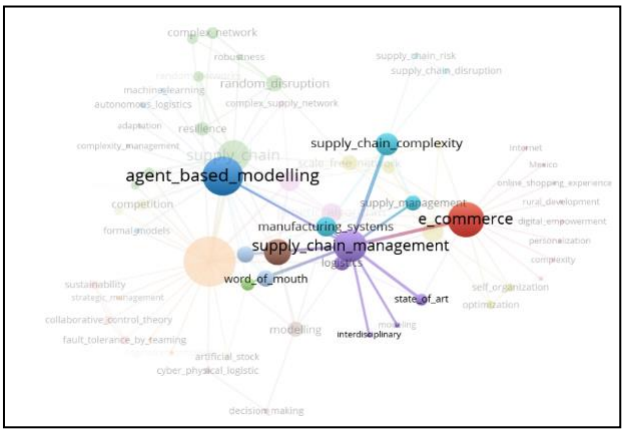


Figure 5: VOSviewer™ software visualization of the keywords network, showing the interactions among agent-based modelling, supply chain complexity, e-commerce and supply chain management.

Figure 4 indicates that in recent years CAS approach on e-commerce supply chain using agent-based models is a trend to confront everyday challenges of modeling and simulating e-commerce process. The topic supply chain management is strongly related to e-commerce of manufacturing systems, agent-based modelling also is used and selected as a tool to continue further investigation. Additionally, Figure 5 demonstrate that interdisciplinary is also included in this concept for multiple solutions.

2.4 Modelling and simulation (M&S) of supply chain as CAS

Supply chain problems have been addressed and confronted by different methods, but we put special attention on recognizing enough ones who helped us to maintain. The following tools act as an alternative to deterministic optimization static models to modelling and simulation of supply chain as CAS [15]:

- Simulation (SD and ABMS)
- Dynamic systems theory
- Observation data-based models
- Dynamic networks

There are other methods for logistics systems modelling using a complex system approach [35], analytic method proposed by [38], game theory presented in the work of [39], dynamic systems proposed in the work of [40]. On the same way, managers look for new ideas for developing collaborated and synchronized network

structures, these networks must adapt to flexible and dynamic environments [14].

3 A novel hybrid M&S methodology based on synthetic microanalysis

In this section, we present a novel methodology for integrating top-down and bottom-up approaches using synthetic microanalysis, of an e-commerce supply chain to perform simulation experiments to find natural emergent properties at certain levels as result of the interactions between the constituent parts. The conceptual modeling and communication hybrid simulation phases are based on [41] using complexity sciences tools. Our main contribution is to adapt recognized knowledge for a complex adaptative e-commerce supply chain. On every phase, advantages of hybrid M&S and CAS approach are described and highlighted.

3.1 Phase 1: conceptual modelling

On this phase, since the hybrid simulation models proposed for supply chain were conceptualized as a CAS, the main objective will be to find emerging patterns due to interactions among the components of the system. Assuming from the beginning that supply chain can be analyzed as a CAS, we took advantage of an additional complexity science theorical framework proposed by [27], but on CAS, detectors and effectors are also taken into consideration along with system interactions. Same as typical system modelling, inputs are considered to obtain certain outputs. Feedback adaptative learning loop and a system learning memory is also included. The conceptual diagram of CAS adapted from [42] is shown on Figure 6. Unlike *black box* system diagrams, interaction rules play a relevant role on CAS modelling, that's why we looked for a software with enough capabilities. After the analysis of software simulation used to modelling and simulation of CAS, we selected Anylogic™ software. Anylogic™ is one of the simulation leading platforms found in the market, native Anylogic™ software's features enables users to integrate DES, ABMS and SD on a single integrated model that can dynamically read and write data to spreadsheets or databases during a simulation run and is capable to develop spatially explicit models integrating GIS functionality [43]. Then, the possibility to build hybrid simulation models on Anylogic™ is not only possible, but it is also many times implied on software's applications. In the e-commerce supply chain case, we took advantage of the two mentioned features: DES simulations, to incorporate time processes and indoor facilities behavior, and ABMS, for interactions and flows between facilities, including transportation network and vehicles. To give a context of what the model includes, CAS components of our proposed modelling are described below:

Input is composed by typical elements incorporated on e-commerce supply chain analysis and modelling:

a) Historical or forecasted demand, taken from both Enterprise Resource Systems (ERP) databases, or any other business information system. Different scenarios might be downloaded and configurated depending on required business answers.
b) Objective service level, included on model data base, or stored on parameters to compare real life indicators.
c) Product characteristics, tagged on every demand line of data base, useful information to interact on discrete indoor processes, or to indicate way of traveling along agent-based network paths.
d) Number of facilities and location, shown on a Geographic Information System (GIS).
e) Processing and traveling time, based on productivity and resources availability.

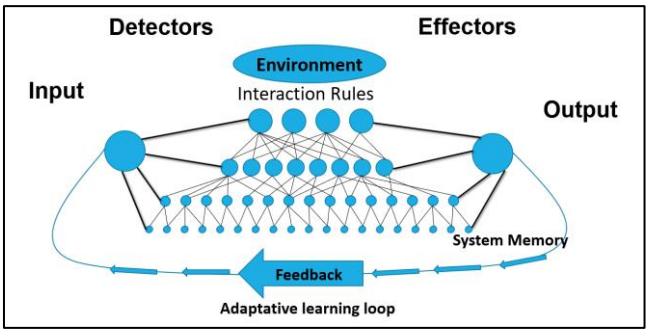


Figure 6: Conceptual CAS diagram adapted for this methodology, adapted from McElroy (2000).

Detectors

a) The counter variables.
b) The statistic Blocks.
c) Time DataSet blocks and plots.
d) Time histogram blocks and plots effectors.

Effectors

a) Agent indoor facilities with DES processes.
b) Agent transportation network.

Output

a) End to end product system life.
b) Specific time (process, travel, waiting) per agents or facility.
c) Backlog indicators.

3.2 Phase 2: simulation modelling interaction rules

In the simulation models, interaction rules are governed by programmed business rules and a predesigned a supply network. Every agent is created based on historical demand, from a source block and then travel to different facilities. On every facility, demand agents could change their attributes or lifetime due to DES processes, and then wait for a vehicle that match the next step or destination. Hybrid simulation models are integrated by agents as following:

Main agent

It hosts every other agent on a GIS, it has also the duty to fill every other agent´s attributes at the start of the simulation. Here, every demand agent is created and then canceled after traveling along supply chain system.

Demand agent

It travels along the system following business rules. It is created with attributes or parameters that will be useful for the software to indicate logistics paths using *tags* to accomplish needed processes. Some support parameters are incorporated to accumulate attributes due to different situations.

Facility agent

Every facility has two purposes:

a) Interaction with other agents.
b) Hosting supply chain processes.

Depending on the complexity of operations, these processes could be as simple as a delay of time followed by a queue, or even a DES complicated simulation that support resources and schedules according business capacity communicating with other agent interactions. Depending on business supply chain, different kinds of facilities could be modelled, in our case studies, facilities were commonly stores, warehouses and x-docks but this is not limited. Depending on simulation designed necessities, interactions with factories or suppliers might be taken also into consideration.

Transportation agent

Vehicles are created also at the beginning of the simulation based on model data base. This option gives you the opportunity to define not only origin and destination of vehicle, parameters of capacity, travel time and schedule can be also included. To obtain specific output, agent-based simulation gave the chance to communicate every state of every vehicle to other agents and then decide when to load or unload products at the start a trip fulfilling business objectives. Vehicles do not only travel based on a defined time on data base, but also could be visible on Main GIS map. GIS map road distance information was useful to simulate time travel and gives to the user a perspective of how fleet is moving along the run.

Agent waiting areas

Two waiting areas, one for products waiting for a vehicle to arrive in a facility and then travel to a specific destination, and inside vehicle agents, where products remain in loaded, unloaded or in transit status between a stop or travel destination.

Messenger agents

Created at a specific time recorded in data base, their purpose is to change values on process parameters or sending messages to some agent following programmed interactions. For example, on facilities capacity is adjusted whether on productivity or on number of resources.

Even when processes or individual programming could be very simple, the combination of several number of agents that are born at the beginning of simulation along with interactions during it, result in complex behavior. The relevance of interactions and understanding of results was the next step of our proposal.

3.3 Phase 3: model communication

Once DES processes agents are created, communication phase includes programming to link variables, identify interactions and then to execute model scenarios. Here is where we identified the most relevant advantage of analyzing and modelling supply chain with a complex adaptative system approach. To be aware of dynamics understanding, and analyze outputs as emerging system patterns, upgraded the results of our methodology. On the same way, synthetic micro analysis gave us the opportunity to understand the most micro level interactions that drive to model better business decisions.

Linking variables and identified interactions

Variables and interactions are linked mainly by demand product flows. These product flows are often a connection between DES processes and using “enters” or “exits” between facilities or vehicles living on a GIS MAP. Natural supply chain designed flows are followed to simulate real life behavior.

Database driven decisions

In real life and in virtual simulation models, products travel along supply chain, according to business defined rules. In real life, these decisions are chosen by systems, or human beings, but in our models, they are previously stored on a data base. Whether the product must accomplish certain process or to travel to another facility, a database query is executed to simulate real life decisions and interactions.

The following diagram explains how demand agents flow through the simulation model.

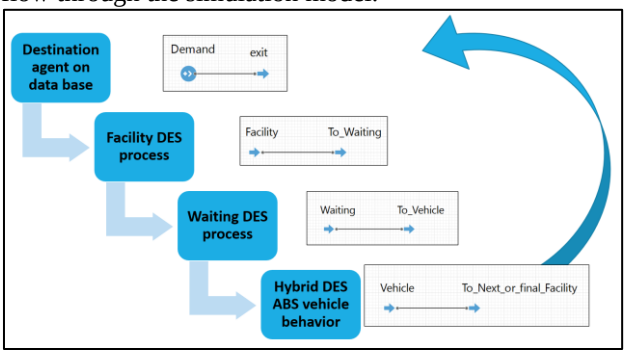


Figure 7: Simulation flow diagram Source.

a) First, the demand is generated, where a source block is configurated with historical or forecasted demand. By using real data, dynamic and nonlinear behavior is entered into the system. Demand agents are created with input parameters that contain characteristics of at least origin and destination previously stored on database, these parameters are also used as tags for product identification and

decisions useful to consider entering or exiting processes between facilities or vehicles. During simulation run, counting variables and datasets are collected and refreshed at a specific time to understand input demand behavior.

b) Second, the demand is communicated to the facilities depending on product requirements. Time on facility, depends on utilization of machines or human resources mayorly. Statistics of every process are collected and plot to understand where capacity is limited and where the agents spend more time. Facilities could be very simple, or also enabled with several complicated processes, for example Figure 8 shows a GIS map containing several facilities that host simple scheduled driven by resources DES processes. Capacity and resources could be static or dynamic, if it is needed, at any time of simulation run, a messenger agent could enter the facility to update conditions. The result of interactions and behavior of these changes, then is observed as emergent patterns of supply chain behavior.

c) Third, the waiting processes of facilities and vehicles are modelled. Their purpose is not only to host a staying process for products and to accumulate time between entering a vehicle, but these processes also allow communication between facilities and vehicle agents. Communication process on simulation could be complicated, the problem is to keep products on a specific destination process until a vehicle has arrived ready to load or unload products. To accomplish this task, an "id" of arrived vehicles is kept on a data collection array and then products are allowed to enter mentioned vehicle that share the next agent destination and at the same time, ensuring choosing the right transportation type with enough left capacity.

d) Finally, products travel from one to another facility or destination. According to waiting processes, every product enters on a specific vehicle depending on the destination. We named this DES processes vehicle compartments. On these agents, processes of loading and waiting inside a transportation unit are simulated. Then, it is possible to measure used capacity, and decide if the vehicle is ready to go using two criteria. The first criteria to start a trip is leaded by programmed schedules fed by departure parameters. Second criteria, to verify if transportation agent is enough loaded according to business rules and physical capacity. Vehicles are programmed as independent agents, so they born with parameters including origin and destination, together with variables recorded in accordance to departure, travel time and capacity limitations.

On every step, outputs are programmed considering special attention on simulation time and different products behavior. Normally output values are shown in dynamic plots during simulation. Now that CAS behavior is programmed on simulation software, adaptative learning loop depends on user experience and iterative *what-if* scenarios. Every iteration stored output on computer will conform our system memory, and retro alimentation will be sustained proposing different agent network paths or capacity.

Model execution and data exchange

Once model and interactions are already programmed execution is very simple. After verifying that every parameter and variable is loaded on model database, next step is to execute the simulation and wait for the result. After every iteration recorded statistics and indicator calculations, are exported to a consolidated *Microsoft Excel spreadsheet*. The relevance of this step is not only to obtain results and record them, but objective is also to identify the influence of every interaction at micro level that are shown at macro level as emergent patterns of the complex adaptative system, in our case these emergent patterns coincide with supply chain performance indicators.

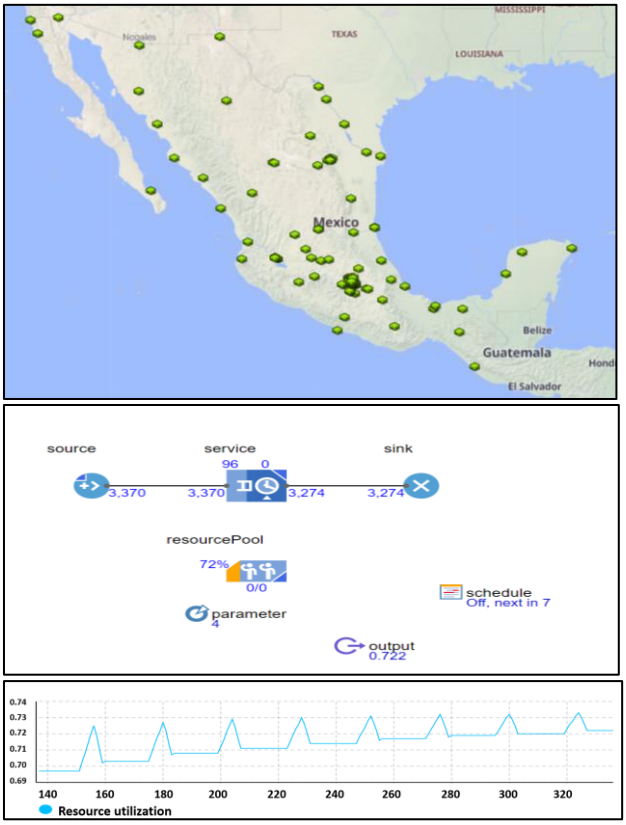


Figure 8: GIS Map containing agents, DES processes and plot outputs.

Then, improvement of supply chain system is implied on the influence of products flows, facilities processes and vehicles interacting with each other. Talking about our supply chain models, every facility will exchange data with every vehicle not only about the status and capacity, but it will also share information about demand dynamic behavior. The result will impact utilization of resources

and bottlenecks at micro level, and the result of end-to-end measured time deliver, at macro level.

Evaluation of outputs

The simulation software offers several options to obtain results and outputs. Statistics and datasets blocks, along with time and histogram plots were successful to understand complex behavior and led us to different scenarios. From the beginning of the run, we take information from data base, placing historical or forecasted demand. The first output is therefore the behavior of demand, where we can understand seasonality and behavior at macro level. These plots could be included by product or by any other interesting demand identification. The most common, apart from products, are related to logistics origin-destination area and required service. Then, at main agent it is possible to observe demand graphics and interesting statistics form its behavior. During simulation runs, demand behavior is compared with accumulated capacity and accumulated statistics help us to understand system performance. Scheduled resources work shifts utilization is also measured to understand everyday dynamics.

Traditional analysis, assuming a typical black box system, will lead us to run several scenarios looking for answers only at macro level, but CAS approach along with synthetic micro analysis tells us that these answers at macro level must be a result of interactions at micro ones. Then, even when performance is enough for accomplishing demand requirements at macro level, interactions at micro mechanisms should be understood to get a better result. On supply chain, installed capacity relates directly to investments and in many times, it is difficult to balance the right place and more difficult to do it at the right time, resulting on money wasting. In our simulation projects, hybrid simulation opens the opportunity to integrate DES processes into agents, and to measure their performance at micro level.

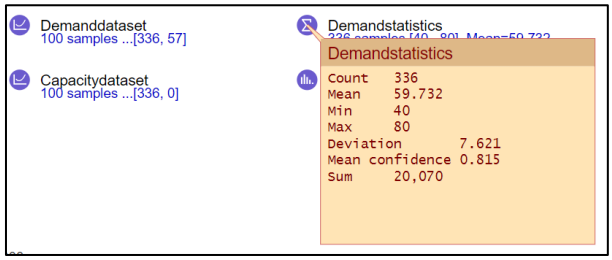


Figure 9: Example of simulation input analysis.

The performance of every agent is collected and then analyzed to understand impact on the whole system. One of the top interests of our simulations is service time, which depends in part on the number of resources and their productivity related to real processes, that's why backlog is measured along every facility to identify bottle necks and prove different scenarios. We measure time on a facility, but also the time between every process. By recording on every hour, the number of agents inside simulation queues, we obtained an average indicator of backlog and the period of recovery. By comparing backlog data with capacity bottlenecks, facility capability can be understood in two dimensions, first to dimension about space capacity and second to ponder if recovery time is according to business service requirements. Figure 11 shows backlog behavior and time recovery due to limited wok shift resources capacity. Simulation results gave us the opportunity to observe and measure impact of limited capacity and its dynamics depending on demand flows. Remembering that the supply chain system could be integrated by several facilities and every piece of data is analyzed after exporting information an integrated *Microsoft Excel Spread Sheet*. The simulation results are organized by flow volume, average backlog, and time of recovery to identify the most relevant interactions.

By collecting the accumulated end to end time of every product on main agent and understanding its causes due to micro level interactions, *what-if* mentioned scenarios are proposed changing parameters. Methodology enables to understand the limitations of the designed supply chain and apply better options to reach business objectives, without thinking about major investments.

Figure 10: Example of simulation outputs.

4 An omnichannel retail business case study

This section is focused on simulation challenges of an omnichannel retail business case study. This section will be useful for the reader to recognize what is necessary to collect, understand and program to deal with a real-life problem.

Omnichannel retail business

Omnichannel problem includes companies who born operating *brick and mortar* stores, where consumers physically visit the place, and purchase their products. Due to customer behavior and growth in Internet use, companies also started an important presence on e-commerce operations. In recent times, e-commerce marketplaces are also opening physical stores, unless this is not ordinary. Because of their nature, both *bricks to clicks* and *clicks to bricks* companies, offer a hybrid way of delivery and shopping experience.

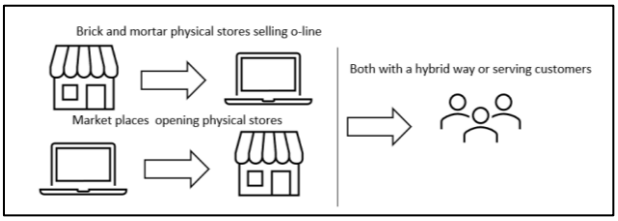


Figure 11: *Brick to clicks* and *clicks to bricks* ways.

4.1 Phase 1: conceptual modelling

In the omnichannel case study, customers are capable to acquire products in different ways depending on their necessities listed below.

a) On store traditional buying, were the customer walk around the place to acquire their products.
b) Online buying, where customer acquires products by the company web page.
c) A combination of previous, where customer could choose to pick up products on store if they are available. Availability could be nearby or foreign, and in both cases, customers could buy and receive product on store or direct to their location.

Depending on customer needs and location there are also different ways of home delivery.

a) The customer is available to receive product in a period of certain hours and the chosen product is available on a store or warehouse nearby.
b) The product could be delivered by the company x-dock in the same city.
c) The product needs to follow a network path between stores and x-docks to find its destination.
d) At an any network node, the product is delivered to a third-party logistics supplier.

4.2 Phase 2: simulation modelling

All of the options indicated before were programed on a hybrid simulation model using historical data as an input looking for improvements by running several capacity scenarios. Before starting to describe simulation inputs and outputs, it is important to remember the relevance and advantages of our methodology, that not only implies to program a hybrid model. It is also important to recognize how these steps can be useful to describe a real life complex adaptative e-commerce supply chain system. CAS are recognized mainly by their nonlinear behavior and a large number or constituents interacting at the same time. In the case study, the constituents were modeled as agents and their interactions using DES processes, because of their nature, agents could communicate with each other increasing complexity of the system. Then stores, x-docks, warehouse and transport assets became the agents and at the same time they were incorporated with historical real measured capacity and productivity on a GIS map location. Input of real demand data, followed by programmed agent business rules, resulted on hybrid simulation outputs emerging patters.

Store agents include DES processes for shelf product picking, customer in-store pick up area, product packaging, for delivery or transfer process, and last a decision block to send product direct to customer nearby or foreign employing a third-party supplier. Decision depends on the tag or parameter previously input on database. Products that must be delivered to another store by a company x-dock, are sent to the waiting area agent, where they will stay until a transportation asset has arrived. At store microlevel, relevant indicators are saved on statistics blocks, the most important were sum of products, queues backlog, resources utilization and finally whole time on store. With the analysis of this information, processes improvements were simulated and implemented, for example, by running different scenarios changing number of human resources working with an average productivity, benefits of service time were found at micro level. We validated those decisions about incrementing crew or working hours on certain season or peak of demand, could be taken based on model runs. With the aid of messenger *agents*, capacity or parameters were changed at any model time looking for better results. Results were not only useful at micro store level, total time macro level indicators were also measured and understood to validate business objectives.

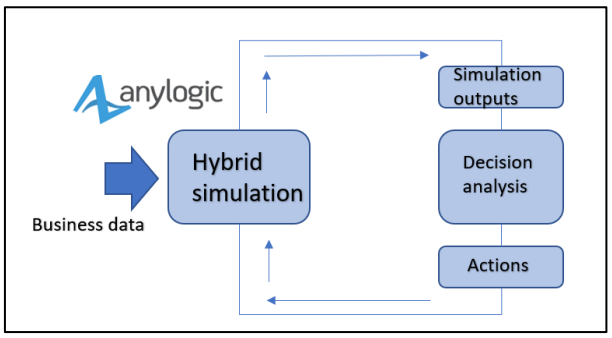


Figure 12: Simulation model of omnichannel case study.

4.3 Phase 3: model communication

Stores face many other challenges on real operations, main is that they are not designed to operate an outbound flow other than what is taken form the shelves aisles by customers. Taking advantage of agent's communication, schedules about receiving product from storage warehouses and sending product direct to customers or x-dock were also simulated. On store waiting area agent, products are hold until the scheduled transport arrive and take them to the next step of simulation. Same as in reality, transportation arrive to the store and report its arrival, in our simulation model, transportation agent communicates an ID to the waiting area by printing it in a collection array. With the help of knowing ID vehicle, the decision of loading product can be made in two ways, the first is to be sure about vehicle left space capacity and the second is if destination is in conditions to receive more products. That's how on this stage, communication between transport vehicles and facilities is emulated. In real system´s operations, it is common to hear about different kind of communications to arrange these operations, like phone calls and email messages, the result is a complex behavior induced by demand who result on micro level findings and in consequence supply chain indicators. Time on waiting area was an important indicator to be aware, but also its causes, especially the contained queue number of products waiting to be loaded on a transportation vehicle.

In certain circumstances, omnichannel problem delivery could be only about leaving the store, because it's common to use third party delivering suppliers to travel along countries and reach the *last mile* until customer receive their product. It could be even simpler when customers have the chance to pick their products on store, but store operations could be complex to study depending on the size of company, city or country. That's why in our study agent facilities were placed on a main agent that includes a GIS map, allowing users to incorporate several numbers of locations. For the reasons listed below we found GIS map useful in an e-commerce supply chain hybrid simulation:

a) Allows users to have a first look of their geographic coverage and number of nearby facilities. Being aware of where your facilities are located, can lead you to understand if they are placed in a correct way, even when this analysis might be inconclusive without an analytic exhaustive one, it is one of the best places to start.
b) GIS maps contain information not only about localization, in our case, Anylogic™ provides a native *OpenStreetmap* service, who enables vehicles to follow real-time downloaded road distances observed during simulation. Combination of localization and observation of vehicles on real time, can also be considered as an emergent pattern of the system seen at macro level.

Finally, GIS map was also useful to explain results and interactions between facilities to stakeholders. Before analyzing products after leaving the store, it is important to stop and interpretate what is causing bottlenecks and traduce it into real life strategies. Increasing the number of human resources might be a solution along with work hours variation, but also to compare productivity and facility space and lack of equipment. Once store variables are understood it is time for analyzing transportation logistics. There were different kind of vehicles depending on their vocation in real life on simulation model. Large capacity vehicles to deliver product from inventory warehouses to stores in synergy with omnichannel product, and local small capacity vehicles for moving product between stores to and from their local assigned x-dock. Distance and volume capacity during trip, was stored on statistics blocks for further analysis. For large capacity vehicles, every day scheduled departure time is programmed along with an average speed to simulate trips using real GIS map road distances. Then, estimated time of arrival of products can be calculated. In the case of small capacity products, delivery and collect *milk run* circuits were programed. At this step of simulation, we already accumulate the store and from store transportation time as a part of the total time of delivery, the rest obey to local x-dock, foreign x-dock, and their vehicle connections. Once products enter x-dock facilities, programmed decisions lead them direct to customer delivery or to travel to another facility, processing times of resources and productivity are also stored and measured along with on queues agent number content. From this point products can visit more than one x-dock and destination stores to finally reach their ending point. On last destination time along whole system is recorded and analyzed using a histogram plot. Something implied on this kind of problem simulation solution is to manage a great amount of agent interactions at the same time, this characteristic is one of the best reasons to choose this method when a supply chain is analyzed.

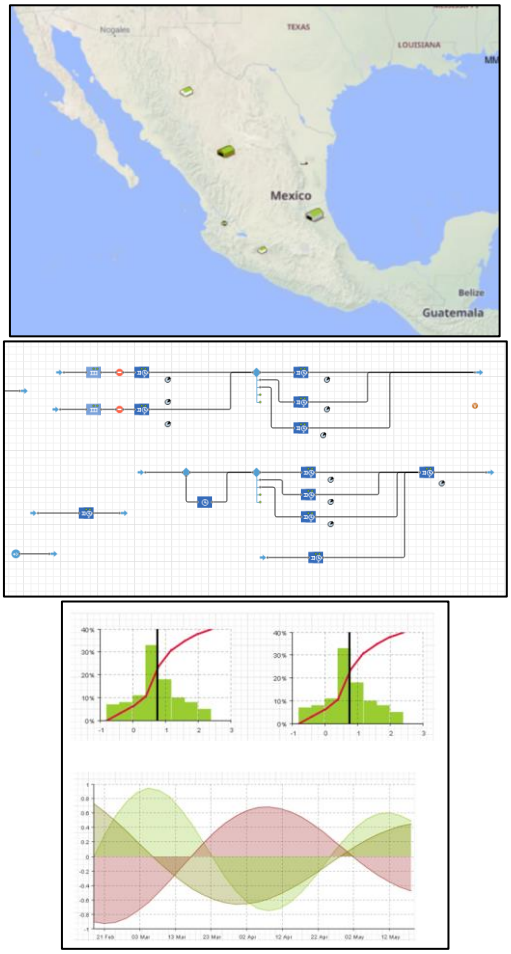


Figure 13: GIS Map containing agents, DES processes and plot outputs of omnichannel case study.

On this kind of modelling, number of facilities and assets is only limited on computer processing time, then by adding more database lines supply chain could be simulated at any scale without extra programming time. The result of thousands of products traveling between dozens of facilities and transportation assets interacting with each other is what finally emerge at macro level of the system. End to end time results of every interaction was also divided on categories and subcategories interesting for the business.

- Local deliveries including the ones from store and stores in the same city
- Foreign deliveries

Results show that even when facilities operate at their maximum capacity, accumulated time combinations (origin-destiny-type of service) delivery 100% on time objectives could be impossible to accomplish. Foreign times and accumulation on demand on particular season, are the main limitations of the supply chain system. Maximum limits of the system were not the only conclusion, more important was to identify offered services and estimated volume encapsulating the problem on certain geographic areas and then to decide the best impact business major investments. With the help of this novel methodology companies could simulate different scenarios to chose a better number of resources in a certain season and decide a better scenario. In our case study a retail company the chance to reduce the gap between on time deliveries in at least 3% by including third party suppliers on peak season.

Unlike traditional modelling of e-commerce supply chain where data is collected and then processed using a static linear programming model, hybrid simulation modelling assuming a CAS behavior, opens the door for a better system understanding and numberless real life dynamic scenarios. The understanding of microlevel interactions and their impact at macrolevel shown as emergent performance patters, can lead business to better strategies, and to predict behavior of their decisions. In early times, size of e-commerce supply chains has caused to operate and plan in a reductionist way, assuming that, solving the problem at micro level could reflect a deterministic upgrade on business macro indicators. Simulation modelling gives not only the *whys*, but also gives answers when might be the correct time to take important business decisions.

Benefits for e-commerce supply chain could be various, but here we resume the ones what we stand out the most:

a) Be aware of the limitations of your supply chain design before commit to business objectives.
b) Size impact of increasing capacity at facilities or transportation at macro level.
c) Understand not only what a better solution is, but also when might be a better moment to implement it.
d) Scale your analysis by running simulations of all your facilities at the same time.

5 Conclusions

The change in consumer behavior has caused supply chain professionals to face more complicated challenges. In the search of alternatives to understand and manage supply chain, the experience of stakeholders has been a great help for businesses. Even when experienced managers had looked for better ways to maintain service, it is difficult to find tools and frameworks to take advantage and test before implement. By understanding e-commerce supply chain as a CAS, multiples possibilities and solutions opens. CAS conceptualization facilitates to look for solutions about non-linear demand and agent interactions, instead of trying to implement average static solutions. After being awake of what CAS understanding could do for your business, the second step will be to find somewhere to experiment and prove interaction hypothesis. Simulation has proven to be a good platform to test nonlinear behavior and try to anticipate what may happen. Is very common to find discrete event simulation on e-commerce supply chain but normally model focusses on specific processes. By incorporating agent-based modelling, a tool born for complex system understanding, hybrid modelling enhances supply chain solutions proposals. On real life case studies, we achieved to point

out benefits of hybrid simulation, first by demonstrating that supply chain behavior can be imitated on a computer software without numberless spreadsheets, and second taking advantage of real-life e-commerce supply chain system agents (operations human resources), who now know they are capable to prove alternatives reducing the risk of wasting inversions or time testing on the go. Future work will address not only experienced bases solution but automated taking advantage of intelligent algorithms adapted for CAS and e-commerce supply chain, the objective will to find better business answers in less time.

References

[1] W. Buckley. Society as Complex Adaptive System, in W. Buckley (Ed.), *Modern Systems Research for the behavioral Scientist*. Chicago, IL: Publishing Company, 1968.

[2] J. H. Holland. Outline for a logical theory of adaptive systems, J ACM, vol. 9, no. 3, pp. 297–314, 1962.

[3] M. Gell-Mann. *The Quark and the Jaguar*. New York: W. H. Freeman, 1994.

[4] J. H. Holland. *Hidden order: How adaptation builds complexity*. New York: Addison-Wesley, 1995.

[5] J. H. Holland. *Signals and boundaries: building blocks for complex adaptive systems*. Cambridge Mass: The MIT Press, 2012.

[6] J. H. Holland. Complex adaptive systems, Daedalus, vol. 121, no. 1, pp.17–30, 1992, http://www.jstor.org/stable/20025416

[7] N. Boccara. *Modeling complex systems*. Berlín: Springer Publ, 2004.

[8] E. Ahmed, A. S. Elgazzar, A. S.Hegazi. An overview of Complex Adaptive Systems, 2005, arXiv:nlin/0506059v1

[9] A. Tolk, A. Harper, and N. Mustafee. Hybrid Models as Transdisciplinary Research Enablers, European Journal of Operational Research, vol. 291, no. 3, pp.1075-1090, 2021.

[10] J. Moffat, M. Bathe, L. Frewer. The hybrid war model: a complex adaptive model of complex urban conflict, Journal of Simulation, vol. 5, no. 1, pp. 58-68, 2011.

[11] A. Huerta -Barrientos and I. Flores de la Mota. Modeling the adoption of sustainable practices in the supply chain: a game theory approach, Journal of Advanced Management Science, vol. 5, no. 4, pp.250-254, 2017.

[12] M. Brandenburg, K. Govindan, J. Sarkis, S. Seuring. Quantitative models for sustainable chain management: developments and directions, European Journal of Operational Research, vol. 233, pp. 299 – 312, 2014.

[13] S. Seuring. A review of modelling approaches for sustainable supply chain management, Decision Support Systems, vol. 54, pp. 1513 – 1520, 2013.

[14] G. Li, P. Ji, L. Y. Sun, W. B. Lee. Modeling and simulation of supply network evolution based on complex adaptive system and fitness landscape, Computers & Industrial Engineering, vol. 56, no. 3, pp. 839 – 853, 2009.

[15] A. Surana, S. Kumara, M. Greaves, U. Nandini Raghavan. Supply-chain networks: a complex adaptive systems perspective, International Journal of Production Research, 2005, https://doi.org/10.1080/00207540500142274.

[16] J. Dunne, U. Brose, R. Williams. Modeling food-web dynamics: Complexity-stability implications, SFI working paper 2004-07-02, 2004.

[17] T. Choi, K. Dooley, M. Rungtusanatha. Supply networks and complex adaptive systems: Control versus emergence, Journal of Operations Management, vol. 19, pp. 351-366, 2001.

[18] S.C. Brailsford, E. Eldabi, M. Kunc, N. Mustafee, A. F. Osorio. Hybrid simulation modelling in operational research: A state-of-the-art review, European Journal of Operational Research, vol. 278, no. 3, pp. 721-737, 2019.

[19] A. Mittal and C. C. Krejci. A hybrid simulation modeling framework for regional food hubs, Journal of Simulation, vol. 13, no. 1, pp. 28-43, 2019, DOI: 10.1057/s41273-017-0063-z

[20] T. Reggelin, S. Lang, S. and C. Schauf. Mesoscopic discrete-rate-based simulation models for production and logistics planning, Journal of Simulation, 2020, DOI: 10.1080/17477778.2020.1841575

[21] W. Jones and P. Gun. Train timetabling and destination selection in mining freight rail networks: A hybrid simulation methodology incorporating heuristics, Journal of Simulation, 2022, DOI: 10.1080/17477778.2022.2056536

[22] C. Barbosa, C. Malarranha, A. Azevedo, A. Carvalho and A. Barbosa-Póvoa. A hybrid simulation approach applied in sustainability performance assessment in make-to-order supply chains: The case of a commercial aircraft manufacturer, Journal of Simulation, 2021, DOI: 10.1080/17477778.2021.1931500

[23] A. Rosenblueth, N. Wiener, J. Bigelow. Behavior, Purpose and Teleology, Philosophy of Science, vol. 10, pp.18-24, 1943.

[24] F. Lara-Rosano. Las Ciencias de la Complejidad en la solución de nuestros problemas sociales, Sistemas, Cibernética e Informática, vol. 13, no.2, pp.43-50, 2016.

[25] J. H. Holland, K. J. Holyoak. *Induction: Processes of inference, learning, and discovery*. Cambridge: The MIT Press, 1989.

[26] J. H. Holland. Studying complex adaptive systems, J. Syst. Sci. Complex, vol. 19, no. 1, pp. 1–8, 2006.

[27] S. Auyang. *Foundations of complex-system. Theories in economics, evolutionary biology, and statistical physics*. Cambridge: Cambridge University Press, 1999.

[28] K. G. Dobson. Complexity science will transform logistics. United States Naval Institute, Proceedings, Annapolis Tomo 130 (4), pp. 74-76, 2004.

[29] F. Nilsson, V. Darley. On complex adaptive systems and agent-based modelling for improving decision-making in manufacturing and logistics settings Experiences from a packaging company, International Journal of Operations &Production Management, vol. 26, pp. 1351-1373, 2006.

[30] C. Wycisk, B. McKelvey, M. Hülsmann. Smart parts, supply networks as complex adaptive systems: analysis and implications, International Journal of Physical Distribution & Logistics Management, 2008, https://doi.org/10.1108/09600030810861198.

[31] D. Ivanov, B. Sokolov. *Adaptative Supply Chain Management*, London: Springer-Verlag, 2010.

[32] A. Nair and J. M. Vidal. Supply network topology and robustness against disruptions – an investigation using multi-agent model, International Journal of Production Research, vol. 49, pp. 1391-1404, 2011.

[33] M. Haghnevis and R. G. Askin. A Modeling Framework for Engineered Complex Adaptive Systems, IEEE Systems Journal, 2012, DOI:10.1109/JSYST.2012.2190696

[34] J. Wojtusiak, T. Warden, O. Herzog. Machine learning in agent-based stochastic simulation: Inferential theory and evaluation in transportation logistics, Computers & Mathematics with Applications, vol. 64, pp. 3658 – 3665, 2012.

[35] Q. Long. Three-dimensional-flow model of agent-based computational experiment for complex supply network evolution, Expert Systems with Applications, 2015, http://dx.doi.org/10.1016/j.eswa.2014.10.036.

[36] R. Reyes Levalle, S. Y. Nof. Resilience in supply networks: Definition, dimensions, and levels, Annual Reviews in Control, vol. 43, 2017, doi:10.1016/j.arcontrol.2017.02.003

[37] N. J. Van Eck and L. Waltman. Visualizing bibliometric networks. In Ding, Y., Rousseau, R., & Wolfram, D. (Eds.) *Measuring scholarly impact: Methods and practice*, pp. 285-320, Springer, 2014.

[38] A. T. Gumus, A. F. Guneri and S. Keles. Supply chain network design using an integrated neuro-fuzzy and MILP approach: A comparative study, Expert Syst. Appl., 2009.

[39] S. Pathak, D. Dilts and G. Biswas. On the evolutionary dynamics of supply chain network topologies, IEEE Transactions on Engineering Management, 2007.

[40] M. Özbayrak, T. Papadopoulou and M. Akgun. Systems dynamics modelling of a manufacturing supply chain system, Simul. Model. Pract. Theory, 2007.

[41] T. Eldabi, A. Tako, D. Bell, A. Tolk. Tutorial on means of hybrid simulation. Proceddings of the 2019Winter Simulation Conference. December 2019, pp. 33-43, 2019.

[42] M. W. McElroy. Integrating Complexity Theory, Knowledge Management, and Organizational Learning, Journal of Knowledge Management, vol.4, no. 3, pp. 195-203, 2000, DOI:10.1108/13673270010377652

[43] A. Ma, A. Zhou, A. Ali, N. Alain. An Agent Based Modelling Approach for Dynamic Risk Modelling in Emergency Response, 2021 IEEE International Conference on Emergency Science and Information Technology (ICESIT) Emergency Science and Information Technology (ICESIT), 2021 IEEE International Conference on. :290-293 Nov, 2021 DOI: 10.1109/ICESIT53460.2021.9696775

[44] https://www.morganstanley.com/ideas/global-ecommerce-growth-forecast-2022 consulted 15/03/23

Dynamic Cost Estimation of Reconstruction Project Based on Particle Swarm Optimization Algorithm

Li Li
Luzhou Vocational & Technical College, Luzhou, Sichuan, 646000, China
E-mail: lili2685@163.com

Keywords: particle swarm optimization, reconstruction works, dynamic cost estimation, PSO- LSSVM model

Received: February 22, 2022

Abstract: This paper proposes the research on dynamic cost estimation of reconstruction project in accordance with the particle swarm optimization procedure in order to predict the value of dynamic cost estimation of reconstruction project. To accomplish the task initially the applicability of example swarm optimization procedure is introduced. The basic principle of particle swarm optimization procedure is described, and PSO (particle swarm optimization) procedure is used to optimize the super factors of LS- SVM. The spss20.0 statistics is used to cluster the sample data to obtain similar engineering classes. BP-NN (Back Propagation Neural Network), LS - SVM, and PSO-LSSVM are implemented to anticipate and simulate the development price for the authentication of request effect of the optimized design in that area. The results show that the relative errors of the three designs are controlled within + 10%, this may be used in the early stages of building to anticipate construction costs accurately. The range of the relative error dissemination interval predicted using the BP-NN design is 13.12% and is between [- 7.46%, 5.74%]. The range of the relative error dissemination interval predicted by the LS SVM design is 14.22% and is between [-8.12%, 6.17%]. According to the PSO-LSSVM design, the relative error dissemination interval is [-2.56%, 2.49%], and its range is 5.21%. In terms of prediction stability and robustness, the prediction design optimized using PSO procedure outperforms the LS SVM design. In conclusion, the predictions design on the basis of PSO optimized LS SVM is better appropriate for predicting construction costs early in the building process and has strong guiding importance for the cost of construction.

Povzetek: Raziskava predlaga dinamično ocenjevanje stroškov obnove projektov z uporabo postopka optimizacije delcev za natančnejše napovedi.

1 Introduction

Cost prediction has an impact on project cost and project plan, and plays a significant role in the field of project management. Therefore, project cost prediction has always been the focus of attention [1]. In the process of project cost prediction, due to the influence of many factors, the designing process of project cost prediction is very complex. Therefore, project cost prediction is also a difficulty in the research of project management and has become an important research direction [2]. At present, the project cost prediction design can be divided into two categories: traditional design and modern design. The traditional design mainly includes quota method, bill of quantities method and so on. The quota method predicts the project cost according to the released budget quota. The project cost prediction result of the design will not be too low or too high, and the prediction error is small. However, this method does not consider market competition factors, human factors and technical improvement factors, and the efficiency of designing and prediction is very low, which is not suitable for the cost prediction of large projects [3]. The bill of quantities method is proposed for the deficiency of the quota method. In practical request, due to the failure to consider the vicious competition between enterprises, the error of project cost prediction is large and the defect is very obvious [4]. Modern project cost prediction designs are divided into two categories: linear design and nonlinear design. The linear design mainly includes fuzzy mathematical design and multiple linear regression design. The change of project cost has certain randomness and nonlinearity, while the linear design cannot describe the nonlinear change characteristics of project cost, and the prediction error of project cost is high. Figure 1 shows the preparation process of construction project investment estimation [5].

When planning a project, by devoting more resources to tasks, the time of the project might be shortened. However, restricting the duration of projects below their standard level is associated with additional costs. The time-price trade-off problem (TCTP) objective is to find a collection of time-cost options that, in specific circumstances, yield the best scheduling [6]. Many studies have concentrated on a discrete variant of this problem, known as the discrete time-cost trade-off problem (DTCTP), because in practice many resources (such as workers and equipment) are accessible in discrete increments [7]. Cost optimization problems, time optimization problems, and Pareto front problems are the three main forms of DTCTP that are frequently discussed

in the literature. A cost optimization issue aims to identify a collection of time cost alternatives that would

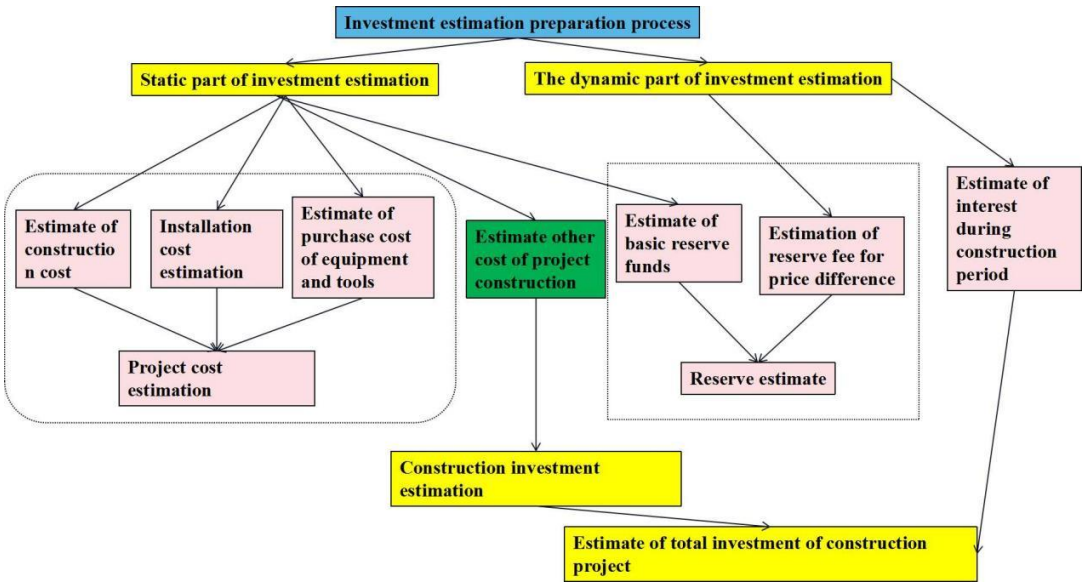


Figure 1: Preparation process of construction project investment estimation.

reduce overall cost under specific constraints, such as given a hard project deadline or penalty for delays. The objective of the budgeting problem is to find time-cost alternatives to shorten the project's timeline while staying within the allocated budget.

The remainder of this essay is structured with a section 2 devoted to literature and a section 3 devoted to research methodologies. Results are shown in Section 4, and the conclusion is offered in Section 5.

2 Related work

This section contains a variety of cutting-edge works in the topic of cost estimate of reconstruction project based on several procedures are discussed.

In view of this research problem, Chen *et al.* start with analyzing the disadvantages of the traditional project cost management system, discuss the necessity, inevitability and realistic environment of implementing bill of quantities pricing in China, and expounds the concept, content, characteristics and preparation methods of bill of quantities, especially how construction enterprises prepare enterprise quota and quotation. In addition, the supporting environment required for the implementation of bill pricing is discussed [8]. Shi *et al.* made an in-depth and detailed analysis on the project cost activities under the bill of quantities pricing mode, made a detailed analysis and Discussion on the project cost control in each stage, and established and improved the project cost management system under the bill of quantities pricing mode [9]. Based on the valuation of bill of quantities and the analysis of project cost composition, Ullah *et al.* put forward the object and content of project cost management in the implementation stage and the management measures in each stage [10]. Peng *et al.* introduced the characteristics and implementation significance of the bill of quantities pricing, and discussed the content, characteristics and precautions of cost management under the bill of quantities pricing mode [11]. Based on the viewpoint of active control, Yan *et al.* systematically analyzed various problems existing in the cost management process of construction projects, analyzed in detail the cost management objectives and main contents of construction projects in different stages, and put forward the measures and methods of whole process cost management, so as to establish the whole process cost management system [12]. Jing *et al.* control the total cost of the project through satisfaction. Because design is an important factor to determine the project cost. Taking the technical route of comprehensive cost management and taking the four aspects of the whole process, the whole risk, the whole team and the whole factor cost management as the starting point, this paper puts forward relevant suggestions and measures for cost management in the design stage, in order to achieve better control effect of project cost [13]. Based on the research and analysis of the whole process, whole life cycle and total cost management, Nanchian *et al.* put forward the cost management theory and method suitable for China's government investment projects, and established the evaluation design for the investment scheme [14]. Ivanov *et al.* discussed in detail and studied the design of cost and quality control objectives in life cycle management [15]. Akimov and Matasov discussed some problems existing in cost control in the design stage of power transmission and transformation project. Combined with the engineering practice of power grid project, this paper analyzes and puts forward relevant measures to control the cost from the aspects of technology, management and

system. It is considered that based on the trinity of technology, management and system, a project cost control system of "reasonable scheme, controllable investment and cost saving" should be established in the design stage [16]. Based on the current research, this paper puts forward the research on dynamic cost estimation of reconstruction project based on particle swarm optimization procedure. Firstly, the applicability of example swarm optimization procedure is introduced. The basic principle of particle swarm optimization procedure is described, and PSO (particle swarm optimization) procedure is used to optimize the super factors of LS-SVM. With the help of spss20.0 to cluster the sample data to obtain similar engineering classes. LS-SVM, PSO-LSSVM, and BP neural networks are used to simulate and estimate the project price for better assessing the request effect of the optimized design in that area. The experimental results verify the efficiency and accuracy of the prediction design on the basis of PSO optimized LS SVM.

3 Research methods

The design process, fundamental principles, and particle swarm optimization technique are covered in this section.

3.1 Applicability investigation of the particle swarm optimization technique

Particle swarm optimization (PSO), which was developed to find the global optimal solution to the optimization issue, was first motivated by the foraging behavior of birds. The law of this biological group movement is that assuming that the birds are randomly foraging in an area, and the birds do not know the specific location of the food, but they can perceive the distance between the food and their current location, the best foraging strategy is to search around the birds closest to the food. Unlike genetic algorithm (GA), PSO procedure does not rely on the evolutionary idea of individual survival of the fittest, but simulates swarm intelligence behavior and finds the global optimal solution through competition and cooperation among individuals [17]. In PSO procedure, the possible solutions of all target problems are in d-dimensional space, and each solution is equivalent to a "bird", which is called "division". The current position of the division is determined by the objective function, and its fitness value can be obtained. On the one hand, each division relies on its own experience to remember its best position, on the other hand, it perceives the best position of divisions other than itself according to the experience of its peers. Combining these two experiences, each division follows the division in the current optimal position at a speed that determines its distance and direction. The velocity of divisions is constantly updated with the change of relevant information, so that the position of divisions is dynamically adjusted at any time until a certain condition is reached, then the procedure is considered to obtain the optimal solution. The PSO method has a straightforward structure, a wide search space, and quick convergence. It can solve the majority of global optimal solution problems. The particle swarm may share the historically optimal division positions with one another during the optimization phase of the goal issue since it has memory. At the same time, there aren't any laborious mathematical processes like crossover and mutation, and there aren't any too complicated factors that have to be defined. It has outstanding advantages compared with other swarm intelligence procedures [18]. The integration with other intelligent procedures is a key development trend for the PSO procedure, which has currently moved beyond its original focus on function optimization to include more general requests like pattern recognition and neural network training. The super factor σ and C setup for the LS SVM procedure is a critical optimization challenge. As a result, PSO is used in this study to optimize the LS SVM factors and increase the design's ability to predict outcomes accurately [19].

3.2 Basic principle of particle swarm optimization

The prime goal of the PSO method is to prepare the position and motion of a collection of casual divisions and, given specific constraints, to continuously iteratively seek for the best solution. The division modernizes its speediness and position in the following round of repetition in accordance with the changes of the two if the finest position passed by each division in the examination procedure is defined as the individual extreme price P_best and the top location found by the current entire group is defined as the global extreme value G_best [20].

Equation 1 and 2 provides a mathematical representation of the particle swarm optimization approach. Assume that a population of m divisions indicating a potential solution to the issue can exist in a d-dimensional search space.

$$X = \{X_1, X_2, ..., X_m\} \tag{1}$$

Among them,

$$X_i = \{x_{i1}, x_{i2}, ...x_{id}\} \tag{2}$$

Signifies the position of the i^{th} division in space, and the corresponding fitness can be calculated by substituting X_i into the goal function connected to the issue at hand. Equations 3 and 4 describe how fast this division will travel in the d-dimensional search space.

$$V_i = \{v_{i1}, v_{i2}, ..., v_{id}\} \tag{3}$$

Use:

$$P_i = \{P_{i1}, P_{i2}, ..., P_{id}\} \tag{4}$$

To represent the best position P_{best} searched by the division itself (the corresponding fitness value is the smallest) by using Equation 5.

$$P_g = \{P_{g1}, P_{g2}, ..., P_{gd}\} \quad (5)$$

To represent the best position G_{best} of the whole population, the velocity and position of the i-th division are updated by Equations 6 and 7 respectively.

$$v_i^{k+1} = \omega v_i^k + c_1 r_1 \left(P_i^k - X_i^k\right) + c_2 r_2 \left(P_g^k - X_i^k\right) \quad (6)$$

$$X_i^{k+1} = X_i^k + v_i^{k+1} \quad (7)$$

Where ω is the weight of inertia and K is the quantity of iterations; c_1 and c_2 indicate, what are referred to as knowledge variables, the steps that divisions take to reach their ideal location while flying and their total optimal position; r_1 and r_2 are two randomly chosen numbers among [0,1]. According to the observation formula (6), essentially, there are three components to the velocity update formula: first, ωv_i^k indicates the speed of the i^{th} division during the earlier time period, that's referred to as the inertia component. Next, $c_1 r_1 (P_i^k - X_i^k)$ indicates the gap between the ideal location of the i^{th} division and its present location, which is its own experience, is referred to as the cognitive portion; Lastly, $c_2 r_2 (P_g^k - X_i^k)$ depicts the distance between the group's ideal location and the current position and the i^{th} division, it develops from peer experience and represents collaboration and information exchange amongst division groups, and is referred to as the social aspect. In this manner, the division's location in the $K + 1$ using the formula above, iteration may be calculated [21]. The pace is typically restricted to prevent doing an excessive amount of blind searching for divisions $[-v_{max}, v_{max}]$ and the search space is limited to $[-x_{max}, x_{max}]$.

3.3 Parameter setting of particle swarm optimization procedure

The PSO procedure needs to have a few factors changed, but reasonable factor setting will also affect its optimization ability and efficiency. Based on extensive reference to relevant literature, this paper obtains the value range of main factors in PSO as follows:

3.3.1 Population size and division dimension

When the population size is too large, the optimization performance of PSO procedure will be enhanced, but at the same time, the convergence speed will be reduced. If the population size is too small, the procedure is easy to fall into local optimization and difficult to jump out. In general, 20 ~ 40 divisions are enough. For special problems such as multi-objective optimization, the group size can be 100 ~ 200 85. The division dimension depends on the specific optimization objective, which is generally the dimension of the solution space of the objective optimization problem.

3.3.2 Inertia weight factor ω

Inertia weight factor ω can efficiently strike a balance between local and global search capabilities. It is mainly used to control the influence of the current flight speed of divisions on the next speed, which is conducive to the procedure to fewer iterations are needed to reach the ideal answer. When the ω value is small, it is convenient for local search, and when the ω value is large, it is convenient for global search [22]. The results show that when the maximum speed is small, the optimization effect is better when the value of ω is approximately 0.9, while when the maximum speed is large, the value of 0.4 is better.

3.3.3 Termination conditions

Usually, when the fitness value of the procedure reaches the maximum value or meets a preset maximum number of iterations, the procedure calculation ends.

3.3.4 Knowledge factors c_1 and c_2

Knowledge factors c_1 and c_2 represent the weights of divisions moving near the optimal position of individuals and groups. Through the setting of knowledge factors, divisions can self review and actively learn other excellent individuals in the group, so as to continuously approach the optimal solution. Generally, the values of c_1 and c_2 are within the range of [0,4], and $c_1 = c_2$.

The following situations need to be explained:

When $c_1 = c_2 = 0$, the division keeps flying at a uniform speed until the boundary of the problem, so the search area is limited and it is difficult to find the solution;

When $c_1 = 0$ and $c_2 \neq 0$, the divisions lack the cognitive part, move to the G_{best} position, and tend to be locally optimal;

When $c_1 \neq 0$ and $c_2 = 0$, the divisions lack the social part, lack the shared information and cooperation in the group, and basically cannot obtain the optimal solution.

3.4 Optimization process of particle swarm optimization procedure

The specific calculation stages of the PSO method may be summed up as follows in accordance with the main concept of following the current optimal position division:

i. Set the necessary algorithmic factors, such as the element swarm size and the inactivity weight feature ω, knowledge elements c_1 and c_2 and close conditions.

ii. Initializing the population, setting the location and velocity vectors to random values of each division in the search space.

iii. Calculate the fitness value $fitness(i)$ of divisions, and find the historical optimal position $P_{best}(i)$ and the current ideal location G_{best} of each division.

iv. According to equations (6) and (7), adjust the division positions and velocities.

v. Determine the fitness value for each division following an update to its speed and position, and then contrast it with the fitness value corresponding to that division's historically best position $P_{best}(i)$. Use the most recent change to the position if the current fitness value is higher $P_{best}(i)$.

vi. A comparison between each division's fitness value and the fitness value corresponding to the global ideal location G_{best} is made. When the global ideal position G_{best} greater than the present ability value, which is modified to reflect the location of the current division.

vii. Check whether the optimization result meets the termination conditions. If not, go back to step (3) and carry out the iteration again. Otherwise, output the P_{best} as the overall best answer and halt the search, equivalent to the current fitness rating.

3.5 Optimization of key factors of LS - SVM based on PSO

To apply the PSO technique to improve the LS SVM's super factors, two important contents need to be considered in the fusion process of the two procedures: one is the representation method of super factors, and the other is the definition of division fitness function.

3.5.1 Representation of hyperparameters

For LS- SVM, since the kernel function selected in this paper is RBF kernel function, there are two hyperparameters to be optimized, namely regularization factor C and intra kernel factor σ. When PSO is used for optimization, each division is required to characterize the possible explanation of the optimization issue, that is, the combination of super factors. Therefore, in the optimization process of LS SVM, a two-dimensional vector is used to represent the combination of the two super factors C and σ, that is, the position of the i^{th} division can be expressed as Equation 8.

$$X_i = (C_i, \sigma_i) \tag{8}$$

3.5.2 Fitness function

Fitness is a measure to evaluate the position of divisions, and also indirectly describes the generalization performance of LS SVM design. In this paper, the mean square error (MSE) of the prediction results is selected as the performance index of the prediction design, and its mathematical definition is presented in Equation 9.

$$f = \frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2 \tag{9}$$

Where y_j represents the actual output value of the j^{th} sample; $\hat{y}_j$ represents the predicted value of its design. By kernel function expression is presented in Equation 10.

$$K(x_i, x_j) = \exp\left(-\frac{|x_i - x_j|}{\sigma^2}\right) \tag{10}$$

And LS- SVM regression expression is shown in Equation 11.

$$f(x) = \sum_{i=1}^{n} a_i K(x_i \cdot x_j) + b \tag{11}$$

It can be seen that f can be regarded as the composite function of C and σ. In the fitness evaluation, the smaller the error, the better the division position. Then the optimization problem can be described as Equation 12 and 13.

$$\min_{C,\sigma} f = \frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2 \tag{12}$$

$$s.t. C \in [C_{\min}, C_{\max}], \sigma \in [\sigma_{\min}, \sigma_{\max}] \tag{13}$$

Where C_{min}, C_{max} and σ_{min}, σ_{max} represent the minimum and maximum values of C and σ respectively. In the designing process of LS-SVM, PSO is used to search in the value range of C and σ to minimize the fitness value f of equations (12) and (13), then the solution vector (C, σ) is the optimal super factor to be found by LS-SVM.

3.6 Construction cost prediction process based on PSO – LSSVM

The following designing concepts may be created using the basic study on PSO procedure and LSSVM design mentioned above:

i. The preprocessed historical statistics are formed into an example matrix by sample clustering and index dimensionality reduction.

ii. The exercise data set and the estimate data set should be mentioned.

iii. Define the factors range of standards C and σ, and initialize the populationn $X = \{X_1, X_2, \ldots, X_m\}$ composed of M divisions by using PSO procedure, where $X_i = \{C_i, \sigma_i\}$.

iv. It calculates and contrasts the division fitness with the training data set, the specific optimum assessment $P_{best}(i)$ and the global optimum

assessment Grest are regular, and the speed and location of every division are efficient.

v. Iterate repeatedly until the end conditions are met (minimum fitness value).

vi. Output optimization hyperparameters C and σ and assigned to LS SVM prediction design.

vii. Train the design with the training regular, input the prediction sample data to predict the design, and achieve the optimum result.

In this manner, based on the actions above and combined with the actual engineering situation, the factors of PSO procedure can be set applicably, and the construction on the basis of factor optimization, project cost projection is possible.

Table 1: Input set of prediction design

Category	**Sample Number**																
Type 1	19	17	3	49	21	32	16	39	11	18	12	38	35	34	14	2	40
	5	29	25	46	42	1	36	31	41	13	30	45	23	50	37	7	
Type 2	4	20	9	43	26	28	47	10	44	33	27	22	6	48	8	24	15

Table 2: Clustering results of sample system

Sample Number	**Principal Component**							
	1	2	3	4	5	6	7	8
1	0.169	-1.011	0.482	0.803	-1.248	-0.555	-0.542	-0.002
2	-1.568	-0.474	0387	-0.328	1.490	-1.002	1.131	-0963
3	-0.393	2.296	-1.077	0.042	-0.010	-1.360	0.137	-2.062
4	0.678	-0.675	-1.247	0.569	0.615	-1.633	0.1 08	0.821
5	1.644	0.570	1.315	-0.252	0.471	-1.446	0.504	-0.717
6	-0.697	0.076	-1.547	-0.215	0.414	-0.419	-0.085	0.102
7	-0.790	-1.428	-1.585	0.557	0.460	0.380	0.680	0.779
8	0.555	0.071	-0.269	0.939	-0.118	1.159	1.321	0.531
9	-1.797	-0.133	2.009	0.314	0.441	0.012	-1 .001	-0.956
10	-0.404	0.688	-0.726	-1.090	-0.182	-0.449	-0. 354	0.417
11	-0.107	2.031	-1.089	0.085	-1.111	0. 664	-0.434	0.143
12	-0.720	-0.855	0.180	0.217	-0.515	0.923	-1 026	1.248
13	-0.810	1.252	0.573	1.467	-0.019	2.743	-0.399	-0.269
14	-0.021	1.638	0.816	-1.529	2.195	1.140	-1.385	0.449
15	1.075	0.444	0.912	0.611	-0.307	-0.963	-2.072	1.034
16	1.252	-0.810	-1.135	-0.155	-0.343	-0.061	-0.430	-1 .043
17	0.489	0.162	-0.924	-0.964	-0.478	0.534	0.858	-1.807
18	0.465	0.382	0.687	1.744	-0.282	-0.494	-0.183	0.794
19	-0.301	-1.664	1.444	-1.663	1.404	0.239	2.188	-0.142
20	-0.112	0.870	-0.741	-1.858	-0.516	1.214	-1.910	0.598
21	-1.709	0.021	-1.006	0.790	1.044	0.119	-0.683	1.244
22	-1. 094	0.056	1.408	-0.499.	-1.306	-1.438	-0.908	0.342
23	0.796	-0.723	0.182	-0.731	1.395	0 989	-0.906	-0.193
24	1.976	0.103	-0.439	0.096	0.825	-0. 994	0.204	0.950
25	-0.568	0.240	1.076	-0.724	-1 .700	-0.789	0.733	-1.348
26	-1.387	-0.409	-0.862	0.912	-0.697	-0.238	0.245	-0.918
27	0.290	-0.957	-1.378	-1.294	-0.297	-0.034	0.939	0.205
28	0.355	0.090	0.191	2. 665	1.338	-0.277	-0.092	-0.529
29	0.073	-1.748	-0.123	-0.782	-0.556	0.082	0.384	0.037
30	0.541	0.517	1.336	-0.520	-2.183	0.395	1.466	2.204
31	-0.180	-1.448	0.500	0.590	-0.606	0.706	1.679	-0.658
32	-0.049	1.137	0.217	-0.295	1.344	-0.847	0.518	1.345
33	2.350	-0.310	0.432	0.498	-0.133	1.699	0.680	-1.635

3.7 Model implementation environment

MATLAB (Matrix Laboratory) is a scientific computing environment for mathematical calculation, programming development and result visualization. It integrates the functions of matrix operation, numerical analysis, design establishment and simulation into a convenient visual window, which provides convenience for the solution of many research problems. Different

from C language, FORTRAN and other computer programming languages, Matlab can interactively accept

various instructions from users to output results, and has developed many toolboxes for specific problems, which is convenient for users to directly learn and apply the corresponding methods. Therefore, this study chooses this platform to complete the extremely complex operation process in PSO - LSSVM prediction design.

3.8 Data processing

The data collected in this paper comes from a reconstruction construction project. This paper systematically clusters the sample data with the help of spss20.0 to obtain similar engineering classes. After 6 times of clustering, all cases can be combined into one category. In the last iteration process, all sample projects can be divided into two categories, including 33 in the first category and 17 in the second category, as shown in Table 1. In order to overcome the error caused by too small sample size and make the prediction design achieve better prediction effect, the first type of similar samples with large sample size are selected for further prediction analysis. The input data set of the prediction design can be obtained according to the attribute value of the sample, as shown in Table 2.

4 Results and analysis

The example matrix states that the main 25 of the 33 matters are utilized as training examples for knowledge and preparation, and the trained design is used to test the remaining eight collections of data. To more effectively assess how the design applies to project cost prediction following factor optimization, three methods of BP-NN, LS- SVM and PSO- LSSVM stay selected for project cost simulation prediction.

4.1 BP-NN design simulation and prediction

To finish the construction of the method, the BP-NN software of MATLAB must be used to forecast and analyze engineering sample data. To construct a new BP-NN net, first call the factor setting function new FF. Next, call the train task to train the BP neural network using the training set of data. The maximum number of iterations is 1000, the knowledge feature is 0.05, and the motion feature is 0.65 all at the same time. Finally, predictions are made using the SIM task and the qualified NN. Figure 2 depicts the relationship between the test sample's actual value and the expected cost value.

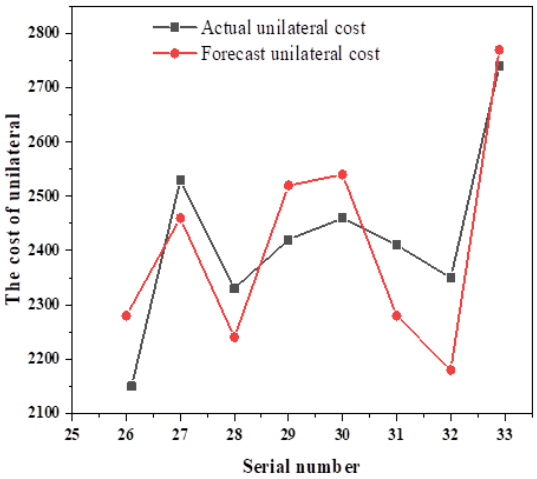


Figure 2: Prediction curve of BP-NN design.

4.2 LS- SVM model simulation prediction

The standard LS-SVM design is established to train and predict the sample data, and the standardized program is compiled under the environment of matlab2016a. The super factor C = 10 is determined by trial-and-error method, σ= 30. The design is created and trained using the trainlssvm function, examination examples are introduced into the trained design, prediction is made using simlssvm, and visualization is done using the plot LSSVM task. The cost prediction findings are displayed in Figure 3.

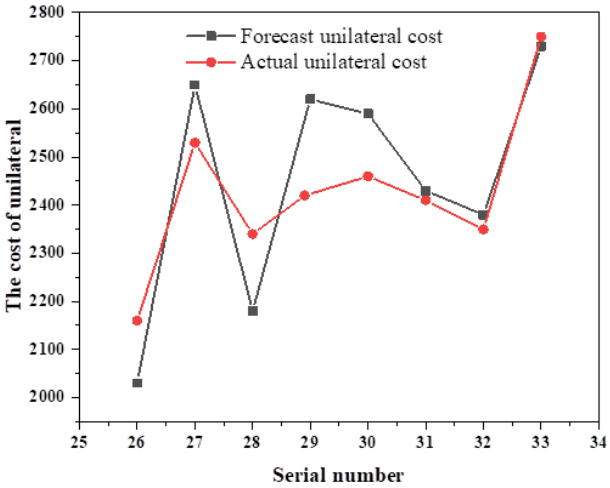


Figure 3: Prediction curve of LS-SVM design.

4.3 PSOLSSVM model simulation prediction

In order to analyze cost prediction, the PSOLSSVM design is utilized. A transfer into the libsvm-3.22 toolbox is the underlying assumption, and research on factor situation in the prediction design is incorporated, the super factor's value range of C and σ is set to [0, 100]. In parallel, the PSO procedure's factors are set to population size pop = 20; Evolution times Max Gen = 200; Inertia weight feature $\omega = 0.5$; Knowledge factor $c_1 = c_2 = 2$. By applying the radial basis task (kernel = rbf_kernel ')

and using the mean square deviation of the predicted value of the examination set as the fitness value, the PSO-LSSVM prediction design is created. The value of the optimized factor is: C = 13.5201, σ = 46.4981. Figure 4 displays the test samples' findings for cost prediction.

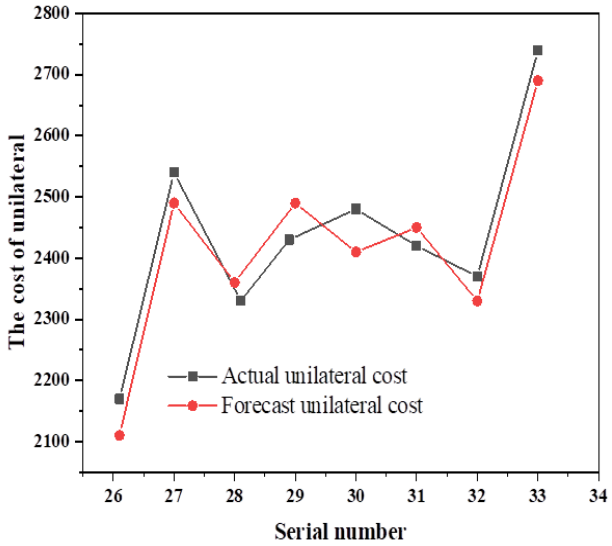


Figure 4: PSO-LSSVM design's prediction curve.

Table 3: Comparison of prediction results of different procedures

Test sample	Actual value of unilateral cost	Predicted value of unilateral cost		
		BP neural network	LS-SVM	PSO-LSSVM
26	2151.23	2274.39	2042. 03	2097.26
27	2537.89	2463.72	2648. 10	2487.52
28	2329.96	2235.11	2188.04	2352.86
29	2422.11	2519.03	2616.97	2486. 13
30	2465.89	2548. 85	2592. 15	2411.71
31	2418.21	2281. 60	2436. 50	2442. 57
32	2356.91	2183.29	2382. 28	2329.18
33	2749.11	2772.78	2721.72	2696.45

Table 4: Error in the test sample's expected value compared to its actual value

Test sample	Relative error			Relative error of mean absolute value		
	BP neural network	LS-SVM	PSO-LSSVM	BP neural network	LS-SVM	PSO-LSSVM
26	-5.61	5.17	2.49	4.23	3.94	1.82
27	2.81	-4.43	1.87			
28	4.10	6.17	-0.87			
29	-4.12	-8.12	-2.56			
30	-3.45	-5.21	2.19			
31	5.74	-0.67	-1.10			
32	-7.46	-1.16	1.21			
33	-0.79	1.21	1.89			

4.4 Analysis of prediction results

Table 3 displays a comparison of the outcomes of the above methodologies calculations. As indicated in Table 4, the relative fault δ and regular total percentage error are used to demonstrate the influence of various procedures on the prediction result of the design.

$$\delta = \frac{y_i - \hat{y}_i}{y_i} \tag{13}$$

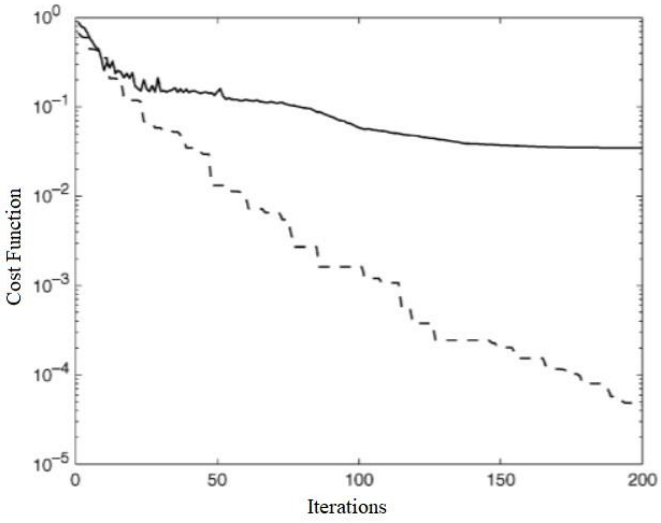
$$MAPE = \frac{1}{N}\sum_{i=1}^{n}|\delta| \times 100\% \tag{14}$$

Where y_i and $\hat{y}_i$ indicate the cost's projected and actual values, respectively of the i^{th} sample, and N is the number of test samples.

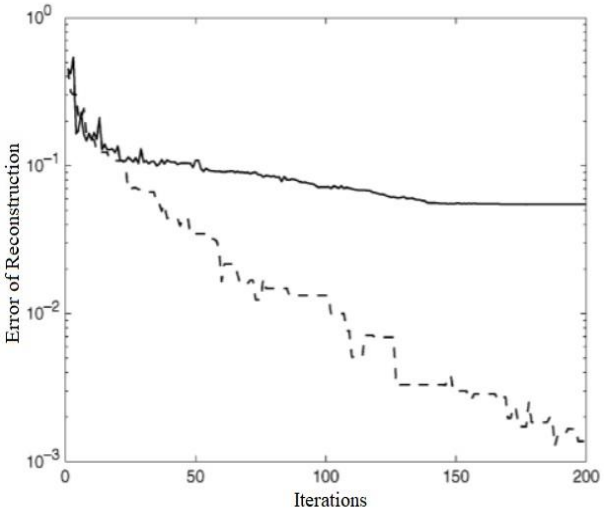
4.4.1 Accuracy analysis

By looking at the data in Table 4, it can be seen that the design performs well and can encounter the precision necessities of production project cost prediction in the early stages of production. The relative errors of LS-SVM, PSO-LSSVM, and BP neural network-based building project cost prediction are controlled within + 10%.

The optimized LS-SVM design significantly lowers the regular complete value relative fault whereas the BP-NN average absolute value relative error is nearly identical to that of the regular LS-SVM design. Consequently, the PSO-LSSVM-based production price prediction design performs superior in error management and has improved prediction precision.



(a)



(b)

Figure 5: (a) Cost task, (b) error of reconstruction with respect to the number of iterations

The accuracy of the proposed procedure is estimated with respect to the cost task and error of reconstruction. The observation of cost task and error of reconstruction with respect to the number of iterations for the proposed procedure is depicted in Figure 5. It is observed from the experimentation that the proposed procedure outperforms for both cost task and error of reconstruction in comparison with existing baseline design [5, 6].

4.4.2 Stability analysis

Table 4 shows that the projected comparative fault dissemination interval using the BP-NN design is [-7.46%, 5.74%], with a range of 13.12%.The range of the comparative fault dissemination interval predicted using the LS SVM design is 14.22%, and it is between [-8.12%, 6.17%]. According to the PSO-LSSVM design, the relative error dissemination interval is [-2.56%, 2.49%], and its range is 5.21%. It is clear that the PSO procedure-optimized prediction design outperforms the LS-SVM design in terms of prediction stability and robustness of the prediction effect.

4.4.3 Forecast time analysis

The prediction period based on the BP-NN design is 9.15 seconds from the start of the programme running to creating the prediction graphic. According to the LS-SVM design, the forecast time is 3.17 seconds, whereas the PSO-LSSVM design predicts the time to be 13.14 seconds. The PSO procedure's optimization increases the design's training time, but the time difference between the three has minimal bearing on how much building projects will ultimately cost in real life.

5 Conclusions

In this paper, the research on dynamic cost estimation of reconstruction project based on particle swarm optimization procedure is proposed. MATLAB platform is used to design simulation and analysis from data collection to preprocessing. The project cost prediction design system based on factor optimization is applied to practical engineering cases, which realizes the good combination of theory and practice. The experimental results verify the efficiency and accuracy of the prediction design based on PSO optimized LS-SVM. Numerical results show that proposed procedure leads to better reconstruction compared to existing design with same number of iterations. Although, both procedures are consistent and deliver accurate reconstruction shape even when the scattered field dimensions are despoiled by additive white Gaussian noise. For different construction projects, how to form a more scientific and perfect construction project cost prediction index system, or establish an index system platform to select and change indicators according to different engineering projects needs further research.

References

[1] Fan, H., Zhan, H., Cheng, S., & Mi, B. (2019). Research and request of multi-objective particle swarm optimization algorithm based on α-stable dissemination. *Xibei Gongye Daxue Xuebao/Journal of Northwestern Polytechnical University*, *37*(2), 232-241. *https://doi.org/10.1051/jnwpu/20193720232*

[2] Wu, K. (2014). A tunable workflow scheduling algorithm based on particle swarm optimization for cloud computing. *https://doi.org/10.31979/etd.wy2s-568v*

[3] Chiang, J. S., Gu, W. S., Chiu, C. C., & Sun, C. H. (2015). Estimation of the two-dimensional homogenous dielectric scatterer in a slab medium using particle swarm optimization and asynchronous particle swarm optimization. *Research in Nondestructive Evaluation*, *26*(4), 208-224. *https://doi.org/10.1080/09349847.2015.1024906*

[4] Xu, J., Noo, F., & Tsui, B. M. (2017). A direct algorithm for optimization problems with the Huber penalty. *IEEE transactions on medical imaging*, *37*(1), 162-172. *10.1109/TMI.2017.2760104*

[5] Miandoab, E. E., & Gharehchopogh, F. S. (2016). A novel hybrid algorithm for software cost estimation based on cuckoo optimization and k-nearest neighbors algorithms. *Engineering, Technology & Applied Science Research*, *6*(3), 1018-1022. *https://doi.org/10.48084/etasr.701*

[6] Rho, S. B., Lee, J. H., Park, M. S., Byun, H. J., Kang, S., Seo, S. S., & Park, S. Y. (2011). Anti-apoptotic protein TCTP controls the stability of the tumor suppressor p53. *FEBS letters*, *585*(1), 29-35. *https://doi.org/10.1016/j.febslet.2010.11.014*

[7] Wuliang, P., & Chengen, W. (2009). A multi-mode resource-constrained discrete time–cost tradeoff problem and its genetic algorithm based solution. *International journal of project management*, *27*(6), 600-609. *https://doi.org/10.1016/j.ijproman.2008.10.009*

[8] Chen, S., Liu, H., Hu, Z., Zhang, H., Shi, P., & Chen, Y. (2015). Simultaneous reconstruction and segmentation of dynamic PET via low-rank and sparse matrix decomposition. *IEEE Transactions on Biomedical Engineering*, *62*(7), 1784-1795. *10.1109/TBME.2015.2404296*

[9] Shi, R. G., Chen, Y., & Wei, M. F. (2018). Least squares estimate algorithm based on particle swarm optimization for position reconstruction of scintillation crystal. *Journal of Instrumentation*, *13*(12), P12029. *http://dx.doi.org/10.1088/1748-0221/13/12/P12029*

[10] Ullah, A., Wang, B., Sheng, J., Long, J., Asim, M., & Sun, Z. (2020). Optimization of software cost estimation model based on biogeography-based optimization algorithm. *Intelligent Decision Technologies*, *14*(4), 441-448. *10.3233/IDT-200103*

[11] Peng, Y., Zhao, S., & Wang, H. (2020). A digital twin based estimation method for health indicators of DC–DC converters. *IEEE Transactions on Power Electronics*, *36*(2), 2105-2118. *10.1109/TPEL.2020.3009600*

[12] Yan, J., Zhang, J., Liu, Y., Han, S., Li, L., & Gu, C. (2015). Unit commitment in wind farms based on a glowworm metaphor algorithm. *Electric Power Systems Research*, *129*, 94-104. *https://doi.org/10.1016/j.epsr.2015.07.018*

[13] Jing, Y., Liu, Y., & Bi, M. (2019). Quantum-inspired immune clonal algorithm for railway empty cars optimization based on revenue management and time efficiency. *Cluster Computing*, *22*(1), 545-554. *https://doi.org/10.1007/s10586-017-1292-7*

[14] Nanchian, S., Majumdar, A., & Pal, B. C. (2015). Three-phase state estimation using hybrid particle swarm optimization. *IEEE Transactions on Smart Grid*, *8*(3), *1035-1045. 10.1109/TSG.2015.2428172*

[15] Ivanov, D., Dolgui, A., Sokolov, B., & Werner, F. (2016). Schedule robustness analysis with the help of attainable sets in continuous flow problem under capacity disruptions. *International Journal of Production Research*, *54*(11), 3397-3413. *https://doi.org/10.1080/00207543.2015.1129467*

[16] Akimov, P., & Matasov, A. (2014). Weight and Time Recursions in Dynamic State Estimation Problem With Mixed-Norm Cost Function. *IEEE Transactions on Automatic Control*, *60*(4), 1050-1063. *10.1109/TAC.2014.2365687*

[17] Paik, J. W., Hong, W., & Lee, J. H. (2020). Direction-of-departure and direction-of-arrival estimation algorithm based on compressive sensing: Data fitting. *Remote Sensing*, *12*(17), 2773. *https://doi.org/10.3390/rs12172773*

[18] Ghaemi, S., Hamzeh Aghdam, F., Safari, A., & Farrokhifar, M. (2019). Stochastic economic analysis of FACTS devices on contingent transmission networks using hybrid biogeography-based optimization. *Electrical Engineering*, *101*(3), 829-843. *https://doi.org/10.1007/s00202-019-00825-6*

[19] Huang, J., Li, Y. F., & Xie, M. (2015). An empirical analysis of data preprocessing for machine learning-based software cost estimation. *Information and software Technology*, *67*, 108-127. *https://doi.org/10.1016/j.infsof.2015.07.004*

[20] Chen, L., Li, Q., Zhao, X., Fang, Z., Peng, F., & Wang, J. (2019). Multi-population coevolutionary dynamic multi-objective particle swarm optimization algorithm for power control based on improved crowding distance archive management in CRNs. *Computer Communications*, *145*, 146-160. *https://doi.org/10.1016/j.comcom.2019.06.009*

[21] Hongwei, Z., Liwei, T., & Dongzheng, W. (2015). Research on improved firefly optimization algorithm based on cooperative for clustering. *International Journal of Smart Home*, *9*(3), 205-214. *http://dx.doi.org/10.14257/ijsh.2015.9.3.19*

[22] Li, N., Wu, Z., & Zhao, Z. (2021). Application of Internet of Things Compressed Sensing and Information Interaction Technology in Intelligent Transportation Layout. *Scientific Programming*, *2021*. *https://doi.org/10.1155/2021/9979198*

Design of Ecological Land Remediation Planning and Remediation Mode Based on Spatial Clustering Algorithm

Yuzhuo Yao
Hebei Bureau of Geology and Mineral Resources second geological brigade, Tangshan, Hebei, 063000, China
E-mail: yuzhuoyao9@163.com

Keywords: spatial clustering algorithm, ecological land, remediation planning, remediation mode, design

Received: February 22, 2022

In this article a method is proposed based on a spatial clustering algorithm, in order to better realize the design of ecological land consolidation planning and regulation mode. Considering the GIS methods such as spatial clustering algorithm and least resistance model, the proposed model selects typical projects and put forward ecological land consolidation planning scheme. The unified engineering construction forms and rigid engineering construction standards aims at regularization and rigidization of traditional land consolidation by means of high-intensity engineering construction. Although in the process of stabilizing the amount of cultivated land and improving agricultural production conditions, played a positive role. However, under the current requirements of ecological civilization construction, there is an urgent need for transformation. Through the improvement of ecological land, the ecological improvement is realized, and the need for landscape improvement is foremost requirement. The study is guided by the theory of landscape ecology, according to the general idea of "landscape pattern evaluation-land remediation function zoning-corridor pattern optimization-patch matrix optimization". The experimental results are generated using Fragstates software for calculating various indicators. The results show that, before and after renovation, the landscape types with the largest landscape ecological security index LESi were all cultivated land, and the smallest one was forest land. Except for the river landscape ecological security index, which dropped by 9.84%, the ecological security of other types of landscapes improved. Among them, the largest increase was road (121.29%), followed by forest land (43.10%). According to formulas (1) and (2), the landscape safety index LES of the project area is 0.45 before the renovation, and 0.61 after the renovation. An increase of 35.56% is observed which shows that through the ecological land consolidation project area, the overall landscape ecological security status has been improved. It is proved that the spatial clustering algorithm can better realize the design of ecological land remediation planning and remediation mode.

Povzetek: Razvita je metoda, ki temelji na algoritmu prostorskega gručenja, za izboljšanje načrtovanja in regulacije ekološke konsolidacije zemljišč z uporabo GIS metod.

1 Introduction

With the rapid development of urbanization, many high-quality arable lands have been occupied in large quantities; In the countryside, there are a large number of hollow villages, houses that are idle or abandoned, rural infrastructure is poor, water pollution is serious, rural landscapes are damaged, and the quality of living environment is poor. Therefore, various regions have successively carried out new rural construction, village appearance improvement, rural land improvement, capital investment has been increased in projects such as agricultural infrastructure construction, comprehensive agricultural development, and returning farmland to forests. These policies and action plans are of great significance to China's rural development and urban-rural integration, and have achieved remarkable results [1]. However, in the process of new rural construction and land consolidation, due to the lack of ecological landscape theory and technical guidance, coupled with the limited business level of management and construction personnel, as a result, the construction of new rural areas is monotonous and has no characteristics. Some peasant resettlement sites are like urban communities, with high-rise and multi-storey barracks-like distribution, and the culture, size and color are seriously inconsistent; Because the ecosystem structure and function composed of local topography, water system and organisms are not brought into play, coupled with a blunt understanding and pursuit of the standardized construction of "Tian Chengfang, Road Chengwang, Canal Interconnection, Trees Forming", excessive transformation of the land in a bulldozer style, despising circulation and symbiosis, as a result, a large number of ditches and roads that need to be ecologicalized have been excessively hardened, and diversified groves have been cut down, the ponds were filled and the rivers were straightened, resulting in serious damage to the biological, ecological and local features of life that nurtured our regional culture [2]. These issues will be important issues and challenges faced by rural land consolidation.

Land consolidation is the arrangement and coordination of land resources and their utilization methods, it is an important platform for promoting new urbanization and implementing the rural revitalization strategy, as well as promoting the coordinated

development of urban and rural areas, it is an important means to build beautiful and livable villages and promote targeted poverty alleviation and poverty alleviation. Objectively analyze the differences in regional natural and social and economic conditions, accurately grasp the direction and goals of land consolidation in the study area, scientific division of land consolidation types, it can effectively guide the smooth development of land consolidation work [3].

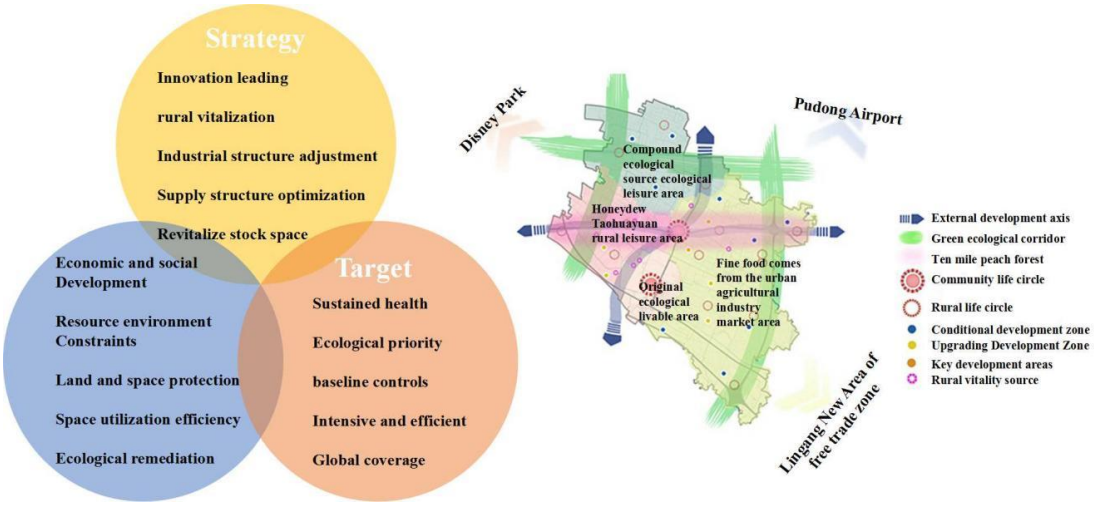


Figure 1: Area mapping based on target and strategy.

The area mapping based on target and strategy is presented in Figure 1. Landscape ecology is a multidisciplinary subject, it covers ecology, geography, environmental science, resource science, management science and other related theories, it is mainly based on the principle of ecosystem, combined with systematic research methods, analyze and study the internal structure, function and dynamic change process of landscape at a certain scale, in order to beautify the landscape pattern and optimize the landscape structure, promote the scientific utilization, rational protection and effective development of landscape. Landscape ecological land remediation, mainly based on the theoretical basis of landscape ecology and ecological security pattern, relying on ecological land remediation technology, by adjusting the functions and structures of patches, corridors and substrates in the landscape, optimize the spatial layout of land remediation, improve the ecological stability and ecological service value of land remediation landscape [4]. The rest of this article is systematized as literature is presented in section 2 followed by research methods in section 3. Section 4 depicts the results and the conclusion is presented in section 5.

2 Related work

In this section various state-of-the-art work in the field of land remediation and planning using several approaches are studied and discussed.

Moravcova *et al.* discusses that artificial land consolidation, not only can the road system be optimized, for improving water resources management system and natural landscape conditions, also played an important role. Some scholars also pointed out that, a series of engineering and biological measures during the implementation of land consolidation, it may affect the quantity and quality of vegetation on the surface [5]. Cong *et al.* pointed out the goal of land consolidation, if only limited to increasing agricultural productivity and improving production conditions, the implementation of the land consolidation, it is likely to have a negative impact on the ecological environment [6]. For example, Long *et al.* research shows that after land consolidation, farmers' labor rate of return has been effectively improved, thus prompting them to support the land consolidation policy [7]. Gong and Tan proposed that the social benefits of land consolidation should include four aspects, for example, the degree of support of local farmers and the implementation of the project, the impact on people's customs, entertainment and social welfare [8]. Guo *et al.* selected the properties of the plot itself, natural resource conditions, economic conditions and other aspects as evaluation indicators, for 3 different land consolidation projects in the Czech Republic, the pre-implementation benefit evaluation and the post-implementation benefit evaluation are carried out respectively. The results show that, the determination of the initial evaluation criteria and evaluation model has a significant impact on the benefits of land consolidation projects [9]. Ng presented a study which presents that according to the regional differences in topography, landforms, land use, and the unbalanced social and economic development of the province, combined with

relevant national development strategies and relevant regional policy requirements, the province is divided into 6 districts for comprehensive development and improvement of land, resources and environment [10]. Tudor taking Qingzhou City, Shandong Province as an example, and taking the comprehensive agricultural productivity as the standard, with the help of related models and methods. Divide Qingzhou into different remediation zones in the remediation planning period, such as the recent land remediation area, and put forward the corresponding remediation suggestions for different finishing areas. Landscape-scale land consolidation zoning research, the main focus is on land consolidation projects implemented in specific spatial areas [11]. Zhai *et al.* suggested that on the basis of the division of the northern low mountains and hills and the North China Plain, the first-class project type area framework, the Tianjin land consolidation project type area is divided into 3 secondary project type areas, it is pointed out that the division of this type of area can provide a basis for the arrangement of engineering construction projects, the preparation of feasibility study reports, and engineering design [12]. Ji and Wang presented a study using the natural quality score of the intermediate results of agricultural land classification, and the type of factor combination, as an entry point, the research on the types of agricultural land remediation in Tianjin is carried out, the difficulty level of farmland remediation in different types of districts, the focus of remediation and the direction of future remediation were clarified [13]. Sankararaman *et al.* analyzed the regional differences in the proportion of land use, consolidation potential, and average size of rural settlements, the research on the remediation zones of rural settlements is carried out, the whole country is divided into 5 remediation areas, and the direction of regional renovation according to local conditions is proposed [14].

The model can further be extended by implementing some Machine learning and Artificial intelligence approaches to enhance the capability of cost estimation of reconstruction. Based on the current research, the author proposes a method based on spatial clustering algorithm. Based on GIS methods such as spatial clustering algorithm and least resistance model, typical projects are selected to propose ecological land consolidation planning [15]. Guided by unified engineering construction forms and rigid engineering construction standards, and aiming at regularization and rigidization, traditional land consolidation by means of high-intensity engineering construction, although it has played a positive role in stabilizing the amount of arable land and improving agricultural production conditions [16-17]. However, it is waiting for transformation under the requirements of the current ecological civilization construction. Through the improvement of ecological land, the ecological improvement is realized, and the need for landscape improvement is more urgent. The study is guided by the theory of landscape ecology, according to the "Landscape Pattern Evaluation-functional Zoning of Land Remediation - Optimization of Corridor Pattern-Plaque matrix optimization" general idea.

3 Research methods

This section includes the design process including pattern analysis, functional division and layout designing of the proposed study.

3.1 Landscape pattern analysis

According to the land use status of the project area and the characteristics of the land remediation project, appropriately merge the original land use types, converted into 2 types of matrix landscapes of cultivated land and pits and ponds, there are two types of patch landscapes, construction land and forest land, and three types of corridor landscapes, namely roads, rivers, and ditches (see Table 1). Among them, among the original land types, hardened rural roads with a pavement width of more than 1 m are defined as road corridors, other roads merge into adjacent other landscapes [18]. The landscape scale selects the number of landscape patches, patch density, average patch area, boundary density, coefficient of variation of plaque area, separation index, landscape richness index and Shannon diversity index reflect landscape characteristics; Type scale Select type area, proportion of project area, number of patches, patch density, the area variation coefficient, shape variation coefficient, separation degree and aggregation degree are used to evaluate the characteristics of each type of landscape; For the corridors in the project area, select the corridor length, corridor density, corridor line point rate, corridor connectivity, and corridor circulation to evaluate its current situation; The scale of the patch was evaluated by selecting the patch area, shape index, and perimeter-area ratio. The above indicator calculations are performed using Fragstates software.

Table 1: Land use type and landscape type conversion

Original land use type	Landscape classification	Landscape definition
Paddy field	Arable land	Matrix
Dry land	Arable land	Matrix
Tian Kan	Arable land	Matrix
Pond water surface	Hang pond	Matrix
River surface	River	Corridor
Ditch	Ditch	Corridor
Rural road	The way	Corridor
Village	Construction land	Plaque
Facility farm land	Construction land	Plaque
Garden	Woodland	Plaque
Scenic spots and special land	Woodland	Plaque

i. On the landscape scale, the project area can be divided into 7 types of landscapes. Compared with adjacent areas, the degree of landscape aggregation and spread is higher, lower

separation, lower Shannon diversity index, it shows that the fragmentation degree of various types of landscapes in the study area is relatively low, landscape diversity and heterogeneity are low, and the same type of landscape has obvious aggregation;

ii. On the type scale, the largest proportion of the area, and the type with the largest number of patches is the pond landscape, and the pond and river landscape together accounted for 54.64% of the total area. The area of arable land and pits has the highest variation coefficient, and the area is uneven. The separation degree of cultivated land and pits and ponds is low, and most of them are concentrated and contiguous, however, the separation degree of ditches, roads and woodlands is much higher than other types, and they are scattered;

iii. In terms of corridor layout, the three types of corridors have low density and poor connectivity. The circulation of the river corridor is relatively good, and the channel channel has not yet formed a network structure. Referring to relevant research, make 100m buffer zones on both sides of roads and ditches, calculate its functional coverage, accounting for 28.78% and 13.23% of the total area of the project area, the current roads and ditches are difficult to fully meet the needs of production and life.

iv. On the spot scale, the cultivated map spot is generally small, the shape index and perimeter-area ratio are higher than those of pits, and the aggregation degree is lower, overall finer. The area of the pits varies greatly, but the shape is more uniform. The distribution of other types of spots is more scattered, on the whole, there is a fragmentation trend from southwest to northeast [19].

3.2 Functional division

According to the results of landscape pattern analysis, the study area has obvious spatial differences in landscape characteristics. In order to propose targeted remediation directions in different regions, arrange project types according to local conditions, select landscape type, patch area, perimeter-area ratio and aggregation degree as clustering conditions, using the KNN (k-nearest neighbor) spatial clustering algorithm, the study area was divided into 3 types of areas with similar landscape types, similar shape of the patches, and high degree of aggregation. Considering the integrity of the zoning, combined with the functional positioning of the study area, the results of spatial clustering and partitioning are corrected to obtain the functional partition of land consolidation, that is, the farmland remediation area, the water surface remediation area and the water town style enhancement area, determine the remediation goals, principles and leading project types for each sub-area:

i. In the farmland remediation area, the landscape type is mainly cultivated land. Taking high-standard farmland construction as the remediation goal, based on the principle of ecological maintenance, under the premise of reducing ecological disturbance and ecological damage as much as possible, through land type adjustment, implement land leveling, field roads, irrigation and drainage projects, etc, it is advisable to combine engineering measures such as biological ridge design, ecological pavement and biological passage design of field roads, ecological slope protection design of ditches and artificial wetlands, etc, on the premise of ensuring ecological connectivity and biodiversity, the productivity of cultivated land can be improved, improve agricultural production efficiency and promote large-scale operation [20].

ii. In the water surface improvement area, the landscape types are mainly ponds and rivers. Taking the construction of ecological sources and the construction of aquaculture bases as the rectification goals, based on the principles of ecological restoration and ecological improvement, through the construction of artificial wetlands, the construction of river bank ecological slope protection, and the deployment of ecological combined purification systems, etc, optimize water system structure, improve water quality, and improve production and ecological functions.

iii. The water town style enhancement area has a variety of landscape types. Taking rural landscape construction as the renovation goal, based on the principle of ecological maintenance, on the basis of maintaining the original natural features, layout road works with recreational functions, maximize the landscape richness within the view of the play trail. It is advisable to set up water-town-style hydrophilic platforms, covered bridges, pavilions, etc. in combination with the road layout. Implement rural landscape improvement projects, form a characteristic landscape experience in the south of the Yangtze River of "rice field-flowing water-people's house", and cultivate new formats such as sightseeing and experience agriculture and leisure agriculture.

3.3 Planning layout

3.3.1 Optimization of corridor pattern

The connectivity and circulation of rivers, roads, and ditch corridors in the study area are low, and their functional coverage is small, so structural optimization is required. By building new roads, ditches and farmland shelterbelts, it can meet traffic and irrigation needs, improve farmland ecological environment, and optimize water system structure. This study uses the least

resistance model to generate new corridors, according to the rectification principles of each rectification zone, the corridors are revised and screened, get the corridor optimization results. Because ditches and shelter forests are laid along with the field roads.

3.3.2 Water surface improvement area

The substrate landscape in this area is a pit and pond, and most of them are aquaculture water surfaces with clear ownership. The southwest corner of the region is the industrial enterprise plaque, which is the key target of environmental pollution prevention and control. According to the actual situation of the aquaculture water surface in the area, some suitable pits and ponds can be adjusted to a depth of 3m at the bottom of the pond, the water depth is controlled within 2.5m, biological channels are set at the bottom of the pool, and a buffer zone of native shrubs and grasses is left beside the pool, while ensuring the breeding efficiency, it can ensure the habitat of aquatic animals and plants, and improve the water quality of the fish pond and the ecosystem environment. In order to prevent point source pollution, at the same time, the non-point source pollution caused by agricultural production is eliminated to a certain extent, based on the survey data of the water system flow in the project area, according to the migration principle of nitrogen and phosphorus pollutants in agricultural shallow drainage system, adopting the five-level load reduction governance model of "Riverside Vegetation Buffer Zone - Biological Pool Contact Oxidation Unit - Wetland Interception System - Estuary Ecological Interception - Plant Purification Integrated Technology", form a series of ecological combination purification system [21].

3.3.3 Water Township Enhancement Area

The area is rich in landscape types, surrounded by water, and the landscapes of villages, fields and forests are intertwined. However, due to the scattered distribution of patches in villages and serious domestic waste pollution, the overall appearance is poor, waiting for remediation. According to the area's rectification goals to enhance the style of Jiangnan water towns, maintain the ecological environment, and guide rural leisure tourism, focus on improving facilities, village greening and village beauty, and implement rural landscape improvement projects. In terms of rural greening, different methods such as tree ponds, bamboo fences or cement masonry are used, carry out road greening, increase the greening of houses, courtyards, and corners, greening the nodes such as residential pools and existing open spaces to improve the ecological environment of the village. In terms of improving facilities, increase garbage collection points, considering the current population of the village and the number of recreational populations in the future, set up garbage bins and public health facilities in the village. In terms of village beautification, guide residents to renovate houses and vegetable gardens with characteristics of Jiangnan water towns, restore the visual style of Jiangnan water town from the aspects of color, morphological structure and greening; In addition, appropriately increase leisure and sightseeing facilities, combined with planning and play trails, water bridges, stone arch bridges, wooden octagonal pavilions and hydrophilic platforms will be built, it not only provides residents with a cool and resting place, but also builds a living space with a beautiful environment, highlight the unique charm of Jiangnan water towns and promote the development of rural tourism. For the cultivated land and woodland patches in the region, the original landscape pattern should be preserved, coordinate with the village landscape to form a multi-level rural landscape experience. Guide the development of experience agriculture, picking agriculture and other new formats with land consolidation infrastructure construction, create conditions for expanding rural tourism and farming experience. Through the optimization of the corridor pattern and the optimization of the patch matrix, the ecological land consolidation planning scheme in the project area is formed [22].

4 Experimental results and analysis

In order to verify the feasibility and implementation effect of the scheme, from the two aspects of engineering construction and landscape ecological security, the ecological land remediation planning scheme is evaluated. Through the project planning, 4 new field roads with a total length of 3102m were built; 13 new production roads with a total length of 3403m; 8 new play trails with a total length of 2693m; 2 new Dougou with a total length of 717m; 3 new bucket canals with a total length of 1479m; 5 new agricultural ditches with a total length of 1116m; 8 new agricultural canals with a total length of 1398m; 3743m of new shelter forest; Implement land leveling in 4 areas with a total area of 25.26hm2, accounting for 64.57% of the total cultivated land area; The newly constructed wetland is 1.24hm2, accounting for 2.91% of the total area of the pit. After the renovation, the functional coverage area of the ditches in the project area accounted for 54.02% of the total area of the project area, an increase of 40.79% compared with that before the renovation; The coverage area of road functions accounted for 88.84%, an increase of 60.06%, and the infrastructure in the project area was improved. Through the project planning, the corridor network pattern in the project area has undergone great changes. The calculation results show that, since the planning scheme does not involve changes in the river channel, the river corridor structure remains unchanged. However, through the construction of ecological ditches, the density of ditches and corridors increased by 447.31% compared with that before the renovation, the circulation degree of ditch corridors increased by 114.91%, and the network structure of ditch corridors was significantly improved. Through the construction of ditches, the water system in the project area has been connected, the water system corridor connectivity has increased by 55.43%, and the water system circulation has increased by 454.95%, the water system structure in the project area

has been optimized. In addition, the road corridor pattern in the project area has also been significantly improved [23]. Landscape ecological security Build an ecological security pattern, maintaining the stability of the ecological environment is an important goal of ecological land consolidation planning, combined with the principles of landscape ecology, various types of landscapes in the project area are used as evaluation units, select the landscape ecological security index, ecological security evaluation is carried out before and after the renovation of the project area. The calculation method is as follows:

$$LES = \sum_{i=1}^{n} LES_i \times P_i \tag{1}$$

$$LES_i = 1 - 10 \times U_{istandardized} \times Q_i \tag{2}$$

$$\begin{cases} U_i = a \times C_i + b \times F_i + c \times D_i \\ C_i = N_i / A_i \\ F_i = \sqrt{S_i} / 2P_i \\ D_i = d \times L_i + e \times P_i \\ S_i = N_i / A \\ P_i = A_i / A \\ L_i = N_i / N \end{cases} \tag{3}$$

Where LES is the landscape ecological security index; LES_i is the ecological security index of landscape type i; Ui is the landscape disturbance index; Qi is the landscape vulnerability index; Ci is the landscape type fragmentation, Fi is the separation degree of landscape type, Di is the dominance degree of landscape type, Si is the distance index of landscape type, Pi is the relative cover of the landscape type, Li is the relative density of the landscape type; $a, b, c, d, and\ e$ are the weights, and the research determines that each weight value is 0.5, 0.3, 0.2, 0.4, 0.6; Ni is the number of landscape type patches, N is the total number of landscape patches, Ai is the area of landscape type patches, the unit is hm2, A is the total landscape area, the unit is hm2 [24].

Use Fragstates software to calculate various indicators (Figure 2 to Figure 8), the results show that, the landscape types with the largest landscape ecological security index LESi before and after the renovation were all cultivated land, the smallest is woodland [25]. Except for the river landscape ecological security index, which dropped by 9.84%, the ecological security of other types of landscapes improved, among them, the largest increase was road (121.29%), followed by forest land (43.10%). According to formulas (1) and (2), the landscape safety index LES of the project area is 0.45 before the renovation, and 0.61 after the renovation, an increase of 35.56%, indicating that through ecological land remediation, the overall landscape ecological security of the project area has been improved, but the optimization of river landscape should be strengthened [26].

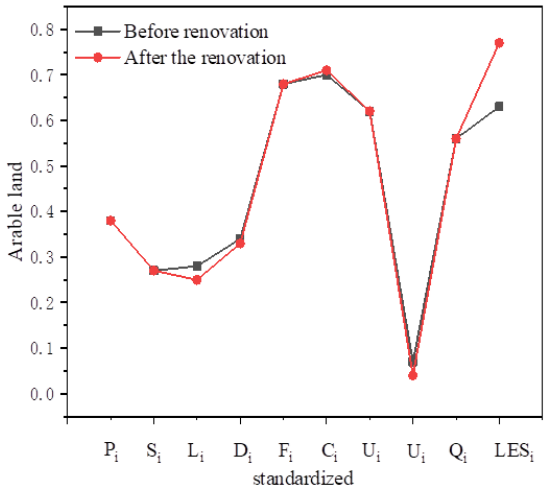


Figure 2: Calculation results of cultivated land safety index.

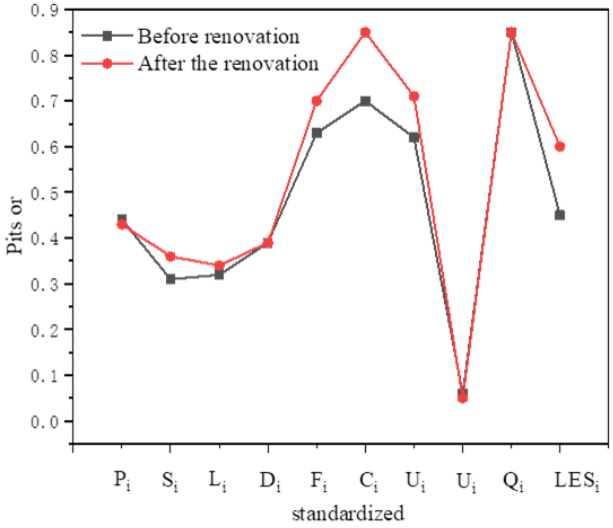


Figure 3: Calculation results of pit pond safety indicators.

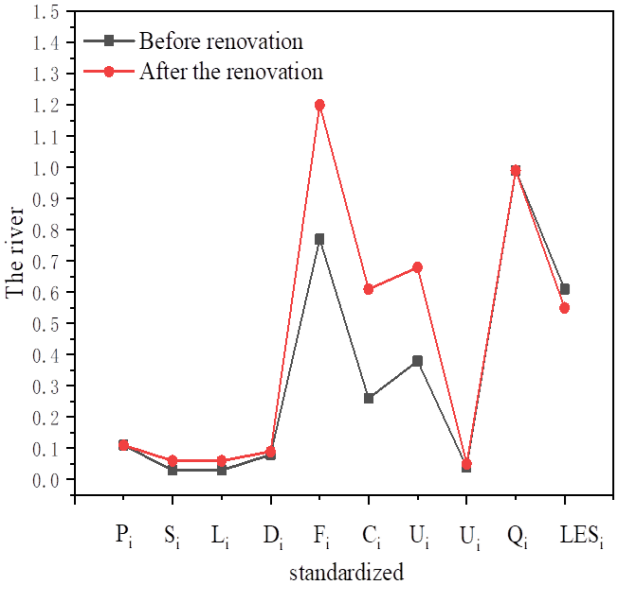


Figure 4: Calculation results of river safety indicators.

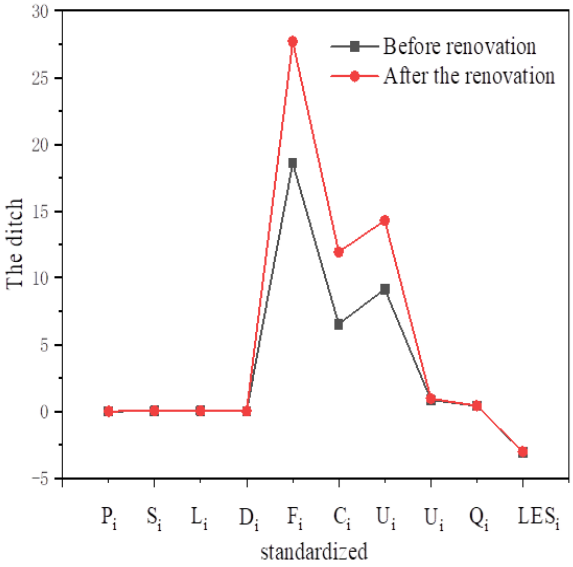


Figure 5: Calculation results of ditch safety indicators.

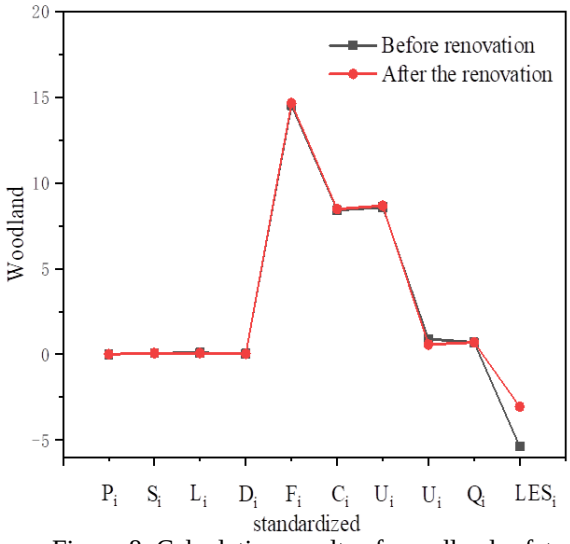


Figure 8: Calculation results of woodland safety indicators.

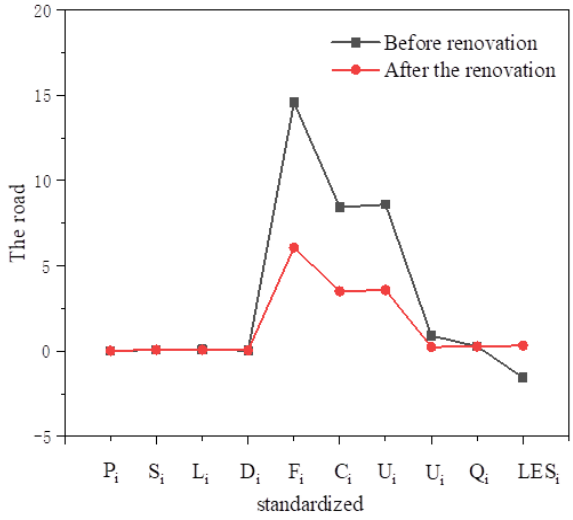


Figure 6: Calculation results of road safety indicators.

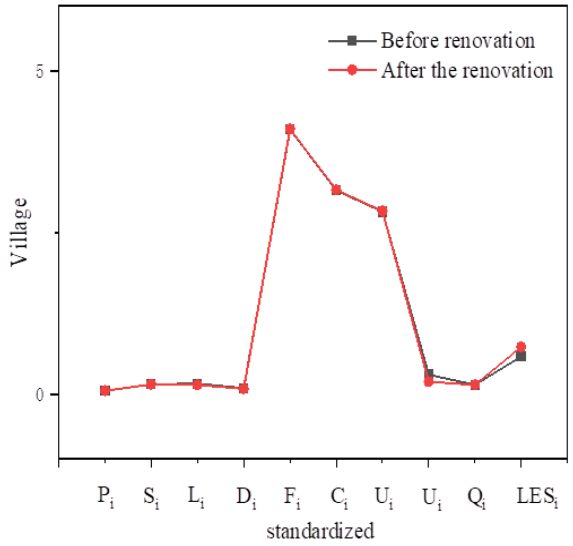


Figure 7: Calculation results of village safety indicators.

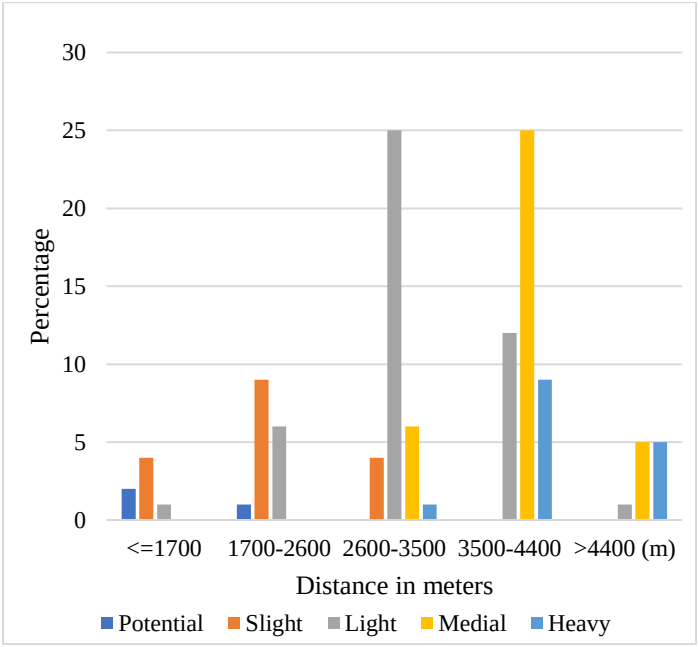


Figure 9: Eco-environmental vulnerability distribution level in different elevation.

Eco-natural weakness in concentrate on region presents distinct distribution of vertical-belt. As depicted in Figure 9, the weakness is connected with height clearly, in which levels III and IV are most broadly disseminated and essentially in rise belt going from 2600 to 4400 m, levels I and II dispersed underneath the rise 2600 m, and level V conveyed over the height 3500 m. Essentially, eco-natural weakness in this district likewise has plainly geological level belt dissemination. As a general rule, the weakness is heavy in the north while it is moderately light in the south.

5 Conclusion

Ecological land consolidation planning should be based on traditional land consolidation. It can be extended in terms of considering goals, modes and methods for landscape pattern analysis. The least resistance model like non-point source pollution control technology and other multi-field methods is considered such as land consolidation zoning. The integration of traditional land remediation planning methods such as road and ditch layout, promotes the realization of multi-functional goals of land remediation. At the same time, considering its agricultural production goals, the status quo of urbanization and industrialization oppression, and the improvement potential of rural landscapes rich in cultural characteristics, determines a reasonable direction of rectification according to local conditions, and combine traditional engineering rectification methods with regional industrial characteristics. The outcomes of this study show that the technique that coordinates the advancements, like RS and GIS, and the SPCA in a numerical way to deal with assess eco-climate weakness in mountainous area, cannot particularly address the input subject spatial circulation of mountain vertical-belt highlight, yet in addition regard the river valley in general framework. Through ecological land consolidation planning, while achieving the goal of improving agricultural facilities. The overall landscape ecological security status of the project area can be improved, and the ecological security indices of other types of landscapes except for rivers are improved.

References

[1] Kwon, S., Kim, S., Tak, O., & Jeong, H. (2017). A study on the clustering method of row and multiplex housing in Seoul using K-means clustering algorithm and hedonic model. *Journal of Intelligence and Information Systems*, *23*(3), 95-118. *https://doi.org/10.13088/jiis.2017.23.3.095*

[2] Xie, J., Yang, G., Wang, G., Song, Y., & Yang, F. (2021). How do different rural-land-consolidation modes shape farmers' ecological production behaviors? *Land Use Policy*, *109*, 105592. *https://doi.org/10.1016/j.landusepol.2021.105592*

[3] Yu, Z., Xiao, L., Chen, X., He, Z., Guo, Q., & Vejre, H. (2018). Spatial restructuring and land consolidation of urban-rural settlement in mountainous areas based on ecological niche perspective. *Journal of Geographical Sciences*, *28*(2), 131-151. *https://doi.org/10.1007/s11442-018-1464-2*

[4] Janus, J., & Taszakowski, J. (2018). Spatial differentiation of indicators presenting selected barriers in the productivity of agricultural areas: A regional approach to setting land consolidation priorities. *Ecological Indicators*, *93*, 718-729. *https://doi.org/10.1016/j.ecolind.2018.05.050*

[5] Moravcová, J., Koupilová, M., Pavlíček, T., Zemek, F., Kvítek, T., & Pečenka, J. (2017). Analysis of land consolidation projects and their impact on land use change, landscape structure, and agricultural land resource protection: case studies of Pilsen-South and Pilsen-North (Czech Republic). *Landscape and ecological engineering*, *13*(1), 1-13. *https://doi.org/10.1007/s11355-015-0286-y*

[6] Cong, L. I., Xiaoyun, L. I., & Wenkang, D. A. I. (2020). Rural Land Consolidation in the Context of Rural Revitalization. *Asian Agricultural Research*, *12*(1812-2021-674), 19-22. *10.22004/ag.econ.310146*

[7] Long, H., Zhang, Y., & Tu, S. (2019). Rural vitalization in China: A perspective of land consolidation. *Journal of Geographical Sciences*, *29*(4), 517-530. *https://doi.org/10.1007/s11442-019-1599-9*

[8] Gong, Y., & Tan, R. (2021). Emergence of local collective action for land adjustment in land consolidation in China: An archetype analysis. *Landscape and Urban Planning*, *214*, 104160. *https://doi.org/10.1016/j.landurbplan.2021.104160*

[9] Guo, B., Su, Y., Pei, L., Wang, X., Zhang, B., Zhang, D., & Wang, X. (2020). Ecological risk evaluation and source apportionment of heavy metals in park playgrounds: a case study in Xi'an, Shaanxi Province, a northwest city of China. *Environmental Science and Pollution Research*, *27*(19), 24400-24412. *https://doi.org/10.1007/s11356-020-08744-x*

[10] Ng, Y. X. (2015). A study of urban heat island using" local climate zones"-the case of Singapore. *British Journal of Environment and Climate Change*, *5*(2), 116-133.

[11] Tudor, M. M. (2015). Small scale agriculture as a resilient system in rural Romania. *Studies in Agricultural Economics*, *117*(1), 27-34. *http://dx.doi.org/10.7896/j.1503*

[12] Zhai, H., Zhang, H., Zhang, L., Li, P., & Plaza, A. (2016). A new sparse subspace clustering algorithm for hyperspectral remote sensing imagery. *IEEE Geoscience and Remote Sensing Letters*, *14*(1), 43-47. *10.1109/LGRS.2016.2625200*

[13] Ji, J., & Wang, K. L. (2014). A robust nonlocal fuzzy clustering algorithm with between-cluster separation measure for SAR image segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(12), 4929-4936. *10.1109/JSTARS.2014.2308531*

[14] Sankararaman, A., Vikalo, H., & Baccelli, F. (2020). ComHapDet: a spatial community detection algorithm for haplotype assembly. *BMC genomics*, *21*(9), 1-14. *https://doi.org/10.1186/s12864-020-06935-x*

[15] Han, B., Jin, X., Xiang, X., Rui, S., Zhang, X., Jin, Z., & Zhou, Y. (2021). An integrated evaluation framework for Land-Space ecological restoration planning strategy making in rapidly developing area. *Ecological Indicators*, *124*, 107374.

https://doi.org/10.1016/j.ecolind.2021.107374

[16] Demetriou, D. (2013). *The development of an integrated planning and decision support system (IPDSS) for land consolidation*. Springer Science & Business Media.

[17] Liu, J., Jin, X., Xu, W., Sun, R., Han, B., Yang, X., & Zhou, Y. (2019). Influential factors and classification of cultivated land fragmentation, and implications for future land consolidation: A case study of Jiangsu Province in eastern China. *Land Use Policy*, *88*, 104185. *https://doi.org/10.1016/j.landusepol.2019.104185*

[18] Wu, Y., Feng, W., & Zhou, Y. (2019). Practice of barren hilly land consolidation and its impact: A typical case study from Fuping County, Hebei Province of China. *Journal of Geographical Sciences*, *29*(5), 762-778. *https://doi.org/10.1007/s11442-019-1626-x*

[19] Wu, X., & Zhang, X. (2019). An efficient pixel clustering-based method for mining spatial sequential patterns from serial remote sensing images. *Computers & Geosciences*, *124*, 128-139. *https://doi.org/10.1016/j.cageo.2019.01.005*

[20] Zhong, W., Tan, D., Peng, X., Tang, Y., & He, W. (2018). Fuzzy high-order hybrid clustering algorithm for swarm intelligence sets. *Neurocomputing*, *314*, 347-359. *https://doi.org/10.1016/j.neucom.2018.03.019*

[21] Jiang, L., Ji, J., Lu, Y., Chen, Y., & Jia, Y. (2019). Mathematical modeling and simulated annealing algorithm for spatial layout problem. *Cluster Computing*, *22*(3), 6383-6391. *https://doi.org/10.1007/s10586-018-2137-8*

[22] Unternährer, J., Moret, S., Joost, S., & Maréchal, F. (2017). Spatial clustering for district heating integration in urban energy systems: Application to geothermal energy. *Applied energy*, *190*, 749-763. *https://doi.org/10.1016/j.apenergy.2016.12.136*

[23] Shang, R., Tian, P., Jiao, L., Stolkin, R., Feng, J., Hou, B., & Zhang, X. (2016). A spatial fuzzy clustering algorithm with kernel metric based on immune clone for SAR image segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(4), 1640-1652. *10.1109/JSTARS.2016.2516014*

[24] Jiao, L. M., Zhang, X., & Mao, L. F. (2015). Self-organizing dual spatial clustering algorithm and its application in the analysis of urban sprawl structure. *Journal of Geo-Information Science*, *17*(6), 638-643. *10.3724/SP.J.1047.2015.00638*

[25] Loehlein Fier, H., Prokopenko, D., Hecker, J., Cho, M. H., Silverman, E. K., Weiss, S. T., ... & Lange, C. (2017). On the association analysis of genome-sequencing data: a spatial clustering approach for partitioning the entire genome into nonoverlapping windows. *Genetic epidemiology*, *41*(4), 332-340. *https://doi.org/10.1002/gepi.22040*

[26] Shen, L., Guo, J., & Wang, L. (2018). A Self-Organizing Spatial Clustering Approach to Support Large-Scale Network RTK Systems. *Sensors*, *18*(6), 1855. *https://doi.org/10.3390/s18061855*

Application of Adaptive Artificial Bee Colony Algorithm in Reservoir Information Optimal Operation

Lijun Cui[1,2*]
[1]Liaoning Open University, School of Public Administration, Shenyang 110034, China
[2]Liaoning Equipment Manufacturing Vocational and Technical College, School of Public Administration, Shenyang 110161, China
E-mail: lijuncui5@163.com
[*]Corresponding author

Keywords: artificial bee colony algorithm, optimal scheduling, cascade hydropower stations, comprehensive learning, multi strategy

Received: February 22, 2022

Abstract: The hydrothermal scheduling is complex issue of nonlinear optimization consisting of several constraints that plays a critical role in the operations of power system. In order to meet the safe operation of hydropower stations, how to reasonably dispatch them to achieve the best comprehensive benefits is one of the main problems in the hydropower industry. Artificial bee colony algorithm has the advantages of simple structure and strong robustness. It is widely used in many engineering fields. However, the algorithm itself still has many shortcomings. Based on the current research, an improved artificial colony algorithm based on standard artificial bee colony algorithm is proposed, and the performance of the algorithm is verified in three benchmark functions and three cec213 test functions. Compared with many well-known improved algorithms, it is proved that the improved procedure has greatly enhances the final solution accuracy and convergence outcome. The experimental outcomes observed by improved artificial algorithm are compared with adaptable artificial bee colony procedure and chaotic artificial bee colony procedure and with other existing works in the literature. It is observed from the experimentation that the proposed algorithm performs better in comparison with established optimization algorithms.

Povzetek: Raziskava obravnava optimizacijo razporeditve vodne energije s pomočjo izboljšanega algoritma umetne čebelje kolonije, kar izboljša natančnost in konvergenčne rezultate.

1 Introduction

Power generation, flood control, water storage and irrigation are the goals that many cascade hydropower stations need to coordinate, but there are complex constraints among them. In solving the scheduling scheme of single reservoir hydropower station, the conventional scheduling method is a feasible solution. However, cascade hydropower stations are a high-dimensional, nonlinear and multi constraint problem. When solving this issue by conventional dispatching methods, the calculation efficiency is usually low, and it is easy to drop into problems such as native optimization and “dimension disaster” [1]. Therefore, it is an important issue to combine the process and dispatching issue of cascade reservoirs with emerging calculation methods (swarm intelligence optimization algorithm and constraint processing method). The research on the theory and method suitable for the optimal operation of cascade reservoirs (groups) is conducive to shorten the gap between the application of the theory and method to practical projects. China's hydropower reserves rank first in the world, but the per capita is scarce. However, China is also one of the countries with more flood disasters. On the premise of having a set of better reservoir operation scheme, in the face of flood problem, reservoir is an important protective measure, which can store and discharge water artificially and reduce the loss caused by flood disaster [2]. Therefore, the solution of the operation and operation scheme of cascade reservoirs (groups) is a issue to be solved for a long time at present and in the future. In the past two decades, with the development of computers, many swarm intelligence algorithms have been proposed by the majority of Chinese and foreign scholars. These algorithms are proposed based on the behavioral characteristics of organisms in the biological world, such as foraging, mating and so on. Social animals are more complex work. Just like this feature, many swarm intelligence algorithms can solve complex practical problems [3]. The artificial bee colony (ABC) algorithm is proposed in 2005 through the observation of bee colony honey collection behavior. The algorithm is to find the optimal solution of the issue by the process of bee colony cyclic iteration. Global search and native development exist in the optimization process at the same time, which is a main feature of the algorithm, as shown in Figure 1.

Much exploration was done to address the downsides of established procedures and enhances the adequacy to determine hydrothermal scheduling in potential frameworks. Wavelet transmission was joined with artificial neural network to work on the capacity of the system in [4, 5]. The genuine coding method was practically implemented to work on looking through capacity of genetic algorithm. Equal computation was

embraced to further develop the particle swarm optimization algorithm. The proposed strategy presents a superior scheduling result in comparison with the existing technique. The chaotic searching method has been embraced by researcher and applied in differential evolution to determine the issue of dynamic dispatch issue.

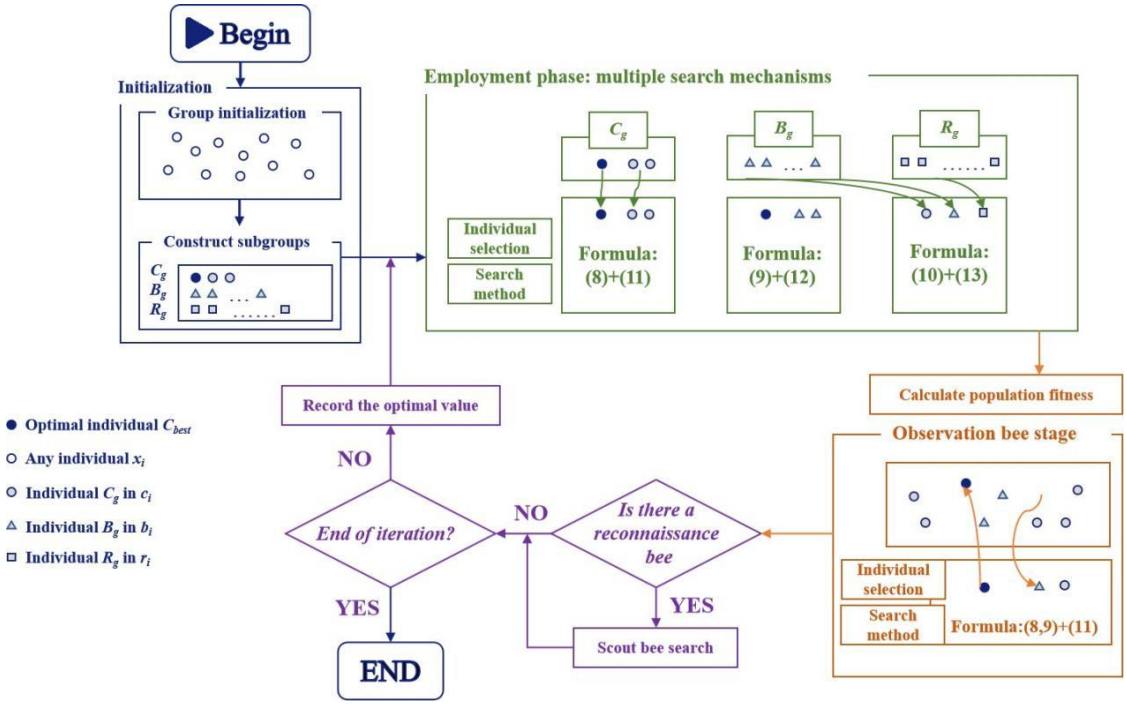


Figure 1: Architecture of artificial bee colony algorithm.

Effectiveness and searching capability of differential evolution was improved in chaotic differential advancement algorithm with an improved outcome. The self-versatile idleness weight instrument was joined with the particle swarm optimization algorithm to decide the age timetable of a flowed hydro framework situated at Narmada stream in India. The outcomes exhibited the adequacy of the planned strategy. These adjustments have worked on different parts of optimization capacity. Though, because of the additional boundaries, the hybrid algorithms might turn out to be more confounded and unsteady.

Recently, another enhancement strategy named artificial bee colony algorithm turns out to be increasingly more adopted due to the fact that it performs strong capacity to deal with different engineering designing issues. The artificial bee colony has a basic design and great union capacity. Though, the coordination of control boundaries is a difficult and uphill errand. Then again, comparable with other current heuristic algorithms, artificial bee colony algorithms additionally have the issue of untimely convergence. Consequently, in this article, we offer a method for an adaptive chaotic artificial bee colony. Control boundary setting technique and chaotic neighborhood search are applied to assist the algorithm with getting away from nearby optimal solution. Additionally, unique adaptation of artificial bee colony did not reflect the obliged optimization issue. Consequently, we propose another limitation taking care of strategy with an adjusted form of artificial bee colony to determine hydropower scheduling with different imperatives. The outcomes show that compared to current methods, the suggested algorithm offers a higher convergence capacity and convergence speed.

This article's remainder is structured as follows: Section 2 displays the recent work done in hydropower scheduling. Section 3 consists the information about the research methods and the algorithmic steps based on chaotic learning. Section 4 deals with the results and examinations. Section 5 describes the concluding remarks.

2 Literature review

The research on optimal operation of reservoir groups in China began in the late 1970s [6]. Ma *et al.* applied stochastic dynamic programming theory to Lancang River reservoir group operation and established an optimization model for dynamic regulation of reservoir information [7]. Jie *et al.* applied the parallel dynamic programming method to the operation of cascade hydropower stations in Lixianjiang River Basin [8]. Mansouri and others established the mathematical model of reservoir group optimal operation for the issue

that the convergence of coordination vector and feedback vector is not synchronized in solving the optimal operation of reservoir group by large-scale system objective coordination method, and improved the solution process of real-time optimal operation combined with accelerated convergence technology [9]. With the deepening of artificial intelligence theory, many intelligent optimization algorithms have been applied to cascade reservoir operation. Aiming at the disadvantages of slow convergence speed and easy to fall into local optimization in the later stage of ant colony algorithm, Gavahi *et al.* gradually determined the scope of the optimal solution by increasing the tabu search area, proposed an ant colony algorithm based on tabu search, and applied the improved algorithm to the optimal operation of Zhangze cascade reservoir [10]. Jamshidi and Shourian improved the standard genetic algorithm by using three elite evolution strategies of floating-point coding, adaptive crossover rate and mutation rate, and applied it to the optimal operation of Fengjiashan and Duanjiaxia cascade reservoirs. As a new swarm intelligence algorithm, artificial bee colony algorithm has significant advantages in dealing with the optimization of high-dimensional functions [11]. Bozorg Haddad *et al.* have shown that compared with ant colony optimization (ACO), frog leaping algorithm (FLA), particle swarm optimization (PSO), differential evolution algorithm and other well-known evolutionary algorithms, the performance of artificial bee colony algorithm is better or equivalent [12]. The research on artificial bee colony algorithm is still in its infancy, and the theory and practice are not perfect. Many scholars have always been committed to improving the performance of the algorithm. Barz *et al.* used chaos strategy and reverse learning strategy to make the initialization of individuals in the solution space purposeful and average the distribution of initial solutions as much as possible. However, the chaotic strategy makes the distribution of individuals have direction, which weakens the global examination enactment of the algorithm to a certain extent [13]. Babu and Kumari proposed a RABC procedure. The population alternately adopts two evolutionary policies with Rosenbrock rotation disturbance to dynamically adjust the search balance between the global and local of the algorithm. Shrabbhal401 draws lessons from the concept of native optimal specific in PSO procedure to make the parent data of honey resource participate in the generation of candidate solutions, so as to improve the local development ability of the algorithm [14]. Aiming at the phenomenon of dropping into native optimization due to the reduction of population diversity, Gao *et al.* proposed a MABC algorithm. The fixed parameter p is used instead of observing the selection probability of bee roulette, and the improved algorithm has achieved good outcomes [15]. Gupta *et al.* improved the artificial bee colony algorithm with the idea of chaotic driving, so that the offspring coefficients are generated by the parent coefficients using the chaotic strategy, which greatly improves the local search ability of the algorithm [16].

3 Research methods

This section includes the discussion of proposed flowchart for addressing the issue of hydro scheduling. The algorithm steps are described in this section.

The standard algorithm simulates the honey collection process of bees, and solves the issue through division of labor, cooperation and information sharing among different individuals. Among them, bee colony is composed of employed bees, onlookers and scout bees. The degree of perfection is dependent on the fitness specified by the optimization issue, and honey source symbolizes a workable solution. Hire bees to search in the global scope. Once a high-quality honey source is found, they fly back to the hive to inform other bees. The observation bee selects a honey source and searches nearby according to the information transmitted by the hired bee. Initialization: assume that the total number of bee colony members for the d-dimensional issue is Sn, the amount of honey bases is FN, SN = 2FN, and the number of hired bees and observation bees are FN. FN initial honey sources are generated from the search space through Equation (1), and each honey source position is a D-dimensional vector, which is recorded as x_{ij}.

$$x_{i,j} = x_{\min,j} + rand(0,1)\left(x_{\max,j} - x_{\min,j}\right) \tag{1}$$

Population update: employ bees to update the position according to Equation (2). Compared with the original honey source, greedy selection strategy is used to record the location of higher quality honey source.

$$v_{i,j} = x_{i,j} + \phi_{i,j}\left(x_{i,j} - x_{k,j}\right) \tag{2}$$

Where, the random number in $k \in \{1,2,\ldots,FN\}$ and $I \neq K, \emptyset_{i,j}$ represents the random quantity between [- 1, 1], x_{ij} represents the original honey source location, $x_{k,j}$ represents the neighborhood honey source location, and $v_{i,j}$ represents the updated honey source location. The flowchart of the proposed algorithm is depicted in Figure 2. The prime focus of the proposed algorithm is to address the issue of hydro scheduling.

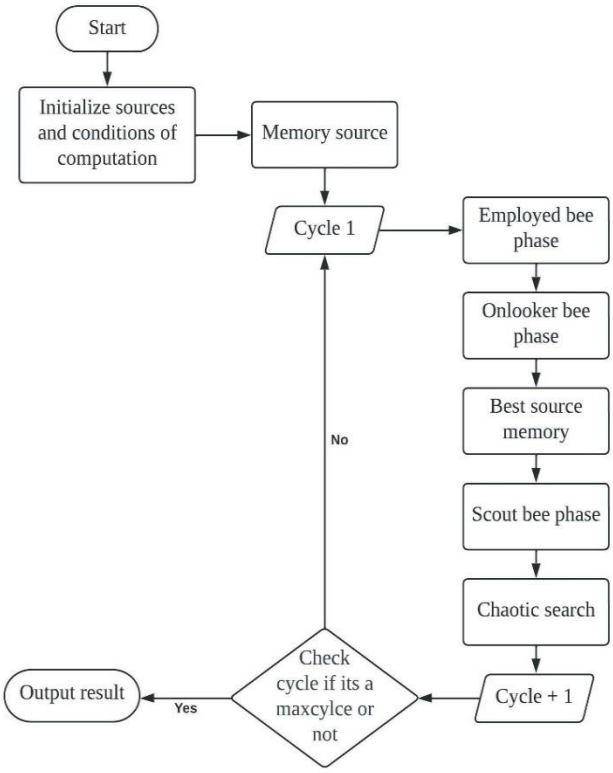


Figure 2: Process flow of proposed algorithm.

Step 1: Determine the swarm size Sn, the variety of sources for honey FN, the size D of the solution to the optimization issue, and set the limit times imit. FN solutions are randomly generated to form the initial honey source position, and the fitness of each honey source is calculated.

Step 2: Employ bees to update according to Equation (2) and evaluate the adaptability of new honey sources. Keep the excellent honey source for the next iteration.

Step3: Select the honey source according to roulette, let the observation bee update according to Equation (2), evaluate the fitness of the new honey source and select the best.

Step 4: If the location of a honey source does not change and is not the current global optimum after the search of limit iterations, the hired bee converts a reconnaissance bee, and in accordance with equation (1), a fresh source of honey is created at random [17].

Step 5: If the specified number of iterations is reached or the global optimal solution meets the needs of the problem, terminate the algorithm. Otherwise, return to Step 2.

Aiming at the shortcomings of slow convergence speed and low convergence accuracy of standard bee colony algorithm, an artificial bee colony algorithm with improved special center (ISC-ABC) is proposed based on the narrow center idea of particle swarm optimization algorithm. First, improve the narrow central thought. By comparing the fitness, the better honey source is selected to form an improved narrow center, and the greedy strategy is compared with the current global optimal position to lead the bee colony to converge. Secondly, change the update strategy of the original bee colony [18]. The employment bee always searches around the current global best, strengthens the ability of the bee colony to develop hidden solutions near the best, and improves the accuracy of the algorithm solution. The special center is formed by the central position of the population and changes in real time with the population movement. The narrow sense center, like other individuals, has the attributes of location update, fitness evaluation and so on. It is found that due to the special position, the narrow center tends to the optimal solution of the issue more than the global optimal position. For every iteration, the narrow center is equated with the current global optimum, and a better position is certain to lead the bee colony movement, so as to speed up the population convergence. The narrow sense center expression is as follows:

$$c = \frac{\sum_{i=1}^{FN} xij}{FN} \tag{3}$$

Where c is the narrow central position in the population, and FN represents the number of honey sources in the population.

In ISC-ABC algorithm, employed bees and observed bees adopt different search strategies [19]. The current global optimal information is added to the update formula to improve the local development ability of the algorithm near the current global optimal; The observation bee update formula remains unchanged to ensure the global search ability of the algorithm. The evolutionary strategy adopted by employing bees and observing bees ensures that the scope of the search of population is mainly around the current global optimum, reduces the search of a large number of random positions, and improves the convergence speed of the algorithm; It also has a certain global development ability, which can maintain the diversity of the population and reduce the possibility of the population falling into local optimization. The employment update formula is as follows:

$$v_{i,j} = gbest_j + \phi_{i,j}\left(gbest_j - x_{k,j}\right) \tag{4}$$

Where, $k \in \{1,2,\ldots,FN\}$ is a random number and $i\#k$, $\emptyset_{i,j}$ represents a random number between [0,1], $w_{i,j}$ represents the original honey source location, gbest represents the current global optimal location, x_k represents the neighborhood honey source, and $v_{i,j}$ represents the updated honey source location. At present, the global optimality is a key position. In the iterative process, all bee individuals are guided to approach the optimality, which directly affects the overall convergence speed of the algorithm and the quality of the final solution. With the progress of search, the current global optimum will be closer and closer to the position of the theoretical optimal solution [20]. The update of the current global optimum in the standard bee colony only depends on the improvement of the extreme value after

the individual evolution of the population. When the extreme value does not change or changes very little after individual evolution, the current global optimum will not change greatly. ISC-ABC algorithm guides the convergence direction of the population by introducing the current global best, ensures that the employed bees gradually approach the current global best, and increases the convergence capability of the procedure [21].

Step 1: Determine the colony size Sn, the amount of honey sources FN, the dimension D of the optimization problem's solution, and set the limit number. FN solutions are randomly generated to form the initial honey source location, and the fitness of each location is estimated.

Step 2: Determine the improved narrow sense center of the current iteration and make greedy selection with the current global optimal.

Step 3: Update according to Equation (4) to evaluate the fitness of the new location. Keep the excellent honey source for the next iteration.

Step 4: Select the honey source according to the way of roulette, let the observation bee update according to Equation (4), evaluate the fitness of the new location, and select the best.

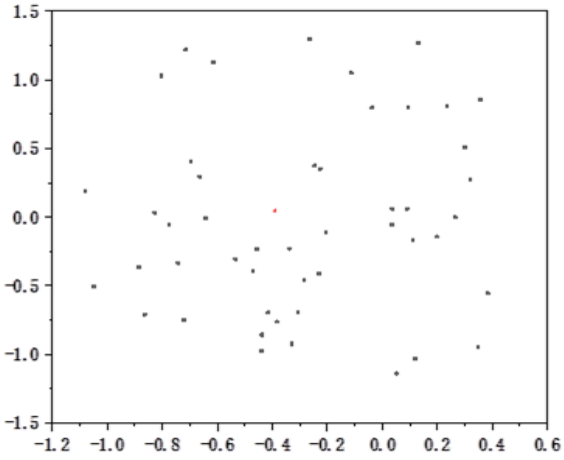
Step 5: If the location of a honey source does not change and is not the current global optimum after the search of limit iterations, the employed bee converts a reconnaissance bee, and according to Equation (3), a fresh source of honey is created at random.

Step 6: If the specified number of iterations is reached or the global optimal solution meets the needs of the problem, terminate the algorithm. Otherwise, return to step 2.

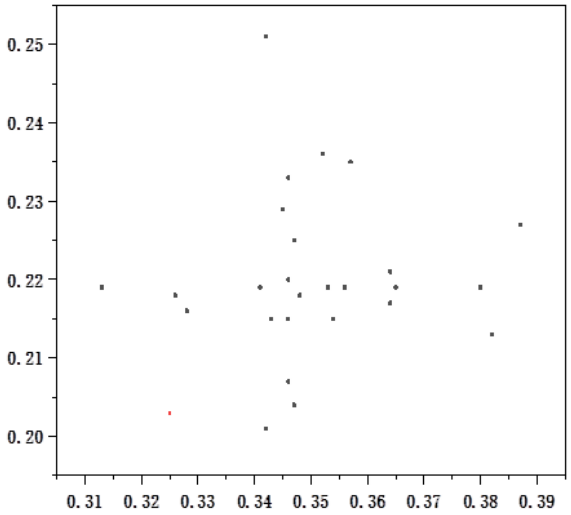
4 Results and analysis

This segment describes the examination of outcomes attained from the proposed algorithm for addressing the issue of hydro scheduling.

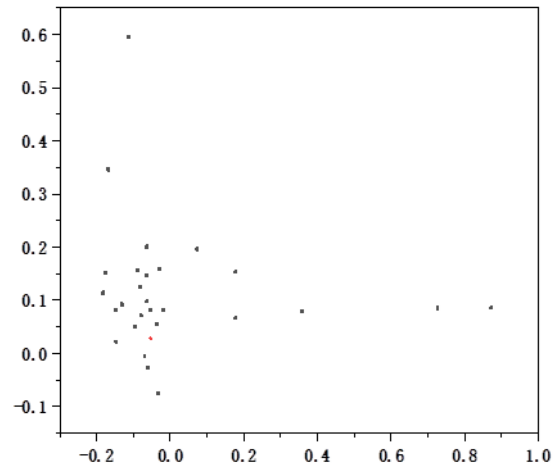
In order to show the population distribution and convergence in different iterative periods, Quatic function is used for testing. Figure 2 and Figure 3 are the population individual motion scatter diagrams of 0ABC algorithm and ISC-ABC algorithm respectively. The dot represents the position of the individual in the population, and the asterisk represents the position of the theoretical optimal solution.



(a): FEs=0

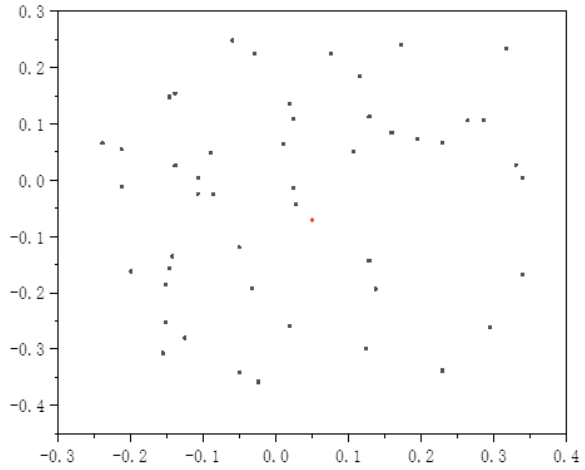


(b): FEs=30000

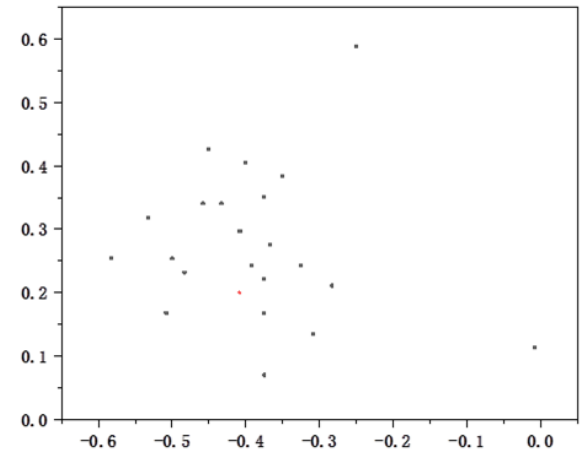


(c): FEs=60000

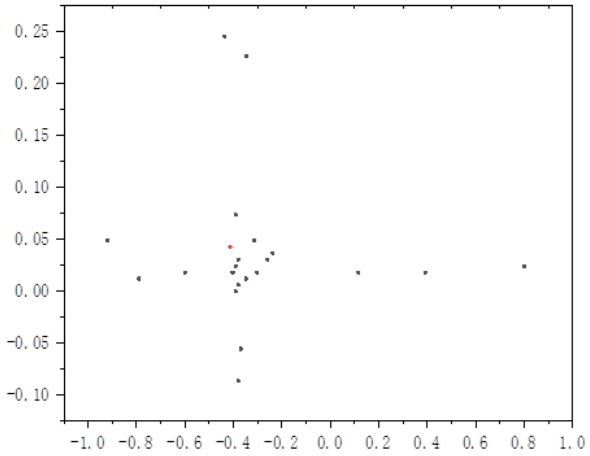
Figure 3: Feature extraction under normal conditions for several iterations.



(a): FEs=0



(b): FEs=30000



(c): FEs=60000

Figure 4: Feature extraction with improved narrow center strategy for several iterations

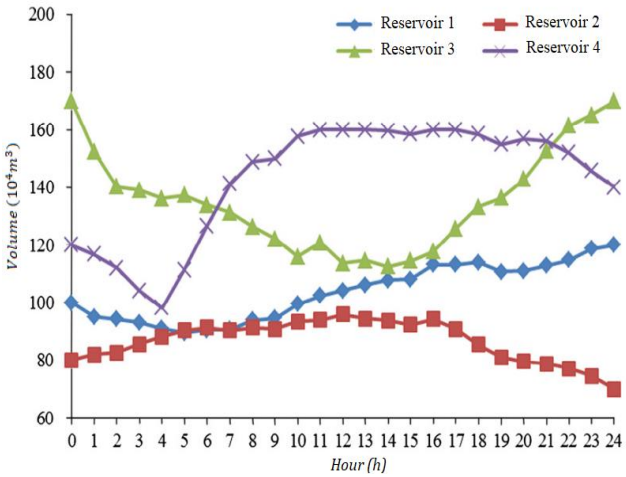


Figure 5: Ideal reservoir storing standards of case 1.

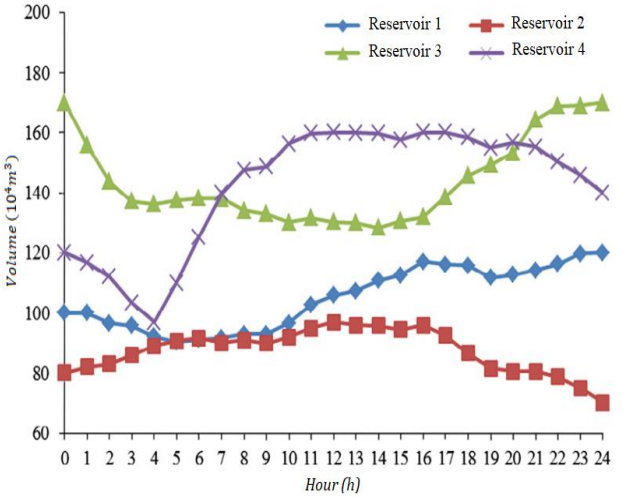


Figure 6: Ideal reservoir storing standards of case 2.

Table 1: Comparison of Friedman test outcomes

Algorithm	Rankings
ISC-ABC	2.12
MABC	2.75
MEABC	2.75
GABC	3.45
ABC	4.37

It can be seen from the comparison between Figure 2 and Figure 3. In the early stage of the algorithm (the number of iterations is before 60000), the improved narrow center strategy can significantly accelerate the convergence speed of the population and quickly find the region where the optimal solution is located; In the middle and later stages of the algorithm (after 60000 iterations), the improved narrow sense center can be used as the advanced part of the global optimum. It can generate a location closer to the optimal solution than the current best advantage, and may also produce a far away

from the current optimal location. The swarm moves around the optimal solution in a large range to enhance the individual's ability to explore the solution space. The diversity of individual distribution is maintained under the condition of ensuring the convergence speed, which reduces the possibility of the population falling into local optimization. With the progress of the search and the continuous change of the population activity area, the current global best point and the improvement of the individual are closer and closer to the optimal solution. Finally, the final solution is obtained to complete the search.

The simulation outcomes are observed for two cases. In case 1 the power transmission loss is not considered whereas it is considered in case 2. The optimal capacity volumes of the hydro capacity are depicted in Figure 5. It is observed from Figure 4 that every one of the imperatives were fulfilled, and that implies every one of the solutions acquired by the proposed technique are practical. The loss of power transmission has been considered in second case. The outcomes got by various techniques were recorded and the reservoir supply volume for case 2 is depicted in Figure 5.

Then the Friedman examination is implemented to analyze the outcomes of the test function, which can effectively distinguish the performance gap between procedures. The smaller the Friedman value, improved the algorithm performance. According to Table 1, ISC-ABC algorithm has the smallest Friedman value and the best comprehensive performance compared to the algorithms ABC, GABC, MABC, and MEABC.

5 Conclusion

Aiming at the shortcomings of slow convergence speed and low final solution accuracy of standard bee colony algorithm, an improved narrow center bee colony algorithm is constructed by referring to and improving the narrow center idea of particle swarm optimization algorithm. First, improve the narrow central thought. Select the better honey source to form an improved narrow center to lead the bee colony to converge. Secondly, change the original bee colony update strategy. Make the hired bee search around the current global optimum, and strengthen the ability of the bee colony to develop hidden solutions near the best. Finally, the simulation test outcomes of three classical benchmark functions and three cec213 functions show that compared with a variety of similar algorithms, the artificial bee colony algorithm based on improved narrow center guided convergence has significantly improved in convergence speed and resolution accuracy, indicating the effectiveness of the improved method.

References

[1] Mat, A. N., İnan, O., & Karakoyun, M. (2021). An application of the whale optimization algorithm with Levy flight strategy for clustering of medical datasets. *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, *11*(2), 216-226. *https://doi.org/10.11121/ijocta.01.2021.001091*

[2] Ahmad, A., Razali, S. F. M., Mohamed, Z. S., & El-Shafie, A. (2016). The application of artificial bee colony and gravitational search algorithm in reservoir optimization. *Water Resources Management*, *30*(7), 2497-2516. *https://doi.org/10.1007/s11269-016-1304-z*

[3] Takano, R., Sato, H., & Takadama, K. (2019). Artificial bee colony algorithm based on adaptive local information sharing meets multiple dynamic environments. *SICE Journal of Control, Measurement, and System Integration*, *12*(1), 1-10. *https://doi.org/10.9746/jcmsi.12.1*

[4] Bhardwaj, C., Jain, S., & Sood, M. (2021). Two-tier grading system for npdr severities of diabetic retinopathy in retinal fundus images. *Recent Patents on Engineering*, *15*(2), 195-206. *https://doi.org/10.2174/1872212114666200109103922*

[5] Bhardwaj, C., Jain, S., & Sood, M. (2021). Hierarchical severity grade classification of non-proliferative diabetic retinopathy. *Journal of Ambient Intelligence and Humanized Computing*, *12*(2), 2649-2670. *https://doi.org/10.1007/s12652-020-02426-9*

[6] Mansouri, A., Aminnejad, B., & Ahmadi, H. (2018). Introducing modified version of penguin's search optimization algorithm (PeSOA) and its application in optimal operation of reservoir systems. *Water Science and Technology: Water Supply*, *18*(4), 1484-1496. *https://doi.org/10.2166/ws.2017.217*

[7] Ma, S., Hu, X., Zhao, Y., Wang, X., & Dong, C. (2021). Design and Evaluation of a Metal-Supported Solid Oxide Fuel Cell Vehicle Power System with Bioethanol Onboard Reforming. *ACS omega*, *6*(43), 29201-29214. *https://doi.org/10.1021/acsomega.1c04698*

[8] Jie, L. I. U., & Xi-huang, Z. H. A. N. G. (2016). Adaptive differential evolution artificial bee colony algorithm based on segmental-search strategy. *Computer and Modernization*, (9), 15. *10.3969/j.issn.1006-2475.2016.09.004*

[9] Mansouri, A., Aminnejad, B., & Ahmadi, H. (2018). Introducing modified version of penguin's search optimization algorithm (PeSOA) and its application in optimal operation of reservoir systems. *Water Science and Technology: Water Supply*, *18*(4), 1484-1496. https://doi.org/10.2166/ws.2017.217

[10] Gavahi, K., Mousavi, S. J., & Ponnambalam, K. (2019). Adaptive forecast-based real-time optimal reservoir operations: application to Lake Urmia. *Journal of Hydroinformatics*, *21*(5), 908-924. https://doi.org/10.2166/hydro.2019.005

[11] Jamshidi, J., & Shourian, M. (2019). Hedging rules-based optimal reservoir operation using bat algorithm. *Water Resources Management*, *33*(13), 4525-4538. https://doi.org/10.1007/s11269-019-02402-9

[12] Bozorg-Haddad, O., Janbaz, M., & Loáiciga, H. A. (2016). Application of the gravity search algorithm to multi-reservoir operation optimization. *Advances in water resources*, *98*, 173-185. https://doi.org/10.1016/j.advwatres.2016.11.001

[13] Barz, T., Sommer, A., Wilms, T., Neubauer, P., & Bournazou, M. N. C. (2018). Adaptive optimal operation of a parallel robotic liquid handling station. *IFAC-PapersOnLine*, *51*(2), 765-770. https://doi.org/10.1016/j.ifacol.2018.04.006

[14] Babu, C. S., & Kumari, V. V. (2017). An Adaptive Optimized Frequent Itemset Mining in Large Databases by means of FPL and Adaptive Artificial Bee Colony (AABC) Algorithm. *http://nopr.niscair.res.in/handle/123456789/43201*

[15] Gao, W. F., Huang, L. L., Liu, S. Y., & Dai, C. (2015). Artificial bee colony algorithm based on information learning. *IEEE transactions on cybernetics*, *45*(12), 2827-2839. *10.1109/TCYB.2014.2387067*

[16] Gupta, M., Kundu, A., & Gupta, V. (2017). Multi-Objective Artificial Bee Colony Algorithm for Multi-Echelon Supply Chain Optimization Problem: An Indian Case Study. *International Journal of Operations Research and Information Systems (IJORIS)*, *8*(4), 76-89. *10.4018/IJORIS.2017100105*

[17] Momtaz, M. I., Amarnath, C. N., & Chatterjee, A. (2020, November). Concurrent error detection in embedded digital control of nonlinear autonomous systems using adaptive state space checks. In *2020 IEEE International Test Conference (ITC)* (pp. 1-10). IEEE. *10.1109/ITC44778.2020.9325229*

[18] Saeed, S., Ong, H. C., & Sathasivam, S. (2019). Self-adaptive single objective hybrid algorithm for unconstrained and constrained test functions: An application of optimization algorithm. *Arabian Journal for Science and Engineering*, *44*(4), 3497-3513. *https://doi.org/10.1007/s13369-018-3571-x*

[19] Rabenberg, A., Schulte, T., Hildebrandt, H., & Wehling, M. (2019). The FORTA (Fit fOR The Aged)-EPI (epidemiological) algorithm: application of an information technology tool for the epidemiological assessment of drug treatment in older people. *Drugs & Aging*, *36*(10), 969-978. *https://doi.org/10.1007/s40266-019-00703-7*

[20] Chen, G., Liu, G., Wang, J., & Li, R. (2012). Identification of water quality model parameters using artificial bee colony algorithm. *Numerical Algebra, Control & Optimization*, *2*(1), 157. *10.3934/naco.2012.2.157*

[21] Gu, W., Yu, Y., & Hu, W. (2017). Artificial bee colony algorithmbased parameter estimation of fractional-order chaotic system with time delay. *IEEE/CAA Journal of Automatica Sinica*, *4*(1), 107-113. *10.1109/JAS.2017.7510340*

LOCUS: A Mobile Tourism Application and Recommender System for Personalized Places and Activities

Duaa AlSaeed
Information Technology Department, College of Computer and Information Sciences, King Saud University, 11451, Riyadh, Saudi Arabia
E-mail: dalsaeed@ksu.edu.sa

Keywords: recommender systems, mobile applications, item-item collaborative filtering, user-user collaborative filtering, tourism

Received: August 23, 2022

The tourism industry is all around keeping tourists happy, occupied and equipped with the things they need during their time away from home. On the other hand, mobile technologies have a considerable impact on user experience, particularly in the tourist and entertainment areas. This paper presents a tourist and entertainment-related mobile application. It utilizes a personalized experience approach and seeks to provide good user experiences by making it adaptable to their unique interests while considering many criteria such as the user's gender, age, location, and other characteristics. The system will propose locations to visit or activities to do in any city to the user. As the user continues to use the application, the suggestions offered will constantly be improved; it will learn more about the user's preferences by recording the user's past and what they enjoy. The application implements and integrates two types of recommender systems, the item-item collaborative filtering algorithm and the user-user collaborative filtering algorithm. The user acceptance testing was conducted on 10 users from a variety of backgrounds and ages. Each participant has performed a set of 17 asks that covers the functionality of the application. Effectiveness results showed that about 70% of the tasks were completed without errors by all participants. And the tasks that were completed with some errors had an average of errors ranges from (0 - 0.4) which is a promising result when compared to the normal average number of errors per which is 0.7. Regarding the efficiency, results show that the longest completion time was in 3 tasks (register task, log-in, and edit profile) which is expected since they require the entry of detailed information. On the other hand, for the remaining tasks the average completion time was 5.4s which is accepted. User satisfaction was measured through a System Usability Scale (SUS) survey, the achieved score was 87.75 which is higher than the threshold to pass the SUS test which is 68, thus LOCUS has fulfilled the user satisfaction measure.

Povzetek: Članek predstavi izvirno mobilno aplikacijo za turizem in zabavo, ki zagotavlja personalizirano uporabniško izkušnjo glede na interese uporabnika, se nenehno uči in izboljšuje.

1 Introduction

The tourism industry is linked to the idea of people travelling to other locations, either domestically or internationally, for leisure, social or business purposes. It is one of the fastest growing global industries [1]. It is closely connected to the hotel industry, the hospitality industry which is a broad category of fields within the service industry that includes food and drink service, event planning, theme parks, and transportation [2]. The tourism industry is based around keeping tourists happy, occupied and equipped with the things they need during their time away from home. On another hand, mobile technologies have recently evolved to significantly influence the user experience especially in the tourism and the entertainment domains [3].

Numerous IT solutions have emerged to enhance the tourism industry. Recommender systems, utilizing personalized recommendations [1], [4], [5], and Natural Language Processing (NLP), which can be used to extract important keywords and phrases or conduct sentiment analysis of tourist reviews and comments [6], are among the different technologies leading to varied solutions. These solutions can enhance the tourism experience and meet the growing demand for personalized and efficient services.

Also, with the increase in usage of social media and social networking applications, companies and business owners are relying more on these platforms to advertise and promote their business and services for the aim of attracting more customers [7].

The rapid growth in the tourism field and the entertainment industry led to a huge demand for technological solutions and has motivated developers to create mobile applications that provide innovative solutions and services that will make entertainment and tourism more pleasing and satisfying for people. One of the areas of innovative solutions is to enhance and facilitate the process of finding and deciding the appropriate places/activity that matches one's needs and preferences.

This paper proposes a mobile application related to the tourism and entertainment fields. The application will use a personalized experience manner and aims to present good user experiences by making it customizable to their specific interests taking into consideration several factors such as the user's gender, age, location, and more. The system will provide the user with recommendations for places to visit or activities to do in any city (including restaurants, cafes, etc.). The suggestions provided will always be enhanced as the user keeps using the application, it will learn more about the user's preferences by tracking the user's history and what they like.

In the context of the current tourism industry, there is a clear need for a reference framework for mobile applications that focus on personalized recommendations. Without such a framework, businesses may struggle to keep up with the evolving needs of tourists and may not be able to offer the high-quality experiences that are now expected in the industry. By providing a reference framework for tourism-related mobile applications, our proposed solution can help businesses stay ahead of the curve and improve their overall user satisfaction, retention, and revenue.

Furthermore, in the Saudi Arabian tourism industry, there are no current attempts to develop a specialized tourism application with recommender systems. This further emphasizes the need for such applications in the local tourism sector, especially given the increasing demand for personalized experiences among tourists.

The rest of the paper is organized as follows. Section 2 presents a brief background of recommender systems. Section 3 introduces the work that has been done in the field of applications using recommender systems. Section 4 we present our proposed solution LOCUS mobile application. Section 5 presents our evaluation results and the discussion. In section VI. We conclude the output of this study.

2 Recommender systems

"Recommendation system" refers to a system that is able to predict the future preference of a group of items for the user, so it aims to suggest relevant items to users and tries to make predictions on user preferences and make recommendations that should interest the users [2], [3]. Recommendation systems are present in most successful internet companies such as Google, Netflix and YouTube [7]. The reason is that users expect targeted marketing, so providing the same notification or the same offers to everyone is no longer working. Recommendation systems are used in a variety of fields and examples of their uses are: the generators for videos and music playlists used in YouTube and Spotify, suggestions for purchasing products that the user may prefer in Amazon, or suggestions for reading content of user interests such as the recommendation system in Twitter [7].

Lately, different approaches have been developed to build recommendation systems, which can use either collaborative filtering, content-based filtering, or hybrid. Figure 2.2 illustrates the different approaches of recommender systems.

Some of the most well-known websites like Amazon, YouTube, and Netflix use one or more (hybrid) of these recommendation techniques [2], [7].

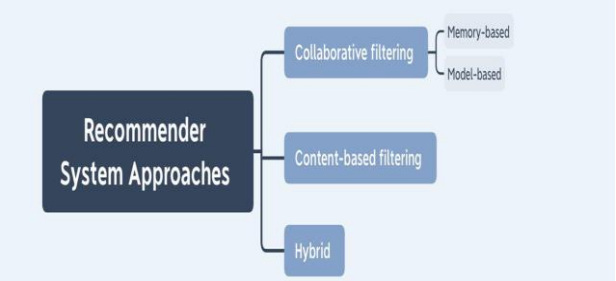


Figure 1: Recommender system approaches.

2.1 Collaborative filtering

This filtering methodology is usually based on collecting and analyzing information on user's behaviors, their activities or preferences and anticipating what they will like depends on the similarity with other users. Collaborative filtering assumes that individuals who concurred in the past will concur later on, and that they will like similar sorts of items as they preferred previously. Further, collaborative filtering algorithms can be categorized into two types: memory-based and model-based [2].

a) Memory-based algorithm.

Another name of the algorithms of memory-based is lazy recommendation algorithms (similar to the k-nearest neighbor method) [2]. There are two different memory-based algorithms:

- User-user collaborative filtering: it finds similar users and recommends items based on items that other similar users previously preferred.
- Item-item collaborative filtering recommends

items similar to those previously favored by the user.

b) Model-based algorithm.

The model-based algorithm is considered as a probabilistic approach, so it is about envisioning the future from learning from the past. The models are developed using different algorithms to predict the user rating of unrated items based on the user's previous ratings[2].

2.2 Content-based filtering

The concept of content-based filtering is that if a user likes an item, the user will also like a comparable item. The filtering methodology depends on the description of an item and a profile of the user's favored choices, so it compares the content to find items that are similar to those previously perused or preferred by the user and determine the similarity between the items from the keywords that are extracted from the item content descriptions [2], [8].

2.3 Hybrid recommendation system

The concept of hybrid recommender systems is combining more than one recommendation technique with one another. Hybrid recommendation systems can be implemented by generating content-based and collaborative predictions separately and then merging them to produce the output. Figure 2 below illustrates the way a hybrid recommender system works.

The accuracy of the hybrid recommendation system is usually higher than the accuracy of any individual ones so combining collaborative and content-based recommendations may be more effective and powerful [2].

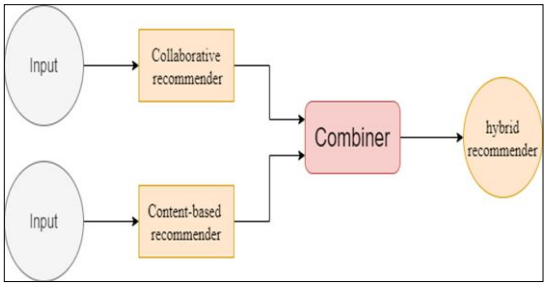


Figure 2: Illustration of hybrid recommender

3 Related work

Several studies have proposed recommender systems (RS) in the tourism domain. Where its one of the ways to deal with deal with the information overload problem. This section contains a review of recent studies that used a mobile recommender system.

Soha et al. (2016) [8] proposed a context-aware recommender system, which recommends places to users based on the current weather, the time of the day, and the user's mood. Places are suggested based on what other users have visited in the similar context conditions. proposed system puts rates for each place in each context for each user. The rates of a location are calculated using the Genetic algorithm, which is based on the Gamma function.

Li et al. (2010) [9] propose a location-aware recommender system for tourism mobile commerce service platform. They integrate GPS into the recommender system to create a location-aware recommender system. The location-aware recommendation system can recommend attractions to the customer with the customer's sensitivity to location and the rating of its attractions.

Abu-Issa et al. (2017) [10] implemented a proactive, multitype, and context-aware recommender application in the environment of Internet of Things (IoT) for smart cities. The recommendation system has the ability to suggest multi-types of services (such as Restaurant, Attraction... etc) and proactively pushes explicit query suggestions to users. The application was developed on Android platform and tested by 50 users. The results shows that the application reach 91.2 % accuracy.

Missaoui et al. (2019) [4] proposed a mobile recommender system for tourism and travel related services called LOOKER. It was developed for the Android platform. It considers basic contextual information, such as location and time. Also, its implements based on content-based filtering (CBF) strategy to make personalized suggestions based on user-related tourism-based user-generated content (UGC) spread across social media. User studies have been conducted to evaluate the usability and the usefulness of the proposed application. A first user study was conducted using two popular questionnaires (SUS and CSUQ) to test the usability of the proposed application with a group of users in four major Tunisian cities. A second user study allowed for a quantitative evaluation of the effectiveness of the proposed system's recommendations. The positive results achieved show the potential of LOOKER. Where, the LOOKER's usability and the recommendations made have been judged to be satisfactory by users.

In conclusion, the reviewed studies have explored the use of recommender systems in the

tourism industry to improve the user experience of mobile applications. Recommender systems have been proposed as a way to address the problem of information overload and provide personalized recommendations to users based on their location, context, preferences, and other factors. These systems have shown promise in improving user satisfaction and retention, as well as increasing revenue for businesses. While there is a body of literature on the topic, there is still room for innovation and unique approaches to design and implementation of recommender systems for tourism-related mobile applications.

We also believe that it is important to highlight the need for the development of a personalized and context-aware mobile application for tourism in Saudi Arabia, which could potentially address the unique needs and preferences of tourists visiting the country. Up to our knowledge, there is a lack of existing tourism-related mobile applications in Saudi Arabia. This further emphasizes the importance of developing a special tourism mobile application with a recommender system tailored to the country's specific tourism industry.

4 Proposed LOCUS mobile application

The rapid growth of the entertainment field in the world and the huge demand on the technological solutions has motivated developers to create a lot of mobile applications. Unfortunately, most applications are using the browsing feature without giving personalized suggestions. Furthermore, these mobile applications require some effort from the user since they depend on filtering things manually to get a list of suggestions.

To enhance and facilitate the process of finding and deciding the appropriate event/activity that matches ones needs and preferences, comes the idea behind developing a mobile application related to the entertainment field that satisfies the users' requirements. The application will use a personalized experience manner and aims to present good user experiences by making it customizable to their specific interests. It will use and integrate recommender system models to provide suggestions for possible activities / based on user's preferences, characteristics, and behavior which will be learned while using the application. The suggestions provided will always be enhanced as the user keeps using the application, it will learn more about the user's preferences by tracking his/her history and what he/she likes. Furthermore, the application will provide information about each activity or place such as pictures, reviews, and location.

The following subsections will discuss the proposed mobile application in detail, covering data sources, architecture, main features, design and development.

4.1 Data sources

For the proposed LOCUS mobile application, real-time and real-world data will be used. The data will be collected mainly from APIs such as google places API [11]. Seasonal or temporary events information will be entered manually from trusted media sources, applications, and websites. Then, the collected information will be stored on a server-based database and will be retrieved upon request from the application. Figure 2 below shows how in general APIs are used to collect data.

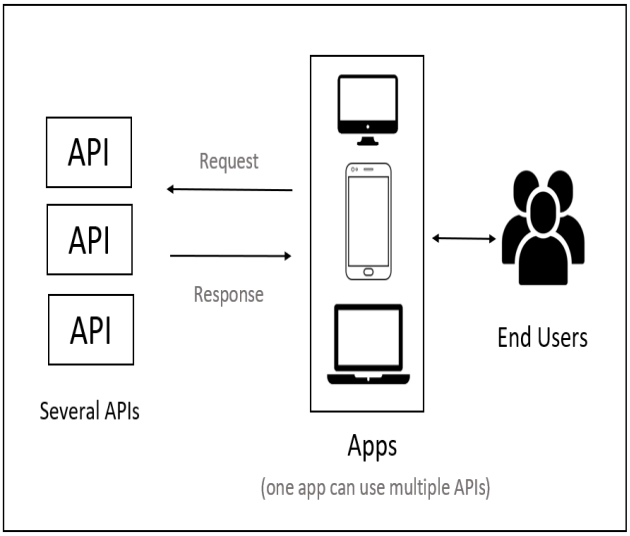


Figure 2: An illustration of how apis are used for collecting data.

4.2 System architecture

The LOCUS system is a client-server architecture in order to enable different users from many different locations to access the shared data that our application provides at any time. Client-server architecture is a shared architecture system in which a group of clients' requests services, a network that allows clients to access these services, and a set of servers that provide services. So, we will implement client-server architecture to fulfill requests by responding to them with the requested services as shown in Figure 3:

LOCUS system architecture Figure 3 below

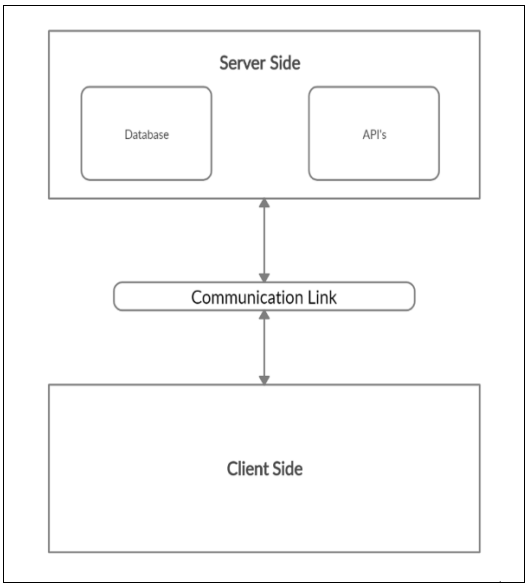


Figure 3: LOCUS system architecture.

On the client side of LOCUS system are the mobile phones that users run the application on, while the server side includes the following:

a) A database to store users' information and manually added activities or seasonal events.

b) The APIs, the application will use multiple APIs to provide information about places, description of each place and its location. Also, the APIs offer some services we want to use such as the nearest places.

4.3 User characteristics and roles

LOCUS application targets the users with the age range from teenagers to adulthood, including both genders and who have interests in finding entertainment activities in Riyadh, whether they are citizens or tourists. there are no restrictions on the education level as long as they have enough knowledge about using smartphones and are able to understand the English language.

4.4 Main features

It has general features such as: add an activity to the favorite list, view ratings, and view reviews/comments. Also, It has customized features such as search by name, Filter the search, and the content. LOCUS uses collaborative filtering (CF) as a recommender system approach. In Figure 4, a use-case diagram is presented, which shows the main features provided by the system and are represented as actions that can be performed by the user.

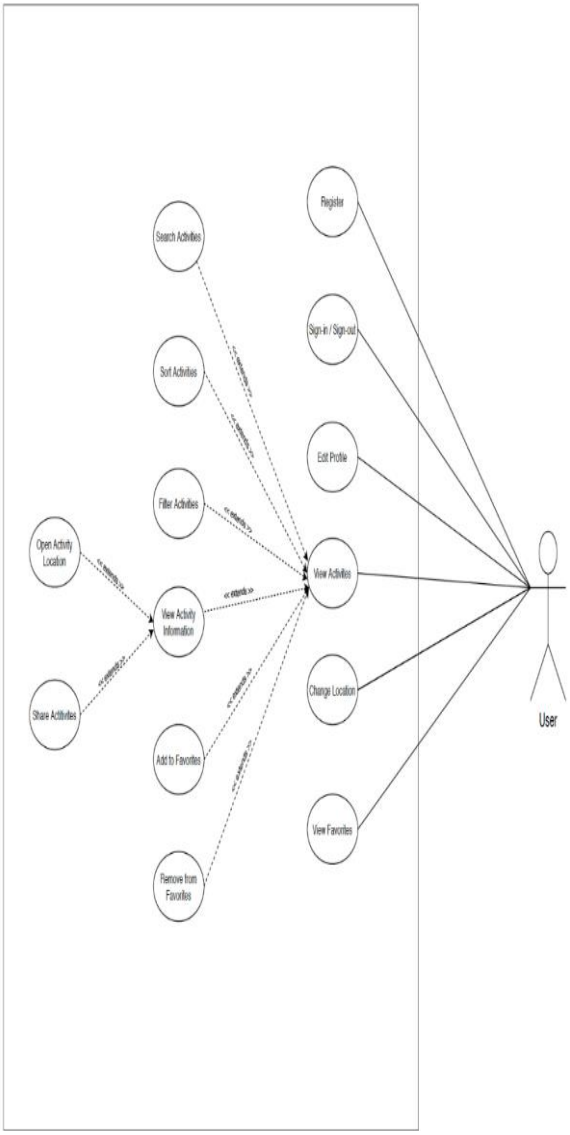


Figure 4: Use case diagram for LOCUS mobile application.

4.5 Implementation

LOCUS application is developed to run on iOS [12] and Android [13] platforms, using React Native framework [14] which uses JavaScript language. In addition, the Firebase database [15] was used to store the data. For the user recommendation functions such as view recommended activities based on user interests or similar users, a module was developed using Python programming language to compute the recommendation algorithms then the Flask Library [16] provided by Python which transforms the code into an API then Heroku [17] was used to deploy it to make the API accessible for LOCUS platform. After

that, the APIs in our React Native application were used using HTTP request calls.

4.6 Algorithms for developed recommenders

As mentioned earlier the proposed system will implement and integrate two types of recommender systems, the item-item collaborative filtering algorithm and the user-user collaborative filtering algorithm. In the following subsection, we will discuss the implemented algorithm for both modules.

c) Item-Item collaborative filtering algorithm

An item-item collaborative filtering recommendation system builds a model to provide a suggested activity based on the user's interest. The model predicts the items, or in this case activities, that the user might be interested in based on computing the similarities between the activity and the user's interests using the vector similar might be interested in based on computing the similarities between the activity and the user's interests using the vector similarity (cosine) and its result will be a value between 0 and 1, the value closer to 1 represents the similarity, and the value closer to zero represents the opposite(cosine) and its result will be a value between 0 and 1, the value closer to 1 represents the similarity, and the value closer to zero represents the opposite [18].

$$s(a,u) = \sum_{j \in 1} \frac{v_{aj}}{\sqrt{\sum_{k \in I_a} v_{ak}^2}} \frac{v_{aj}}{\sqrt{\sum_{k \in I_a} v_{ak}^2}} \quad (1)$$

$$similarity = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}} \quad (2)$$

Where a means the activity vector and b the user vector.

In our application, we give each activity and user a vector representing a number of features (category, key words of activity features such as Wi-Fi, credit card, outside seating, ... Etc.) Those features will be extracted from the activity information and for the user it will be extracted from the user's (interests, viewed activities and favorite list). Our system will compute the cosine vector similarity between the activity and user vectors, then test if it passes a 0.5 threshold to decide if it shall recommend it to the user [18]. Figure 5 shows a flowchart for the developed module for item-item recommender.

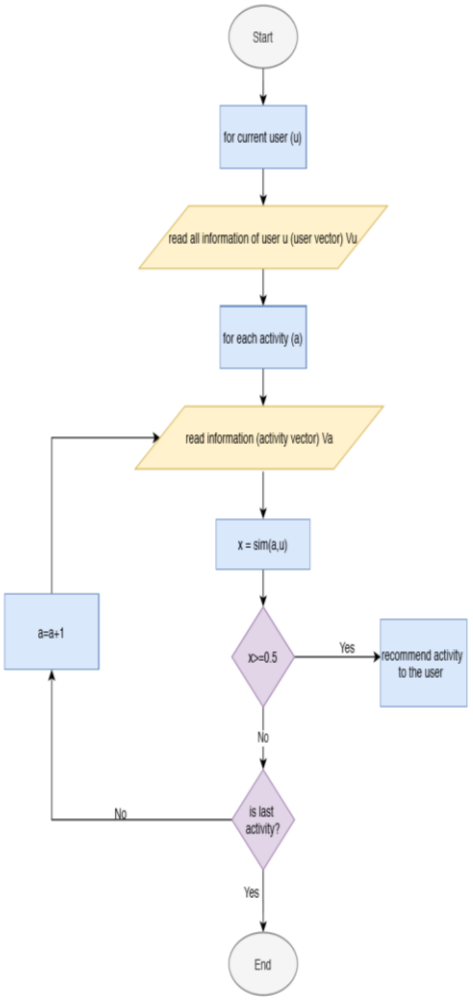


Figure 5: A flowchart for the developed item-item recommender.

d) User-user collaborative filtering algorithm

A user-user collaborative filtering recommendation system builds a model to find similar users and suggest items (activities) based on items (activities) that similar users previously preferred. In our case first, we will select all the users that have the same gender, city and age group, then we will measure the similarity between those users' preferred activities and the user's preferred activities and we will assume that the preferred activities are user's (interests, viewed activities and favorite list) using the vector similarity (cosine) where a means the users' activities (activity) vector and b the user vector using same technique in the item-item collaborative filtering that we mentioned previously [18]. Figure 7 shows a flowchart for the developed module for user-user recommender.

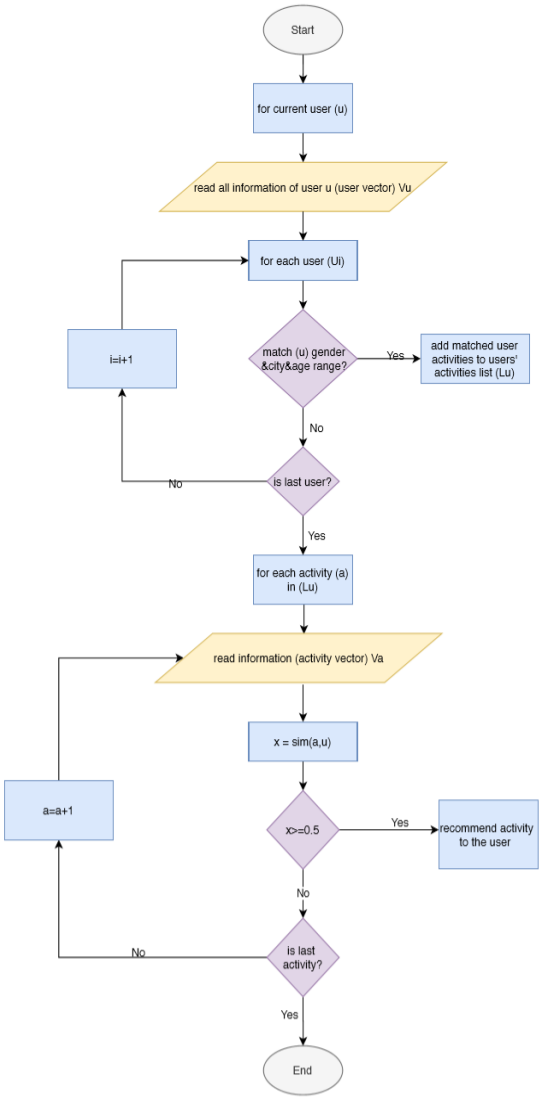


Figure 7: A flowchart for the user-used developed recommender module

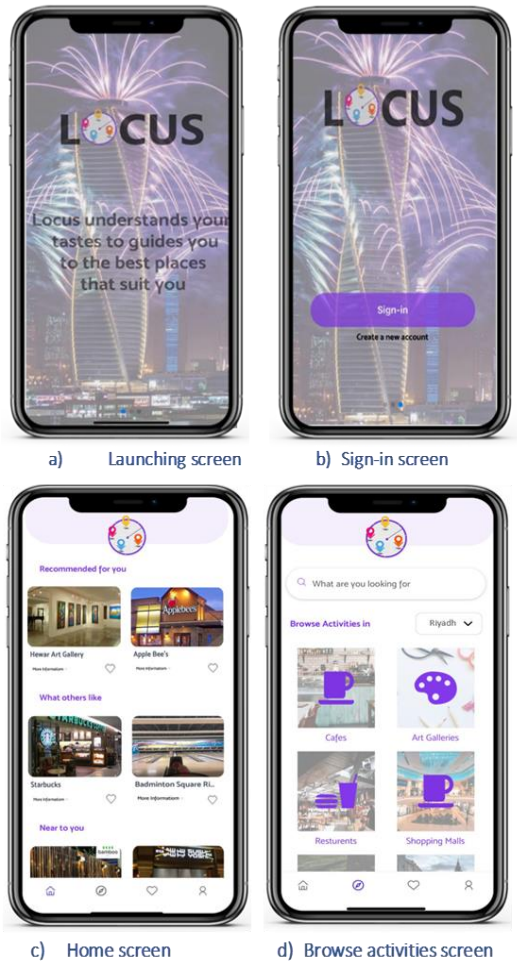


Figure 6: Sample screens of LOCUS mobile application.

4.7 Application layout and user interface

While developing LOCUS, maintaining the ease of use with the functionality of the proposed application were a priority. This section provides a detailed description of the application's user interface. The visual presentation of the user interface (UI) elements has a great impact on the user experience of a product\application. For a good user experience, the content components should be well organized to help people navigate easily within a product or interact with it properly. The main screen of LOCUS gives the user access to the main functions of the application. Screen shots of some of the main screens in the application are shown in Figure 6Figure 6 below.

5 User acceptance testing and discussion

User Acceptance Testing (UAT) is an important step in the software testing process where actual users test the software to assure that functions work as expected and the application can process the tasks in real-world scenarios [19]. In addition, a user acceptance study serves as a proof-of-concept that demonstrates the applicability and effectiveness of the proposed reference framework.

In this work, the user acceptance testing was conducted on 10 users from a variety of backgrounds and ages. The test measured the system usability in three criteria [20]:

- Effectiveness: counting the number of errors, the participant makes when attempting to complete a task.

- Efficiency: calculating the time that participant takes to successfully complete a task.

- Satisfaction: using System Usability Scale (SUS) form to filled by the user, which includes 10 questions to evaluate user satisfaction.

The following subsections describe the conducted user acceptance testing, including the participants' information, experiment design, and testing results.

5.1 Participants

For the user acceptance testing in this study, the participants were a random sample of 10 people covering the characteristics of the targeted users. Their ages are between 20 to 32 with an average of age 24 years. In this study, only nine females and a male volunteered to participate. All participates can understand the basic English language, and capable of using smartphones. Table 1 below shows the demographics of the participants.

Table 1: Participant's demographics

Participant number	Age	Gender	Can understand basic English language	Capable of using smart phones
1	23	Female	Yes	Yes
2	23	Female	Yes	Yes
3	25	Female	Yes	Yes
4	20	Male	Yes	Yes
5	32	Female	Yes	Yes
6	24	Female	Yes	Yes
7	23	Female	Yes	Yes
8	25	Female	Yes	Yes
9	22	Female	Yes	Yes
10	23	Female	Yes	Yes

5.2 Experiment design

The user acceptance testing was conducted with the participants. participants tested the application using an iOS iPhone prepared for the testing (Have an internet connection, Locus app downloaded). In addition, different data collection methods were used to collect quantitative and qualitative data. These methods included pre-questionnaires, direct observation during the test sessions, and post-questionnaires.

- Pre-questionnaire: A pre-questionnaire was used to collect data on the participants. Participants were asked about the places they visit often. Also, if they are facing any difficulties finding suitable places with full information about the places. The pre-questionnaire also included a question about what is the features they would like to have in the recommendation system that could help them to determine their destination easily?
- User acceptance testing sessions: During the UAT sessions, several types of data were collected while observing the participants performing a total of 17 tasks. A timer was used to record the completion time for each task, also we took notes of their mistakes and difficulties faced by the participants during performing the tasks and their feedback.
- Post-questionnaire: After performing the tasks for the LOCUS application, we requested the participants to fill in a postquestionnaire that includes questions to assess their satisfaction. As well as gathering their feedback and suggestions for improvement.

5.3 Testing results and discussion

Each participant has performed a set of tasks that covers the functionality of the application. For each participant, quantitative measures were recorded, this includes: the number of errors and completion time for each task. The list of the 17 tasks performed by participants and a summary of users' performance results are shown in Table 2. Next, the results for effectiveness, efficiency and user satisfaction will be discussed.

- **Effectiveness**: To measure the effectiveness (task completion), we counted the number of participants who completed the tasks without errors and the number of participants who completed the tasks with errors, and the average number of errors made by those participants for each task. Around 70% of the tasks were completed without errors by all participants. And the tasks that were completed with some errors had an average of errors ranges from (0 - 0.4), According to MeasuringU Research Centre [21], the normal average number of errors per task is 0.7,the achieved results are promising since the highest average number of errors for a task is 0.4 and the average number of errors for all tasks is 0.059 and all are below 0.7 which is a good indicator of the effectiveness of the system. Figure 8 shows a summary of number of errors results.

- **Efficiency**: In order to evaluate the efficiency was measures by the time taken to complete a task by participants. The completion time for each task for each participant was recorded and analyzed, the maximum, minimum, and average time per task were computed. Figure 9 shows a summary of completion time results. The results show that the longest time spent was in 3 tasks register task, log-in, and edit profile which is expected since these tasks require the user to enter detailed information, looking into the remaining tasks the average completion time was 5.4s which is accepted.

- **User satisfaction**: To evaluate the user satisfaction of LOCUS app all 10 participants have filled the System Usability Scale (SUS) survey which has 10 questions and results were collected and analyzed. Participants ranked each of the 10 questions from 1 to 5, based on their level of agreement. To calculate the user satisfaction score using SUS we performed the following steps:
    - For each of the odd numbered questions, subtract 1 from the score.
    - For each of the even numbered questions, subtract their value from 5.
    - Add up the total score. Then multiply the value by 2.5.

The SUS score for LOCUS was found to be 87.75, and the threshold to pass the SUS test is 68 according to MeasuringU Research Centre [22] which means that LOCUS app has fulfilled the satisfaction measure.

below shows the survey's measuring scale and results. After completing the survey, we asked them about their opinions and suggestions regarding Locus application interface, most of them (87%) agreed on the necessary of the following:

- Developing an Arabic version interface.
- Separating the filter and sort buttons for better accessibility.
- Adding more user interaction such as adding reviews and ratings.

Those suggestions will be considered in the development of the next version of LOCUS as a future work.

Table 2: Summary of user testing results

Measure\ Task	**Effectiveness**			**Efficiency | Task completion time (seconds)**		
	# of users complete the task without errors	# of users complete the task with errors	Average of errors	Max.	Min.	Average
Register	6	4	0.4	120.0	30.0	59.7
Sign-in	8	2	0.2	59.6	16.7	34.9
Edit profile	9	1	0.1	38.0	6.2	19.3
View categories	10	0	0	17.4	3.0	7.7
View activities based on category	10	0	0	7.4	1.0	4.3
Filter activities	9	1	0.1	27.4	5.6	14.3
Search for activities	10	0	0	15.6	3.1	9.4
Sort activities	8	2	0.2	27.8	5.0	9.2
View recommended activities	10	0	0	11.1	2.6	6.9
View nearby activities	10	0	0	11.3	2.1	4.5
View activity details	10	0	0	8.3	2.9	5.6
Share activity	10	0	0	14.5	2.3	7.4
Open in maps	10	0	0	20.0	1.6	8.3

Add to favorite	10	0	0	6.1	2.2	4.2
View favorites	10	0	0	5.10	2.2	3.8
Remove from favorite	10	0	0	5.5	1.2	3.6
Sign-out	10	0	0	10.3	1.1	4.0

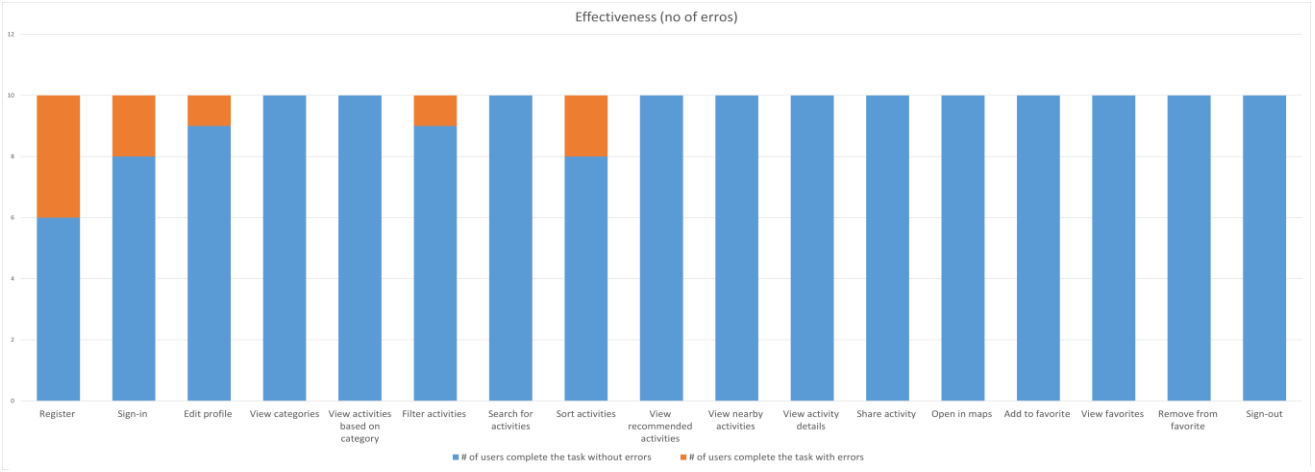


Figure 8: Summary of user's performance (no of errors) – effectiveness results.

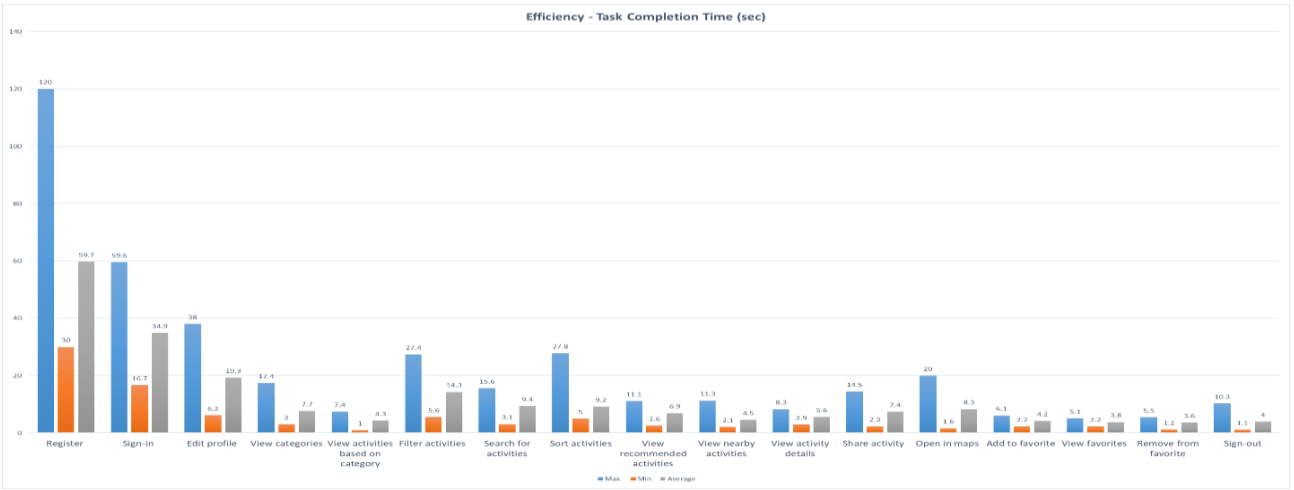


Figure 9: Summary of user's performance (completion time) - efficiency results.

Table 3: User satisfaction survey results

Questions \ Measure scales	# Of users who				
	Strongly agree. (5)	Agree (4)	Neutral (3)	Disagree (2)	Strongly disagree. (1)
I think that I would like to use LOCUS App frequently	6 (60%)	3 (30%)	1 (10%)	0	0
I found LOCUS App unnecessarily complex	0	0	0	6 (60%)	4 (40%)
I think LOCUS App was easy to use	6 (60%)	3 (30%)	1 (10%)	0	0
I think that I would need the support of a technical person to be able to use LOCUS App	0	0	0	3 (30%)	7 (70%)
I found the various functions in LOCUS App were well integrated	4 (40%)	4 (40%)	2 (20%)	0	0
I think there was too much inconsistency in LOCUS App	0	0	1 (10%)	4 (40%)	5
I would imagine that most people would learn to use LOCUS App very quickly	8 (80%)	1 (10%)	1 (10%)	0	0
I found LOCUS App very cumbersome to use	0	0	0	2 (20%)	8 (80%)
I felt very confident using LOCUS App	5 (50%)	4 (40%)	1 (10%)	0	0
I needed to learn a lot of things before I could get going with LOCUS App	0	0	1 (10%)	2 (20%)	7 (70%)
I found LOCUS App unnecessarily complex	0	0	0	6 (60%)	4 (40%)
I think LOCUS App was easy to use	6 (60%)	3 (30%)	1 (10%)	0	0
I think that I would need the support of a technical person to be able to use LOCUS App	0	0	0	3 (30%)	7 (70%)
I found the various functions in LOCUS App were well integrated	4 (40%)	4 (40%)	2 (20%)	0	0
I think there was too much inconsistency in LOCUS App	0	0	1 (10%)	4 (40%)	5
I would imagine that most people would learn to use LOCUS App very quickly	8 (80%)	1 (10%)	1 (10%)	0	0
I found LOCUS App very cumbersome to use	0	0	0	2 (20%)	8 (80%)
I felt very confident using LOCUS App	5 (50%)	4 (40%)	1 (10%)	0	0
I needed to learn a lot of things before I could get going with LOCUS App	0	0	1 (10%)	2 (20%)	7 (70%)

6 Conclusion and future work

The continuous growth in the tourism industry in addition to the high impact of mobile technologies on user experience in many fields and in particular the tourist and entertainment areas have led to high demand on mobile applications that provide a personalized experience and services to the tourist.

Derived by this demand, LOCUS was developed as a mobile application that targets tourists and aims to suggest for them activities, places and events that suit their interests and are based on the user's characteristics, behavior. It implements and integrates two types of recommender systems, the item-item collaborative filtering algorithm and the user-user collaborative filtering algorithm. It proposes places and activities to the user and those suggestions will constantly be improved

as the user keeps using the application.

Through user acceptance testing with 10 participants, we found that the system achieved promising results in terms of effectiveness, efficiency, and satisfaction. We plan to expand the application's capabilities by adding multilingual support and new features such as rating, writing reviews, and crowdsourcing to allow users to add places and activities. Overall, the development of LOCUS shows the potential for mobile applications to enhance the tourist experience, and we hope to continue exploring ways to improve and expand its functionality in the future. This could include, for example, exploring the potential benefits of incorporating user-generated content and social media features, as well as investigating ways to integrate the application more effectively with other tourism-related technologies and platforms.

References

[1] B. Cvetković *et al.*, "e-Turist: An Intelligent Personalised Trip Guide," *Informatica*, vol. 40, no. 4, Art. no. 4, Dec. 2016, Accessed: Apr. 20, 2023. [Online]. Available: https://www.informatica.si/index.php/informatica/article/view/1385

[2] A. H. Nabizadeh, N. Rafsanjani, N. Salim, A. Rezaei Aghdam, and K. Fard, "Recommendation Systems: a review," May 2013.

[3] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh, "Recommendation systems: Principles, methods and evaluation," *Egyptian Informatics Journal*, vol. 16, no. 3, pp. 261–273, Nov. 2015, doi: 10.1016/j.eij.2015.06.005.

[4] S. Missaoui, F. Kassem, M. Viviani, A. Agostini, R. Faiz, and G. Pasi, "LOOKER: a mobile, personalized recommender system in the tourism domain based on social media user-generated content," *Personal & Ubiquitous Computing*, vol. 23, no. 2, pp. 181–197, Apr. 2019, doi: 10.1007/s00779-018-01194-w.

[5] A. Tavčar, A. Csaba, and E. V. Butila, "Recommender System for Virtual Assistant Supported Museum Tours," *Informatica*, vol. 40, no. 3, Art. no. 3, Dec. 2016, Accessed: Apr. 29, 2023. [Online]. Available: https://www.informatica.si/index.php/informatica/article/view/1433

[6] H. N. L. Huy, H. H. Minh, T. N. Van, and H. N. Van, "Keyphrase Extraction Model: A New Design and Application on Tourism Information," *Informatica*, vol. 45, no. 4, Art. no. 4, Dec. 2021, Accessed: Apr. 29, 2023. [Online]. Available: https://www.informatica.si/index.php/informatica/article/view/3493

[7] R. Python, "Build a Recommendation Engine with Collaborative Filtering – Real Python." https://realpython.com/build-recommendation-engine-collaborative-filtering/ (accessed Apr. 20, 2023).

[8] S. A. El-Moemen, T. Hassan, and A. A. Sewisy, "A Context-Aware Recommender System for Personalized Places in Mobile Applications," *International Journal of Advanced Computer Science and Applications*, vol. 7, Mar. 2016, doi: 10.14569/IJACSA.2016.070360.

[9] "A location-aware recommender system for Tourism mobile commerce | IEEE Conference Publication | IEEE Xplore." https://ieeexplore.ieee.org/document/5690282 (accessed Apr. 30, 2023).

[10] A. Abu-Issa *et al.*, "A smart city mobile application for multitype, proactive, and context-aware recommender system," in *2017 International Conference on Engineering and Technology (ICET)*, Aug. 2017, pp. 1–5. doi: 10.1109/ICEngTechnol.2017.8308181.

[11] "Overview | Places API | Google Developers." https://developers.google.com/maps/documentation/places/web-service/overview (accessed Apr. 20, 2023).

[12] "ISO - IT tools for standards development," *ISO*. https://www.iso.org/it-tools-for-standards-development.html (accessed Jun. 20, 2021).

[13] "Android 11 | Android Developers." https://developer.android.com/about/versions/11?gclid=CjwKCAjwzruGBhBAEiwAUqMR8NPJvjC5qcihfp9pWlqwm0SvXyF043xzzhUrYorNA_X9KjXM-q1k0BoCWXMQAvD_BwE&gclsrc=aw.ds (accessed Jun. 20, 2021).

[14] "React Native · Learn once, write anywhere." https://reactnative.dev/ (accessed Apr. 20, 2023).

[15] "Firebase Realtime Database." https://firebase.google.com/docs/database?hl=en (accessed Jun. 20, 2021).

[16] "Welcome to Flask — Flask Documentation (1.1.x)." https://flask.palletsprojects.com/en/1.1.x/ (accessed Apr. 20, 2023).

[17] "Cloud Application Platform | Heroku." https://www.heroku.com/ (accessed Apr. 20, 2023).

[18] F. Cacheda, V. Carneiro, D. Fernández, and V. Formoso, "Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems," *ACM Trans. Web*, vol. 5, no. 1, p. 2:1-2:33, Feb. 2011, doi: 10.1145/1921591.1921593.

[19] "What is User Acceptance Testing (UAT Testing)? - Usersnap," *Usersnap blog - the feedback platform*, Feb. 17, 2020. https://usersnap.com/blog/user-acceptance-testing-right/ (accessed Feb. 07, 2021).

[20] "Usability Metrics - A Guide To Quantify The Usability Of Any System - Usability Geek." https://usabilitygeek.com/usability-metrics-a-guide-to-quantify-system-usability/ (accessed Feb. 07, 2021).

[21] "MeasuringU: 10 Benchmarks for User Experience Metrics." https://measuringu.com/ux-benchmarks/ (accessed Feb. 07, 2021).

[22] "MeasuringU: Measuring Usability with the System Usability Scale (SUS)." https://measuringu.com/sus/ (accessed Feb. 07, 2021).

An Integrated Approach for Analysing Sentiments on Social Media

Vrinda Tandon, Ritika Mehra*
Department of Computer Science and Engineering, Dev Bhoomi Institute of Technology, India
E-mail: vrinda1804tandon@gmail.com, riti.arora@gmail.com

Keywords: sentimental analysis, natural language processing (NLP), SVC, scikit-learn

Received*:* September 11, 2022

Sentiment analysis is an analytical subfield of Natural Language Processing (NLP) to determine opinion or emotion associated with the body of the text. The requirement for social media sentiment analysis has exceptionally increased with the growing extent of online activities in form of user generated content like posts or comments on social networking platforms. People often share their thoughts, opinions and reviews openly which can further be leveraged to analyze what they feel about a particular topic or their reviews/ feedback about a certain service. This study covers different approaches to conduct social media sentiment analysis on Twitter dataset both balanced and imbalanced obtained from Kaggle. For text analysis, we have implemented various classification techniques such as: Naive Bayes Classification and Support Vector Classification (SVC). It was concluded that SVC on twitter dataset surpassed other classification techniques in terms of performance.

Povzetek: Predstavljena je študija različnih pristopov za analizo sentimenta / razpoloženja na Twitterju, pri čemer se je najbolje izkazala metoda Support Vector Classification.

1 Introduction

Sentiment analysis also referred as emotion AI, helps to determine author's mentality or attitude by classifying their piece of writing as positive, negative or neutral. Sentiment analysis has displayed an intrinsic influence in various domains like product analysis, political campaigns, marketing and competitive research, brand inclination and monitoring. Content posted by the user on social media can be leveraged to make noteworthy interpretation of author's opinion, tone or emotions associated with their posts. The requirement for social media sentiment analysis has exceptionally increased with the growing extent of online activities in form of user generated content like posts or comments on social networking platforms. People often share their thoughts, opinions and reviews openly which can further be leveraged to analyse what they feel about a particular topic or their reviews/ feedback about a certain service.

In social media analysis often known as opinion mining, everything revolves around diving into words to understand the context of the user generated content and the opinions they reveal on such platforms. In this study we have summarized distinct classification techniques on twitter tweets dataset. The objective of this study is to classify sentiments of tweets with and without class weights for both balanced and imbalanced dataset hence, deriving the difference between the results obtained and its impact.

2 Literature review

The data sources on which sentiment analysis can be performed has grown exponentially with time hence; large amount of opinionated data can be fetched from different types of social media and websites. Content posted on different platforms like web-based entertainment platforms for example: movies reviews, social networking platforms or even product reviews can be utilized to make critical translation of author's tone, feelings or feedback related to the text they post. In this section, we have examined the work presented by different researchers on sentiment analysis and briefly discussed their approaches and observations after conducting the study.

Divij et al. in his work covered distinct pre-processing and classification techniques on binary and multi-class movie review dataset. Along with traditional classifiers, modern classifiers like RNN were also implemented. The performance for various approaches was compared where SVM with word embedding was proved to surpass other classification techniques [2]. Naresh et al. in his work determined social media user's opinions by using optimization-based machine learning algorithms. They found that the proposed technique sequential optimization with decision tree provides 89.47% of accuracy compared to other algorithms. Tweets were collected and classified into three categories i.e., positive, negative and neutral. According to the authors, for larger dataset this model will perform faster and will take less time [3]. Mullen et al. conducted a sentiment analysis study on movie review dataset. 1380 movie reviews were collected from a website

named epinions.com. The dataset consisted both negative and positive reviews. Support Vector Machine algorithm were applied to train and test the model. For validation purpose three-fold and ten-fold cross validation were used. 84.6% accuracy was obtained using three-fold cross validation whereas 86% accuracy was obtained using tenfold cross validation [1]. The study done to understand different classification techniques on various datasets are tabulated as Table 1.

Table 1: Study to understand different classification techniques on various data sets

Authors	**Paper Title**	**Models /Algorithms**	**Discussion**	**Datasets**	**Year**
Divij Gera, Amita Kapoor [2]	Sentiment Analysis using Scikit Learn: A Review	RNN and BERT models	To perform sentiment analysis on movie reviews	Binary classification dataset from IMDb and multi-class dataset from Rotten Tomatoes)	2022
Naresh, A. and Parimala Venkata Krishna [3]	An efficient approach for sentiment analysis using machine learning algorithm	Sequential minimal optimization with decision tree, Multivariate vehicle regression models	To classify the twitter data.	Airline twitter dataset	2021
Yuxuan Wang, Yutai Hou, Wanxiang Che & Ting Liu [4]	From static to dynamic word representations: a survey	Static and dynamic embedding models	Survey on evaluation metrics and applications of these word embeddings	TOEFL [13], ESL [11], RDWP[14], BM[12], AP and ESSLLI-2008[10]	2020
Kapoor, Amita [5]	Hands-On Artificial Intelligence for IoT: Expert machine learning and deep learning techniques for developing smarter IoT systems.	Machine Learning, Deep Learning and genetics Algorithms	Implement IOT to make their IOT solution Smart	UCI ML (Combined cycle poer plant)	2019
Vanaja, S., & Belwal, M. [6]	Aspect-level sentiment analysis on e-commerce data.	Naïve Bayes algorithm and Support Vector Machine (SVM) algorithm	Aspect-level Sentiment Analysis	Amazon Customer reviews data	2018

Jianqiang, Zhao, and Gui Xiaolin [8]	Comparison research on text pre-processing methods on twitter sentiment analysis	Naive Bayes and Random Forest, Logistic Regression, support vector machine	URLs do not contain useful information for sentiment classification	Stanford Twitter Sentiment Test (STS-Test), SemEval2014, Stanford Twitter Sentiment Gold (STS-Gold), Sentiment Strength Twitter (SS-Twitter) , Sentiment Evaluation (SE-Twitter)	2017
Gamallo, P., & Garcia, M. [9]	A Naive-Bayes Strategy for Sentiment Analysis on English Tweets.	Naive Bayes	For detecting the popularity of English tweets	SemEval2014 organization (tweeti-b.dist.tsv)	2014
Tony Mullen and Nigel Collier [1]	Sentiment analysis using support vector machines with diverse information sources	SVMs based on unigrams and lemmatized versions of the unigram models.	To assign semantic values to phrases and words within a text to be exploited in a more useful way	Epinions.com	2004

2.1 Concern with imbalanced data

One of the major challenges to deal with is imbalanced data. Imbalanced datasets are those datasets in which the observations distribution associated with the target class is not even. In other words, one class label possesses large number of observations as compared to the other class label. The main concern is to accurately and efficiently obtain the likelihood for minority as well as majority class. Imbalanced datasets are prone to give biased results hence to mitigate the issue distinct approaches are utilized.

Model is susceptible to fail when fed poor data, imbalanced data leads to inconsistent results and is considered as one of the major obstacles faced to obtain genuine results. In a study conducted by Alation [16] it was found that more than 80% of the participants were concerned about the quality of the data affecting the progress of their AI executions. https://www.alation.com/blog/alation-sodc-bad-data-spells-trouble-for-ai/

Imbalanced, mislabelled data and data gathered from unknown or non-reliable sources for training and testing tools is the major factor to produce flawed results. Some real-life failure examples induced by flawed data are:

An automated experimental hiring model by Amazon ended up as a failure due to imbalanced training data. The system designed for hiring was found to be biased against women candidates and trained itself by inferring male candidates better.

A racial inclination was found in health prediction algorithms used by US hospital and insurance organisations. The study published in science unveiled that the algorithm was found to recommend white patients over Black patients.

A predictive tool to identify covid-19 and diagnose patients was found to be not fit for clinical use by its own researchers. Derek Driggs' group observed that the trained dataset consisted scans of patients in lying and standing positions which inferred patients in lying position as seriously ill. The algorithm to identify covid-19 risk was inefficient as it was solely giving results based on the position of patients scanned.

In case of imbalanced dataset, the chances of algorithm being biased to the majority class are quite high and the main objective becomes to mitigate misclassification by minority class by setting a higher-class weight to minority class and simultaneously lowering the class weight to majority class. In this study, different weights were assigned to classes to improvise the performance for both binary and multiclass imbalanced data.

3 Working flowchart of proposed work

We implemented various classification techniques for text classification like: Multinomial Naïve Bayes, Bernoulli NB and SVC on the tweets posted by the user. The algorithm and working flowchart (Figure 1.) for the proposed work is as follows:

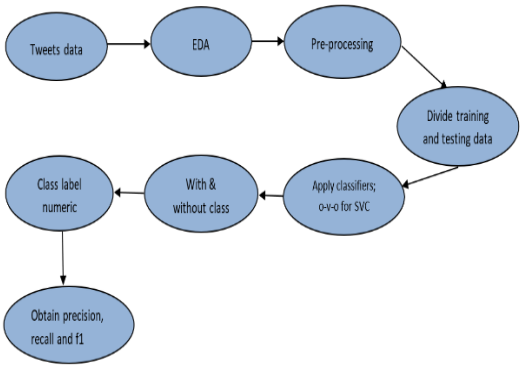


Figure 1: Proposed working flow chart.

3.1 Algorithm for proposed work

Step 1. Divide data into training and testing in 80 and 20 proportions.
Step 2. Determine nature of the dataset i.e., balanced or imbalanced using seaborn. countplot () for sentiment distribution.
Step 3. Apply Naive Bayes Multinomial, Bernoulli classifier and SVC classifier using o-v-o approach as per the nature of the dataset.
Step 4. Predict the test dataset using predict ().
Step 5. Classify tweets with numerical labels.
Step 6. Determine f1-score, precision and recall for each classifier.

4 Dataset utilised

Determining sentiment score is one of the prominent approaches to access emotion or tone of the text. This scaling system assigns scores corresponding to the tone of the text i.e., positive, negative or neutral making it easier to understand. We have utilized twitter Twitter tweets Sentiment Dataset obtained from kaggle to classify tweets sentiment. This multiclass dataset consisted of four columns i.e., textID, text, selected_text and sentiments associated with the tweets. The sentiments were labelled as positive, neutral and negative. In total there were 27.5k tweets present in the dataset.

4.1 Data pre-processing

The sentiments column was originally labelled as neutral, positive and negative so, in order to get sentiment distribution of the dataset we replaced neutral, positive and negative labels with numerical values 0, 1 and -1 respectively. The data containing 27.5k tweets was split into three numerical categories -1 to 1 from negative to positive sentiments associated with the tweets. And the sentiment distribution (Figure 2) of the dataset was represented using seaborn. countplot to draw the ordinal positions on the axis. This was done after importing the seaborn module in collab.

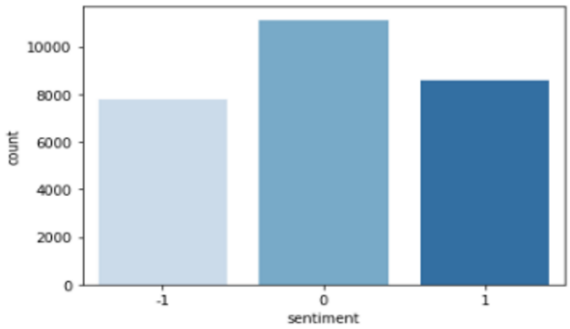


Figure 2: Sentiment distribution of the twitter data.

na values were present in the dataset were removed using fillna (). All the numerical values were also removed from the textual dataset followed by the punctuation removal for pre-processing the data. Conversion to lowercase and expanding contractions were also performed. The dataset was then divided into training and testing data (Figure 3) where plotting was done in order to ensure stratified data split.

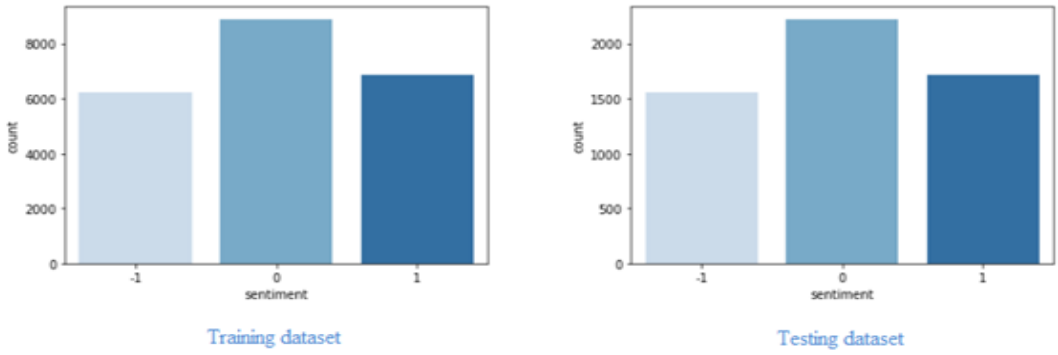


Figure 3: Training and testing dataset.

4.2 Imbalanced multiclass dataset

Imbalanced datasets are those datasets in which the observations distribution associated with the target class is not even. In other words, one class label possesses large number of observations as compared to the other class label. The main concern is to accurately and efficiently obtain the likelihood for minority as well as majority class. Imbalanced datasets are prone to give biased results hence to mitigate the issue distinct approaches are utilized. After obtaining the summary of the dataset using info() method similar operations were performed. The columns were labeled as -1, 1 and 0 for negative, neutral and positive labels after which seaborn_countplot was utilized to represent sentiment distribution as represented in the Figure 4. Numerical values as well as punctuations were removed; lowercase conversion was also done for pre-processing the data.

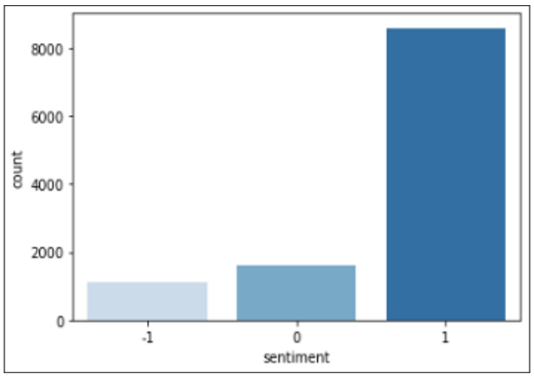
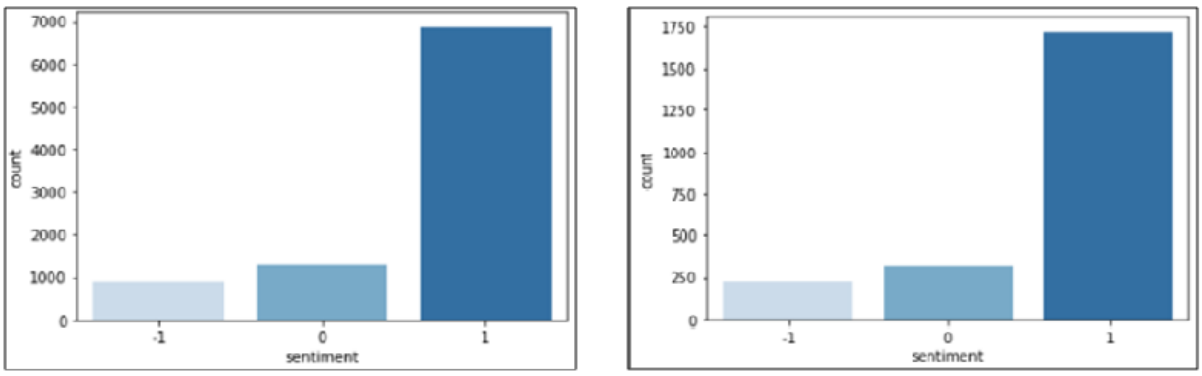


Figure 4: Sentiment distribution of the imbalanced twitter data.

This imbalanced dataset was divided into training and testing data (Figure 5) where plotting was done in order to ensure stratified data split.



Training dataset

Testing dataset

Figure 5: Training and testing dataset

5 Classification techniques implemented

5.1 Naive-Bayes classifier

Naive Bayes is a supervised learning algorithm based on Bayes Theorem. This probabilistic machine learning algorithm predicts on the basis of the probability of an object and is used to solve classification problems. Bayes Theorem also known as Bayes law is a mathematical formula for determining conditional probability as:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

The fundamental Naive Bayes assumption is that each feature each holds independent and equal contribution to the outcome. The three types of naive Bayes models are:

Gaussian Naive-Bayes classifier**:** This model assumes a normal distribution of features when working with continuous data and likelihood of the features is given as:

$$p(x = v \mid C_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}}\, e^{-\frac{(v-\mu_k)^2}{2\sigma_k^2}}$$

Multinomial Naive-Bayes classifier: As the name suggests, this classifier is utilized when we have multinomial distributed data. This is specifically used for document classification and conditional probability formula is given as:

$$p(\mathbf{x} \mid C_k) = \frac{(\sum_{i=1}^{n} x_i)!}{\prod_{i=1}^{n} x_i!} \prod_{i=1}^{n} {p_{ki}}^{x_i}$$

Bernoulli Naive-Bayes classifier: It is a multivariate event model also popular for document classification where binary term occurrence features are used instead of term frequencies. Here features are independent booleans describing inputs. The likelihood of features is given as:

$$p(\mathbf{x} \mid C_k) = \prod_{i=1}^{n} p_{ki}^{x_i}(1 - p_{ki})^{(1-x_i)}$$

5.2 Naive-Bayes classifier on the data

Data cleaning was done followed by stop-words removal. In order to display frequently occurring features in the corpus a WordCloud (Figure 6) was also produced. The dataset was then vectorized using TF-IDF vectorizer.

The Multinomial and Bernoulli Classifiers provided by Scikit-Learn library were trained and validated using the obtained vectors.

Figure 6: WordCloud for twitter dataset.

In case of imbalanced dataset, the chances of algorithm being biased to the majority class are quite high and the main objective becomes to mitigate misclassification by minority class by setting a higher-class weight to minority class and simultaneously lowering the class weight to majority class. Therefore, different weights were assigned to classes to improvise the performance.

5.2 Splitting multiclass classification into binary classification

The algorithms designed for binary classification cannot be leveraged in multi-class classification problems so, to mitigate this issue we use heuristic methods like one-vs-rest and one-vs-one methods to make binary classifiers work as multiclass classifiers.

One-Vs-Rest Classification Model: also known as one-vs-all is a heuristic method to enable binary classification algorithms work as multi-class classification algorithms. In this technique the multi-class data is split as binary classification data in order to apply binary classification algorithms so that it can be ultimately converted into binary classification data [15].

One-vs-One Classification model: this approach is similar to o-v-r as it also functions by splitting the data i.e. by splitting multi-class dataset into binary classification problem. The primary difference between o-v-r and o-v-o is that this classification model groups dataset into one single data file as an opponent to every other class [15].

In our dataset we used one-vs-one (ovo) classification strategy in order to split multi-class classification into binary classification problem per each sets of classes. We trained the underlying classifier as it is after which class weights were presented. Adjusted class weights were utilized to train the data.

6 Results

After completing Sentiment Analysis on twitter's dataset, testing and validation on various classifiers used in the study was done by determining the F1 score, precision and recall. It was evident in this study that the versions of Naive Bayes Classifier gave inaccurate results for imbalanced dataset which were then removed by introducing weights to the class. The results after conducting the study for balanced and imbalanced multiclass dataset are displayed in Table 2. Results for each algorithm for balanced multiclass dataset with and without weights were presented in which no major variation after introducing class weights was observed due to balanced nature of the dataset. SVC displayed better results.

Table 2: Observation Table F1-score for balanced dataset

S No.	Classifier	F1-score
1.	Bernoulli Naive-Bayes (without classweights)	[0.5473 0.6598 0.6851]
2.	Multinomial Naive-Bayes (without class weights)	[0.4914 0.6555 0.6280]
3.	Bernoulli Naive-Bayes (With class weights)	[0.4955 0.6371 0.5163]
4.	MultinomialNaive-Bayes (with class weights)	[0.4292 0.6356

		0.3952]
5.	Support Vector Classifier (without class weights)	[0.6683 0.6944 0.7489]
6.	Support Vector Classifier (with class weights)	[0.6803 0.6738 0.7547]

Inaccurate results for each algorithm were found in case of highly imbalanced dataset so, to mitigate the issue class weights were presented after which SVC and Bernoulli displayed improved results (Table 3).

Table 3: Observation Table F1-score for imbalanced dataset

S.No.	Classifier	F1-score
1	Bernoulli Naive-Bayes (without classweights	0.6608
2	Multinomial Naive-Bayes (without class weights)	0.6560
3	Bernoulli Naive-Bayes (with classweights)	0.6880
4	Multinomial Naive-Bayes (with class weights)	0.6714
5	Support Vector Classifier (without class weights)	0.7024
6	Support Vector Classifier (with class weights)	0.7446

7 Conclusion and future scope

In this study we covered different approaches to conduct social media sentiment analysis on Twitter tweets dataset fetched from Kaggle. The twitter dataset was multiclass data which required more data pre-processing and hence was closer to the dataset in real life situation where sentiment analysis is conducted. Binary classification algorithms were leveraged in multiclass dataset with the help of o-v-o heuristic technique. These tweets were classified in positive, negative and neutral categories using distinct classification approaches. We implemented various classification techniques using Scikit-learn library for comparative analysis like Bernoulli Naive-Bayes, Multinomial Naive- Bayes and SVC using TF-IDF vectorizer.

A model is susceptible to fail and generate inaccurate results when fed poor i.e. imbalanced data. Highly imbalanced data can create a huge impact on the model's performance and in real life situation it is not surprising to encounter unbalanced datasets. Therefore, it is very important to select the right evaluation matrix in such scenarios. In our study we have utilised F1 score as our evaluation matrix and class weights were also introduced to obtain improved results and enabled us to study Multinomial, Bernoulli and Support Vector classifier with both class weights and without class weights for balanced and imbalanced dataset. For balanced dataset, no major variation was observed after introducing class weights while for imbalanced dataset, the results improved significantly where the Support Vector Classifiers (SVC) ended up being the best performing classifier with class weights. Further, we are focused to test GloVe word embeddings along with different approaches to handle imbalanced data for sentiment analysis.

References

[1] Mullen, Tony, and Nigel Collier. Sentiment analysis using support vector machines with diverse information sources. International conference on empirical methods in natural language processing, 412-418. 2004. https://doi.org/10.3115/1219044.1219069

[2] Divij Gera, D., & Kapoor A. Sentiment Analysis using Scikit Learn: A Review, 2022. https://doi.org/10.13140/RG.2.2.26189.10720

[3] Naresh, A. and Parimala Venkata Krishna. An efficient approach for sentiment analysis using machine learning algorithm. Evolutionary Intelligence. 14, 725- 731, 2021. https://doi.org/10.1007/s12065-020-00429-1

[4] Wang, Y., Hou, Y., Che, W., & Liu, T. From static to dynamic word representations: a survey. International Journal of Machine Learning and Cybernetics, 11(7),1611-1630, 2020. https://doi.org/10.1007/s13042-020-01069-8

[5] Kapoor Amita, Hands-On Artificial Intelligence for IoT: Expert machine learning and deep learning techniques for developing smarter IoT systems. ISBN-978-1-78883-606-5, 2019, Packt Publishing Limited, Birmingham, UK.

[6] Vanaja, S., & Belwal, M. Aspect-level sentiment analysis on E-commerce IEEE International Conference on Inventive Research in Computing Applications (ICIRCA), 1275-1279. 2018. https://doi.org/10.1109/icirca.2018.8597286

[7] Kumar, Y., Sharma, H., & Pal, R. Popularity Measuring and Prediction Mining of IPL Team Using Machine Learning. IEEE 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) ICRITO, 1-5, 2021. https://doi.org/10.1109/icrito51393.2021.9596405

[8] Jianqiang, Zhao, and Gui Xiaolin. Comparison research on text pre-processing methods on twitter sentiment analysis. IEEE Access 5: 2870- 2879, 2017.
https://doi.org/10.1109/access.2017.2672677

[9] Gamallo, P., & Garcia, M. Citius: A Naive-Bayes Strategy forSentiment Analysis on English Tweets. Semeval@ coling, *pp.* 171-175, 2014.
https://doi.org/10.3115/v1/s14-2026

[10] Baroni M, Evert S, Lenci A. Bridging the gap between semantic theory and computational simulations. In: Proc. of the ESSLLI workshop on distributional lexical semantic, 2008.
https://archive.illc.uva.nl/ESSLLI2008/Materials/BaroniEvertLenci/BaroniEvertLenci.pdf

[11] Baroni M, Murphy B, Barbu E, Poesio M. Strudel: a corpus-based semantic model based on properties and types. Cogn Sci 34:222–254, 2010.
https://doi.org/10.1111/j.1551-6709.2009.01068.x

[12] Landauer TK, Dutnais ST. A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychol Rev 211–240, 1997.*
https://doi.org/10.1037/0033-295x.104.2.211

[13] Turney PD. Mining the web for synonyms: PMI-IR versus LSA on TOEFL. In: De Raedt L, Flach P (eds) Machine learning: ECML 2001., Springer, Berlin Heidelberg, 491–502, 2001.
https://doi.org/10.1007/3-540-44795-4_42

[14] Jarmasz M, Szpakowicz S. Roget's Thesaurus and semantic similarity. In: Proc. of RANLP, pp 21, 2003.
https://doi.org/10.1075/cilt.260.12jar

[15] Multiclass classification:
https://www.analyticsvidhya.com/blog/2021/05/multiclass-classification- using-svm/

[16] Alation study:
https://www.alation.com/blog/alation-sodc-bad-data-spells-trouble-for-ai/

A Two Phase Ultrasound Image De-speckling Framework by Nonlocal Means on Anisotropic Diffused Image Data

Niveditta Thakur, *Nafis Uddin Khan, Sunil Datt Sharma
Jaypee University of Information Technology, Solan, India
E-mail: niveditta.thakur08@gmail.com, nafisuk97@gmail.com, sunildatt.sharma@juit.ac.in
*Corresponding author

Keywords: anisotropic diffusion, non-local mean filter, image de-speckling, contrast enhancement

Received: September 6, 2022

Image de-speckling is one of the most challenging issues in multimedia imaging systems. All of the available speckle noise reduction filters are nearly noise reduction capable, but they fail to restore subtle features such as low grey level edges and fine details against a low contrast background. Non-local mean filtering and anisotropic diffusion are two popular and effective methods for image de-speckling while preserving detail. This paper presents a two-phase ultrasound image de-speckling framework by utilizing the capability of the non-local mean filtering method for de-speckling and edge preservation on anisotropic diffused images. The prior image smoothing along with edge preservation and contrast enhancement by anisotropic diffusion is carried out in the first phase, which is then followed by the non-local means method for de-speckling and edge sharpening in the next phase. The degree of speckle noise attenuation is measured on low-contrast standard and ultrasound images and compared to state-of-the-art and advanced anisotropic diffusion techniques and non-local means methods. The percentage improvement of PSNR over the existing methods is found to be in the range of 2.06% to 46.68%. The experimental results show that the proposed method is capable of reducing noise and preserving edges better than existing speckle reduction filters.

Povzetek: Predstavljen je dvofazni pristop za odstranjevanje šuma iz ultrazvočnih slik, ki se je v primerjavi z obstoječimi filtri za zmanjšanje šuma izkazal za boljšega v ohranjanju robov in zmanjševanju šuma.

1 Introduction

Today, in the world of computer vision and artificial intelligence, a massive number of images are now required to train a deep learning model. Database management in block-chain technologies is another emerging multimedia application that stores a large number of videos and images. However, the performance of all these technologies is mostly dependent on the quality and natural aspects of the images. These images are mainly degraded due to the defects that arise during the image sensing and acquisition processes. This ill-posed problem affects ultrasound imaging and MRI in medical science, SAR images in remote sensing, and a variety of other scientific and industrial images [1] [40] [41]. Most of the methods available in the literature aim to remove these degradations from images, but doing so without disturbing the image's essential features has always been difficult. The effect of noise in ultrasound images with poor contrast is a serious issue that has been addressed in the proposed work. The most difficult type of noise in real world medical imaging is speckle noise [2][3] which is due to the multiplicative behaviour of unwanted pixels introduced into the images under poor illumination or environmental conditions. The multiplicative nature of the speckle tends to vary the mean value of pixels in a local area of the image, which degrades its visual information.

A generalized model of speckle noise [2][3] is given as:

$$g_{i,j} = I_{i,j} n_{i,j} \quad (1)$$

where $g_{i,j}$ is the speckled noisy image, $I_{i,j}$ is the original image at the pixel position (i, j) that has been perturbed by speckle noise $n_{i,j}$. This is mathematically represented as a white Gaussian noise of zero mean and variance at position (i, j). Ultrasound imaging suffers from this multiplicative noise, which degrades its usefulness in medical diagnosis and modality. The major problem in such speckle reducing filters is inaccurate localization of edges under low contrast regions and thus the loss of image information during the filtering process, which is called "over-filtering" or "over-smoothing." Thus, it has always been observed that it is very difficult to de-speckle such a low contrast image without disturbing the essential features like edges and boundaries.

Traditionally, the problem of image de-speckling had been tackled with the help of order statistic filtering like the mean, median, or Gaussian filter [4][5][6] for the reduction of Gaussian noise. Later on, some stochastic and wavelet-based methods are also developed for

effective Gaussian noise reduction [7][8][45][46][47]. However, the blurring of edges, lines, and boundaries degrades the image's perception quality. Classical de-speckling filters [9][10][11][12] inhibit smoothing near edges by using adaptive filters that calculate coefficients of variation, but these filters do not possess directional properties, which creates the problem of poor reconstruction of edges and boundaries.

Recently, the use of partial differential equation based anisotropic diffusion has become a very convincing edge-preserving smoothing scheme since the earlier work of Perona and Malik [13][27][38]. The anisotropic diffusion is a non-linear and space-variant transformation of the original image that aims to smooth the noise adaptively. However, the performance of anisotropic diffusion in image de-noising is dependent on a more robust mathematical exploration of the diffusion equation, which deals with the image structures consisting of homogeneous as well as heterogeneous characteristics.

The Non-Local Means (NLM) method is another well-known and effective technique for de-speckling ultrasound images [28][36][37]. The NLM method reduces the impact of noise in an image by computing a weighted average of all the pixels in the non-local areas of the image. The mechanism of the NLM filter is to compute the weight based on the surrounding pixels of the test pixel in a local window along with the other windows of similar appearance in the same image. Consequently, the similar-looking patches or regions in an image would influence the weight assignment more for the test pixel as compared to the regions that are different from the test pixel. This filter, unlike local mean filtering, produces a better smoothed image and tends to recover the pixels representing edges. The major drawback of non-local mean filtering is the computational cost, which includes selecting the number and size of similar patches in the image.

The proposed work presented here is a hybrid framework for de-speckling ultrasound images along with sufficient recovery of edges and boundaries by utilizing the properties of anisotropic diffusion and non-local mean filtering. As it has been observed, the two techniques have been independently applied for noise removal and edge preservation, but achieving both targets equally by either of the two methods is not satisfactory. Therefore, the proposed method has been framed to adopt the properties of the two methods to achieve the goal of image enhancement along with sufficient speckle reduction.

The contributions of the proposed method are as follows:

- The proposed method is a two-phase process where anisotropic diffusion is used as a pre-filtering step to enhance the contrast by reformulating the diffusion coefficient function, taking speckle noise characteristics into consideration.
- The resulting diffused image is further passed through the NLM filter in the next phase in order to de-speckle the image with sufficient edge preservation.
- The use of the anisotropic diffusion process prior to applying the NLM filtering has been shown to be quite effective in synthetic and ultrasound images.

This two-phase hybrid algorithm maintains an optimal balance between speckle removal, fine details, and edge preservation, along with contrast improvement and less computational complexity.

The organization of the paper is followed by Section 2 which describes a brief understanding of ultrasound de-speckling with related anisotropic diffusion and NLM filtering methods. The proposed two-phase method has been presented in Section 3. The experimental analysis and discussion are made in Section 4. Finally, the paper is concluded in Section 5.

2 Related ultrasound de-speckling methods

Many efficient methods have been developed so far for ultrasound de-speckling. The two very popular methods, anisotropic diffusion and non-local mean filters, have been briefly explained in this section.

2.1 Anisotropic diffusion

The anisotropic diffusion [13][27][38] introduced by Perona and Malik, due to its directional smoothing properties is quite effective in the images affected by Gaussian noise and now it has become quite popular in medical imaging especially in the enhancement and de-speckling of ultrasound images. Anisotropic diffusion model as suggested by Perona and Malik [13] is given by:

$$\partial_t I = \text{div}\,(C(\nabla I).\nabla I) \tag{2}$$

$$C(\nabla I) = \frac{1}{1 + \left(\frac{|\nabla I|}{k}\right)^2} \tag{3}$$

where$t$ is iteration, I is input image, ∇I is image gradient, $C(\nabla I)$ is diffusion coefficient function and k is edge threshold parameter. The two equations (Eq. (2) and (3)) show that the diffusion process is controlled by $C(\nabla I)$ which is varied with respect to the gradient magnitude at each pixel of the image. The gradient magnitude specifies the direction of smoothing as per the image structures. The edge threshold parameter k and no. of iterations t are two important parameters which affects the performance of anisotropic diffusion. The values of k and t are chosen depending on the particular image

application. If $\nabla I > k$ at a pixel, then the smoothing stops and the pixels at higher gradient are preserved otherwise, if $\nabla I < k$, then anisotropic diffusion become isotropic and behave like Gaussian smoothing. This concept of smoothing was first used in ultrasound de-speckling by Yu and Acton [14] where non-homogeneous diffusive heat phenomenon was utilized. This method has been named as "speckle reducing anisotropic diffusion" (SRAD) where the diffusion coefficient function has been modified as follows:

$$c(\mathrm{q}) = \frac{1}{1 + [q^2(x,y,t) - {q_0}^2(t)]/[{q_0}^2(t)(1 + {q_0}^2(t))]} \quad (4)$$

or

$$c(q) = \exp\left\{-\frac{[q^2(x,y,t) - {q_0}^2(t)]}{[{q_0}^2(t)(1 + {q_0}^2(t))]}\right\} \quad (5)$$

where, $q(x,y,t)$ is named as instantaneous coefficient of variation which is dependent on ∇I and is determined as:

$$q(x,y,t) = \sqrt{\frac{\left(\frac{1}{2}\right)(\nabla I/I)^2 - \left(\frac{1}{4}\right)(\nabla^2 I/I)^2}{1 + \left(\frac{1}{4}\right)(\nabla^2 I/I)^2}} \quad (6)$$

and $q_0(t)$ is speckle scale function. The edge preservation sensitivity of this method was further examined [15] and presented as detail preserving anisotropic diffusion where the orientation of edges was made to stabilize while removing speckle noise as indicated in Eq (7).

$$q(x,t) = \frac{|\alpha\|\nabla I\|^2 - \beta(\nabla^2 I)^2|^{1/2}}{[I + \gamma\nabla^2 I]} \quad (7)$$

The $q(x,t)$ in Eq (7) denotes the edge stabilizing function with α, β and γ are the regularization parameters. However, the improper selection of these regularization parameters in the diffusion equation suffers from over-filtering and blurring of edges. Fernandez and Carlos [16] provide the estimation of these parameters using local statistics of the image features for better anisotropic diffusion de-speckling. A fuzzy optimization algorithm as suggested by Puvanathasan and Bizheva [17] has been used to properly select the parameters in order to compute the edginess of a pixel, which is effective for optical coherence tomography images specifically. Wu and Tang [18] suggested a new selective degenerate diffusion model using fidelity and speed functions based upon ENI (edge, noise, interior pixels) for impulse noise reduction. A similar type of improved edge-enhancing diffusion approach was developed by Febrinni et al. [19] to minimize noise in homogeneous regions while keeping weak edges. The statistical properties of speckle noise have been investigated for proper selection of parameters in the diffusion equation by Ramos et al. [20]. A double degenerated nonlinear diffusion model [21] was developed by reframing the diffusion coefficient function as shown in Eq (8).

$$c(\nabla I) = \frac{2|I|^{\alpha}}{M^{\alpha} + |I|^{\alpha}(1 + |\nabla I|^2)^{(1-\beta)/2}} \quad (8)$$

The above diffusion coefficient function is calibrated by taking $\alpha > 0, 0 < \beta < 1$ and M indicates the maximum intensity range of the image. This work was extended by introducing a gray level indicator [22] in the diffusion coefficient function as given below:

$$C(\nabla I) = \frac{\nabla I}{1 + \left(\frac{|\nabla I|}{k}\right)^{g(I)}} \quad (9)$$

where $g(I)$ is defined as $(2 - ({2I^{\alpha}}/{M^{\alpha} + |I|^{\alpha}}))$. This promotes the de-noising process with gray level extraction of the image. K-means clustering has also been used to filter out the noisy pixels from the image during the diffusion process [23]. In this method, the cluster-based speckle scale function and the homogeneous sample region are recursively chosen based on the previous clustering results. In order to control the diffusion process separately along the edges and across the edges, Mishra et al. [24] use the probability density function of edge and pixel relativity information, which is quite effective in ultrasound speckle filtering. Gao et al. [25] decompose the divergence term of the diffusion equation and modify the iteration stopping criteria to meet the requirements of speckle filtering along with edge preservation. This method diffuses erroneous pixels that appear in a uniform background. This method, however, is ineffective in low-contrast ultrasound images. Xu et al. [26] suggested Gabor-based anisotropic diffusion, supporting the advantages of the Gabor edge detector on edge preservation and the advantages of the Lattice Boltzmann method on rapid parallel implementation of the diffusion equation. Very recently, a fuzzy based approximated anisotropic diffusion [38] has been presented, which is quite effective for impulse noise removal from standard scientific and medical images. A detailed overview of various anisotropic diffusion techniques [48][49][50][51][52][53] used for image enhancement and de-noising is provided in [39]. In all the above anisotropic diffusion filters, the smoothing results in ultrasound images are effective but fail to restore fine details. In addition to this, the low contrast of the ultrasound image after diffusion produces false

artifacts that lead to wrong medical diagnoses. The reason for this problem is that the presence of speckled, noisy pixels in the image exhibits high gradient magnitudes, which get restored at the output image due to the slow diffusion process. Therefore, there is a need for a more robust physical and mathematical interpretation of the anisotropic diffusion equation to deal with speckled, noisy images.

2.2 Non-local mean filtering

The non-local mean (NLM) [28][32][33] filter utilizes the pattern redundancy in the image content by analyzing large pixel data and collecting information from the whole image, looking for similar features. NLM filters are successful in filtering the noise and preserving edges, but their performance degrades at higher noise levels. Basically, the NLM algorithm [28] estimates a new value for a pixel in a given noisy image. The computation estimates the new value for a test pixel based on the surrounding pixels of the test pixel in a local window as well as other windows of similar appearance. Mathematically, the new estimated value $f_{NLM}(i)$ at a pixel i can be computed as the weighted average of all the pixel values of the image I or some predefined region of the image space as given below:

$$f_{NLM}(i) = \sum_{j \in f} w(i,j) f(j) \tag{10}$$

where, $w(i,j)$ is the weight which corresponds to the similarity between the neighborhoods around the pixel i and the pixel j located at some similar patch of the image. The weight $w(i,j)$ is computed as:

$$w(i,j) = \frac{1}{z(i)} exp\left(-\frac{\|N(i) - N(j)\|_{2,\sigma}^2}{h^2}\right) \tag{11}$$

The above calculation of the weight is exponentially decreasing function of weighted Euclidean distance $\|N(i) - N(j)\|_{2,\sigma}^2$ where $N(i)$ and $N(j)$are the fixed size neighborhoods centered at pixel i and j respectively. σ is the standard deviation of Gaussian kernel. The weight obtained through above calculation indicates the similarity between the pixels i and j which must satisfy the following conditions:

$$0 \le w(i,j) \le 1; \sum_j w(i,j) = 1$$

The term $z(i)$ defined in Eq (11) is a normalizing constant which is stated as:

$$z(i) = \sum_j exp\left(-\frac{\|N(i) - N(j)\|_{2,\sigma}^2}{h^2}\right) \tag{12}$$

The weights calculated are a decaying exponential function of Euclidean distances. This decay in the exponential function is controlled by a parameter h in Eq (12), which is termed as smoothing parameter of the filtering operation. A small value of h leads to inefficient noise filtering, whereas a high value tends to smooth the entire image. Thus, the proper tuning is required to preserve the image details while filtering the noise. Coupe et al. [29] adapted this NLM algorithm for ultrasound de-speckling by using a Bayesian framework (OBNLM), where a block-wise approach for weight computation was used. This method computes the weighted average of patches instead of the weighted average of all the pixels in the image. This tends to reduce the computational complexity, but the preservation of low contrast edges is not satisfactory. A shape-adaptive patch in NLM was developed using Stein's unbiased risk estimate [30] to further decrease the computational burden.

Recently, the KS-NLM filter [31] has become very popular in NLM filtering because it is based on the evaluation of Kolmogorov-Smirnov distance between the image pixels instead of Euclidean distance. This tends to compute the filtered value from the pixels of the patches, which are very similar to the target pixel, and the results are completely free from the ghost effect that has been observed in Euclidean distance-based NLM approaches. A three-stage hybrid algorithm (HSR) for de-speckling ultrasound images was recently introduced by Singh et al. [32], where the NLM filter is used in conjunction with guided filtering and bilateral filtering. A maximum likelihood-based method [33] and a fuzzy logic-based computational model (FSR) [34] have also been developed to attain the same goal. However, maintaining the balance between noise removal and contrast improvement with low computational cost is still a challenging issue in NLM filtering.

Speckle noise's multiplicative nature has a significant impact on SAR image signals, which intercept information retrieval from SAR imaging systems. During the monitoring and investigation of target elements, the backscattered echo captured by the system produces interference because of sunlight and different weather conditions, resulting in speckles in the images. These speckles introduce a granular appearance in images, which intercepts the information retrieval from SAR systems. The use of NLM filtering in SAR image de-speckling is extremely effective [42][43][44]. A recent study of implementing the NLM filter for SAR images was presented by Penna et al. [42], where stochastic distances have been embedded in the NLM filter in place of Euclidean distance in the wavelet domain. The method works well in homogeneous portions of real SAR images, but fails to restore the fine details and yields a slower computation time. Various Deep Learning (DL) and Convolutional Neural Network (CNN) based methods have also been utilized in computer vision tasks

for achieving fast computation. Many of them are used to deal with SAR image de-speckling. Recently, a CNN-based self-supervised SAR image de-noising method was developed [43], with the restored images exhibiting poor contrast and loss of low grey level edges and fine details due to the unknown speckle distribution. Mullissa et al. [44] provide a DL based image de-speckling model that estimates the de-speckling behaviour in accordance with the speckle noise distribution. This method is effective in speckle noise removal, but the major issue is its non-adaptability to different regions of SAR images. In addition to this, the computational requirements for estimating and evaluating the unknown de-speckling parameters make this method not convincing.

3 Proposed method

The proposed two-phase hybrid algorithm applies the anisotropic diffusion process in the first phase for contrast improvement of the image. The NLM filtering is then applied in the next phase for de-speckling and edge preservation. The functional block diagram of the proposed filtering method is given in Fig. 1. The two phases of the proposed hybrid algorithm are explained below.

3.1 Phase I: Pre-processing by anisotropic diffusion

Most of the ultrasound images used in medical diagnosis are of low contrast, where all the gray level intensities are scattered towards the darker side of the histogram. This makes distinguishing speckle defects and low grey level edges in ultrasound images extremely difficult. In the first phase of the proposed approach, a modified anisotropic diffusion is used to improve the contrast of the ultrasound image, along with low gray level edge preservation. This is achieved by employing gray level variance to control the diffusion process. The diffusion coefficient function has been updated so that it will vary in accordance with the gradient magnitude as well as the gray level variance of image pixels.

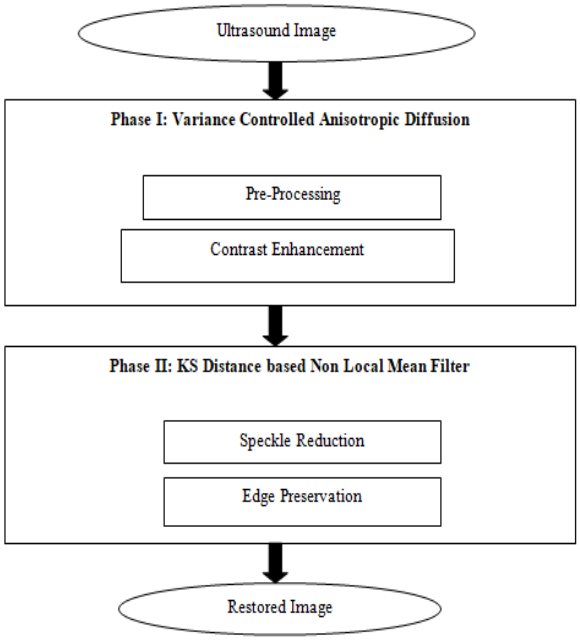


Figure 1: Block diagram of the proposed two-phase method.

Let r be the gray level of a pixel of an image whose gray levels are varying in the range l which is 0 to $(L-1)$ and $p(r_l)$ be an estimated histogram component corresponding to the gray levelr_l. The mean of the gray levels is then calculated as:

$$m_g = \sum_{l=0}^{L-1} r_l p(r_l) \tag{13}$$

The n_{th} moment of r about m_g can be written as:

$$\mu_n(r) = \sum_{l=0}^{L-1} (r_l - m_g)^n p(r_l) \tag{14}$$

In particular, the second moment of r as computed using Eq (14) can be considered as gray level variance of the image denoted by ${\sigma_g}^2$ which can be defined as:

$${\sigma_g}^2 = \mu_2(r) = \sum_{l=0}^{L-1} (r_l - m_g)^2 p(r_l) \tag{15}$$

The mean m_g and the variance ${\sigma_g}^2$ as calculated through Eq. (13) and Eq. (15) respectively are termed as global gray level mean and global gray level variance of the entire image. The m_g and ${\sigma_g}^2$ of a pixel at any point can also be calculated directly from the discrete image of size $M \times N$ as given in Eq (16) and Eq (17) respectively.

$$m_g = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(i,j) \tag{16}$$

$$\sigma_g{}^2 = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[I(i,j) - m_g]^2 \quad (17)$$

The assumption made here is that the pixel (i,j) in the image needs to be enhanced if gray level variance of that pixel in its neighborhood (sub-image) is comparatively lower than the gray level variance of the entire image. For this, we computed local gray level variances $\sigma_W{}^2$ at each pixel in their neighborhood W and measured their differences with the global gray level variance $\sigma_g{}^2$ of the entire image. The local gray level variance $\sigma_W{}^2$ of a pixel at location (i,j) at a neighborhood W of size $m \times n$ is calculated as:

$$\sigma_W{}^2 = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - m_W]^2 \quad (18)$$

where m_W is the mean of the pixels in W and the size $m \times n$ of W is very small as compared to that used in Eq.(17) for calculation of $\sigma_g{}^2$. The difference between the two variances $\sigma_g{}^2$ and $\sigma_W{}^2$ at a pixel of the image indicates about the contrast information of that particular pixel area. It has been assumed here that at a pixel location(i,j), if$\sigma_W{}^2 \ll \sigma_g{}^2$, then that particular pixel needs to get preserved in the output image. The overall contrast of the image also gets enhanced if all such pixels get preserved. This assumption is utilized in the proposed approach to enhance the low gray level inter-region edges with overall contrast enhancement.

The edge threshold parameter k in Eq (3) of Perona - Malik anisotropic diffusion has been reformulated in this work so as to smooth the ultrasound image in accordance with gradient as well as gray level variance of each pixel. The k in Perona - Malik anisotropic diffusion is replaced by edge threshold function $k(\sigma^2)$ in the proposed approach which is defined as follows:

$$k(\sigma^2) = \frac{1}{1 + \left(\frac{\sigma_W{}^2}{\sigma_0\sigma_g{}^2}\right)^q} \quad (19)$$

The above edge threshold function $k(\sigma^2)$ controls the diffusion process in accordance with the ratio $\frac{\sigma_W{}^2}{\sigma_0\sigma_g{}^2}$ where σ_0 is a scale factor with values $\sigma_0 \leq 1$ for low contrast areas whereas $\sigma_0 > 1$ for high contrast areas. If at any pixel position,$\frac{\sigma_W{}^2}{\sigma_0\sigma_g{}^2} > 1$, then $k(\sigma^2)$ goes down which tends to stop the diffusion and preserves the low gray level inter-region edges. In reverse manner if, $\frac{\sigma_W{}^2}{\sigma_0\sigma_g{}^2} < 1$, then $k(\sigma^2)$ rises to increase the diffusion. This variation of controlling function in anisotropic diffusion is summarized as follows:

$$k(\sigma^2) = \begin{cases} low & \sigma_W{}^2 > \sigma_0\sigma_g{}^2 \\ constant & \sigma_W{}^2 = \sigma_0\sigma_g{}^2 \\ high & \sigma_W{}^2 < \sigma_0\sigma_g{}^2 \end{cases} \quad (20)$$

The parameter $q(1 < q < 5)$ has been tuned here through experiments for best result. The revised diffusion coefficient function by substituting $k(\sigma^2)$ in place of k in Eq. (3) is now expressed as:

$$C(|\nabla I|, \sigma^2) = \frac{\frac{1}{\left(1+\left(\frac{\sigma_W{}^2}{\sigma_0\sigma_g{}^2}\right)^q\right)^2}}{\frac{1}{\left(1+\left(\frac{\sigma_W{}^2}{\sigma_0\sigma_g{}^2}\right)^q\right)^2} + |\nabla I|^2} \quad (21)$$

The above diffusion coefficient function varies in accordance with the combined effect of $|\nabla I|$and$k(\sigma^2)$. The following four cases have been observed:

1) *Case 1*: If at a pixel position,$|\nabla I| > k(\sigma^2)$ and $\sigma_W{}^2 > \sigma_0\sigma_g{}^2$, then the high value of $|\nabla I|$ and low value of $k(\sigma^2)$tends to stop the diffusion and preserve the low gray level inter-region edges.
2) *Case 2*: If at a pixel position, $|\nabla I| > k(\sigma^2)$ and $\sigma_W{}^2 < \sigma_0\sigma_g{}^2$, then both the$|\nabla I|$ and $k(\sigma^2)$ are high and diffusion undergoes as per the gradient magnitude and preserve the high gradient edges.
3) *Case 3*: If at a pixel position, $|\nabla I| < k(\sigma^2)$ and $\sigma_W{}^2 > \sigma_0\sigma_g{}^2$, then both the $|\nabla I|$and $k(\sigma^2)$are low and diffusion takes place with respect to variations in $k(\sigma^2)$. This tends to recover the low gradient and low gray level inert-region edges.
4) *Case 4*: If at a pixel position, $|\nabla I| < k(\sigma^2)$ and $\sigma_W{}^2 < \sigma_0\sigma_g{}^2$, then $|\nabla I|$ is low but $k(\sigma^2)$ becomes high which makes the diffusion faster to smooth out the entire homogeneous area.

The variation of diffusion coefficient function as per the above four cases controls the performance of smoothing and preserving low gradient as well as low gray level edges in low contrast ultrasound images.

3.2 Phase II: post-processing by non-local mean filter

In this phase, the image obtained from phase I undergoes NLM filtering. As discussed in Section 2(B), the performance of NLM filtering is mostly dependent on the estimation of weight function which in turn depends on the measurement of patch similarity. The image obtained from *Phase I* of the proposed method is contrast enhanced image along with edge preservation. The purpose of NLM filtering as a post processing phase is to

recover the original de-specked image with reduced ghost effects and well-located fine details. This has been attained by utilizing well popular "Kolmogorov – Smirnov" (KS) distance [31] in comparing the similar patches in search window. Unlike the conventional NLM filtering where the distance between similar patches have been measured, the proposed method measures the distances between the similar pixels. For this, the cumulative distribution functions (CDFs) are estimated for each pixel and the distances between their curves are calculated. The pixels which are very similar to the test pixel are merged to calculate the weight for the test pixel. The estimation of CDF of a pixel is probability based statistical measurement where we consider $p_f(r)$ asa probability of occurrence of a pixel of gray level r in a discrete image f and is defined as:

$$p_f(r) = p(f = r) = \frac{n_r}{n} \tag{22}$$

This is an estimated histogram component of $f$corresponding to gray level r in the range 0 to $L-1$. Heren_r and n are the number of occurrences of r_{th} gray level and total number of pixels in f respectively. The CDF corresponding to $p_f(r)$can be defined as:

$$CDF_f(r) = \sum_{s=0}^{r} \left(p_f(f = s)\right) \tag{23}$$

The above Eq. (23) is also termed as image's accumulated normalized histogram where s denotes another gray level with less probability of occurrence compared to r. The CDF of any random variable represents a monotonically increasing curve. The KS test provides the vertical difference between the CDF curves of two pixel's gray levels lying at different locations of f. In the proposed case, the KS test for distance measurement can be taken as:

$$D_{(r(i1,j1),s(i2,j2))} = MAX\left|CDF_{(i1,j1)}(r) - CDF_{(i2,j2)}(s)\right| \tag{24}$$

where, $r$and s are two gray levels at position $(i1, j1)$ and $(i2, j2)$ respectively. Based on the distance D measured in Eq (24) with the help of CDFs estimated at each point, the points closer to the test pixel are merged to calculate the corresponding weight. The weight calculation process is same as given in Eq. (11) and (12) except the weights computed are decaying exponential function of the distance measured based on CDFs as per the Eq (24).

4 Experimental results and analysis

The proposed two-phase method was run in *MATLAB R2015a* on a Windows 7 computer with *4 GB* of RAM and a *2.50 GHz Intel(R) Core (TM) i5-3210* processor. We tested the proposed method on varieties of ultrasound images taken from internet resources. However, for the purposes of this paper's demonstration, three real and one synthetic ultrasound images of $8-bit$ length and of size 256×256 were used. All the test images have been corrupted by additive noise of variance σ^2 varying from 0.2 to 0.8. Through experiments, the optimal parameters of the two phases of the proposed method were selected and, accordingly, the performance was evaluated. The criteria for parameter selection are explained in the following section.

4.1 Selection of parameters

In Phase I of the proposed method, explained in Section 3.1, it has been observed that the performance of the proposed method is dependent on two parameters σ_0 and q of the diffusion coefficient function defined in Eq. (21). The two parameters have been tuned to a particular value based on the image under test. As per the assumptions made in Eq, (21) explained in Section 3.1, $\sigma_0 \leq 1$ and $1 < q < 5$ are taken for initial experimentations. Initially, the value of σ_0 has been kept fixed and the different values of $q$have been varied in between 1 and 5. The results have been compared in terms of two well-known quality assessment parameters PSNRs and SSIMs at varying noise densities of $\sigma^2 =$ 0.2 to 0.8 respectively. In the similar fashion, the value of q has been kept fixed and different values of σ_0 varying between 0 and 1 are used and again the results are compared with respect to PSNRsand SSIMs at the noise densities$\sigma^2 = 0.2$ to 0.8 respectively. The experiments shows that the best result is obtained at $\sigma_0 =$ 0.8and$q = 2$. Therefore, we kept $\sigma_0 = 0.8$ and $q = 2$ in Eq. (21) of the proposed diffusion coefficient function for all the other test images used in this experimentation. The size of the neighborhood W for calculation of local gray level variance ${\sigma_W}^2$ as used in Eq. (18) and the no. of diffusion iterations are other important parameters which affects the result of the proposed method. We compared the three smallest standard local window sizes $3 \times 3, 5 \times 5$ and 7×7 in terms of PSNRs and SSIMs and the best results are obtained at $W = 3 \times 3$which has been taken in the entire experimentation. Similarly, the results are compared at iterations 30, 50,70 and 100 where the best performance was achieved at 50 iterations.

In *Phase II* of the proposed method explained in Section 3.2, the parameters in NLM filtering which need to be set for optimal performance are the patch size $N(i)$ for a pixel i and the smoothing parameter h during weight assignment. Through experimentations, the patch size $N(i)$ for a pixel i is decided by using all the standard sizes. However, the best results are obtained when the image is divided in to patches of size 3×3 and as per KS-NLM algorithm [31], the CDFs are estimated for each point of the 3×3window and their KS-distance is measured as given in Eq. (24). Similarly, during the weight assignment in Eq. (11) and (12), we have chosen the range of h to be between 4 to 10. At $h = 5$, the result

is quite acceptable but if we increase the value of h towards 10, then restored image become over-smoothed. Therefore, finally we decided to keep $h = 5$ in weight calculation of NLM filtering for the entire experimentations.

All the optimal parameters chosen in the proposed method are tabulated in Table 1.

4.2 Simulation results

The performance of the proposed two-phase hybrid method has been tested on varieties of real and synthetic ultrasound images. However, for the sake of demonstration in this paper, the two real ultrasound images of $8 - bit$ length and of size 256×256 have been used. The two original images are shown in Fig. 2 where the first image is Skull Osteoma ultrasound image and second one is Thyroid Lesion ultrasound image showing hypo-plastic left thyroid lobe and a benign cystic nodule in the right lobe. The two original images have been corrupted by additive noise of variance $\sigma^2 = 0.2, 0.4$ and 0.6 which are shown in the left

Table 1: Optimal Parameters

Phase I (Anisotropic Diffusion)		**Phase II (NLM Filtering)**	
Parameters	**Optimal Value**	**Parameter**	**Optimal Value**
Constants σ_0	0.8	Patch size $N(i)$	3×3
Constant q	2	Smoothing Parameter h	5
Local Neighborhood W	3×3		
Diffusion Iterations	50		

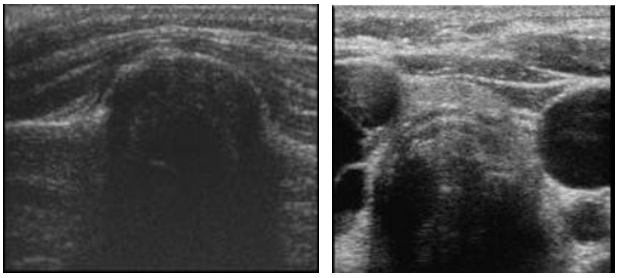


Figure 2: Skull osteoma ultra-sound image (left); thyroid lesion ultrasound image (right).

columns of Fig. 3 and Fig. 4 respectively. The middle column is showing the results of phase I where as the corresponding restored images obtained by phase II of the proposed two-phase hybrid method are demonstrated in the right columns of Fig. 3 and Fig. 4 respectively.

As seen in the restored images in Figs. 3 and 4, noise has been easily reduced while low gradient edges and fine details have been preserved. The behaviour of the proposed method in contrast enhancement can also be observed through the experimental results. The Skull Osteoma image (Fig. 2 (left)) is of quite low contrast, which then shows improvement in overall contrast after processing by the proposed method as shown in the right column of Fig. 3. The contrast of the Thyroid Lesion image (Fig. 2 (right)) is slightly better than that of the other test image, and the results obtained using the proposed method show significant contrast improvement, noise reduction, and detail preservation, as shown in Fig. 4.

In order to evaluate the performance of the proposed method more robustly, the two well-known image quality assessment parameters, PSNR and SSIM [35] of the restored images have been computed. Table 2 demonstrates the values of PSNRs and SSIMs of the two noisy test images as well as the corresponding two restored images shown in Fig. 3 and 4 respectively at noise densities of $0.2, 0.4$ and 0.6.

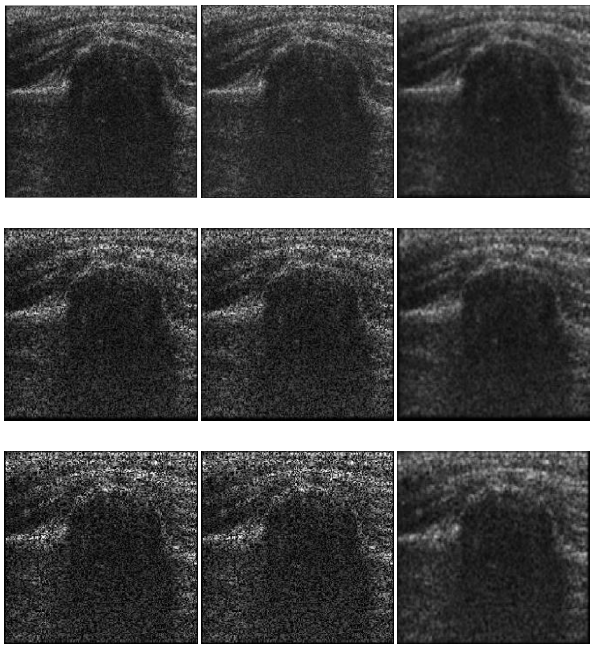


Figure 3: (left column: from top to bottom respectively) skull osteoma ultrasound noisy image with noise variance $\sigma^2 = 0.2, 0.4$ and 0.6; (middle column; from top to bottom respectively) intermediate results of phase i; (right column; from top to bottom respectively) restored skull images by proposed method (phase ii).

It can be observed from Table 2 that there is a considerable amount of increment in the values of PSNRs and SSIMs of the restored images as compared to the noisy images and the intermediate results obtained by phase I of the proposed method.

4.3 Performance comparison

The proposed method has been validated by comparing its performance with the existing state-of-the-art ultrasound de-speckling methods as well as some recently developed methods available in the literature. In the demonstration of comparison analysis, we used one synthetic and one real ultrasound image of $8-bit$ length and of size 256×256 each. The synthetic image includes oval, cardioid, line, triangle and a rectangle whereas the real ultrasound image is a kidney ultrasonic image generated by linear acoustic program [32][34][36]. The two original images are shown in Fig. 5 and both are contaminated with different levels of speckle noise ranging from variance $\sigma^2 = 0.2$ to 0.8. The restored images are compared perceptually as well as quantitatively with existing techniques: SRAD [14], OBNLM [29], FSR [34] and INLM [36]. Fig. 6 demonstrates the results obtained by the above four existing methods and the proposed method when applied to the synthetic image corrupted by speckle noise of variance $\sigma^2 = 0.6$.

Table 2: Quantitative results of the proposed method

Skull Osteoma Ultrasound Image						
Noise Variance (σ^2)	**Noisy Image**		**Phase I**		**Phase II-Restored Image**	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
0.2	20.60	0.4763	21.21	0.5234	22.26	0.7256
0.4	18.04	0.3899	20.54	0.4763	21.52	0.6846
0.6	16.77	0.3578	19.55	0.4554	20.56	0.5965
Thyroid Lesion Ultrasound Image						
Noise Variance (σ^2)	**Noisy Image**		**Phase I**		**Phase II-Restored Image**	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
0.2	16.11	0.4680	19.21	0.5667	20.35	0.6899
0.4	13.54	0.3639	17.63	0.4552	19.73	0.6394
0.6	12.56	0.3262	16.97	0.4112	19.40	0.6121

Table 3 and Table 4 respectively shows the average values of PSNRs and SSIMs of the restored synthetic and ultrasound images obtained by the all the five comparative methods. It can be seen from Fig 6 that the proposed method has shown the significant improvement in contrast of the input image in addition to the de-speckling as compared to other existing methods.

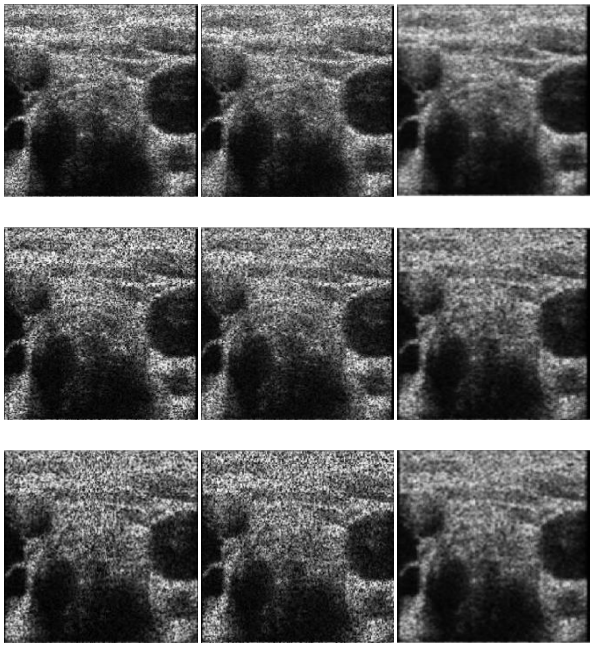


Figure 4: Thyroid Lesion ultrasound noisy image with noise variance $\sigma^2 = 0.2, 0.4$ and 0.6 (left column: from top to bottom respectively); (middle column; from top to bottom respectively) Intermediate results of Phase I; (right column; from top to bottom respectively) Restored images by proposed method (Phase II).

Moreover, Table 3 indicates that the PSNR values of the proposed method are much better than SRAD [14], OBNLM [29] and FSR [34]. However, if we compare the values with INLM [36], the PSNR obtained by the proposed method at higher noise variances are larger than that of the INLM [36] which shows that average improvement of PSNR at increasing noise densities is more in the proposed method as compared to INLM [36] and even other methods.

Figure 5: Synthetic ultrasound image (left); Real ultrasound image (right).

The similar observation can be seen in Table 4 where the SSIM values of all the comparative methods have been represented at different noise variances, which again validates that the proposed method has outperformed the existing methods at higher noise levels. Figure 7 depicts the restored results of the kidney ultrasonic image that had been corrupted by speckle noise of variance $\sigma^2 = 0.6$. This indicates, however, that while the PSNR and SSIM values obtained on the real ultrasound image are lower than those obtained on the synthetic ultrasound image, the values obtained by the proposed method are still better than those obtained by the existing methods. The average PSNR and SSIM values of the proposed method in comparison to the existing methods at varying noise densities on the two images are graphically represented in Figs. 8 and 9, which validate the effectiveness of the proposed method's de-speckling capability.

4.4 Computational complexity

The computational complexity of the proposed method depends upon two individual phases. In the *Phase I* of the proposed method, the computational complexity depends on the computation of gradient and local gray level variance of the image. If image is of size $N \times N$, then the complexity of calculating gradient in four main directions of the test pixel is $O(N^2)$ whereas the calculation of local gray level variance in the neighborhood size $w \times w$ yields the complexity of $O(w^2)$. Accordingly, the overall computational complexity of the first phase of the proposed method is $O(N^2 w^2)$. The NLM filter in *Phase II* of the proposed method has computational complexity of $O(N^2 M^2 L^2)$for image of size $N \times N$, search area $M \times M$ and local window size is $L \times L$. The proposed algorithm has been implemented in the system with Intel (R) Core (TM) i5 – 3210M processor of speed 2.50 GHz and 4 GB RAM. The overall average computation time of the proposed algorithm on an image of size 256×256 is found to be 36.13 seconds that is comparatively better than the existing methods.

The key feature of the proposed method is to de-speckle the low-contrast ultrasound image along with edge and fine detail preservation, which makes the proposed method different from the existing comparable methods, as can be observed from the above experimental analysis. There are, however, numerous opportunities to investigate the properties of anisotropic diffusion and non-local mean filtering in order to deal with high density speckle noise in edge-abundant images. as can be observed from the above experimental analysis.

Table 3: Comparison of average PSNR

Method	Noise Variance			
	0.2	**0.4**	**0.6**	**0.8**
Noisy Image	0.57	0.31	0.22	0.15
SRAD [14]	0.95	0.88	0.46	0.26
OBNLM [29]	0.94	0.92	0.85	0.80
FSR [34]	0.97	0.94	0.88	0.85
INLM[36]	0.96	0.92	0.88	0.75
Proposed Method	0.96	0.94	0.90	0.88

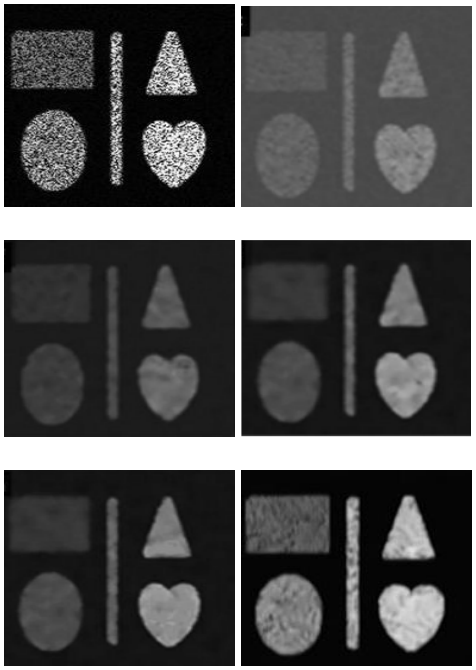


Figure 6: Noisy synthetic ultrasound image with $\sigma^2 = 0.6$, results of obnlm [29] and inlm [36] (left column: from top to bottom); results of srad[14], fsr [34] and proposed method (right column: from top tobottom)

Table 4: Comparison of SSIM

Method	Noise Variance			
	0.2	**0.4**	**0.6**	**0.8**
Noisy Image	18.54	16.02	14.96	14.13
SRAD [14]	28.10	21.90	13.10	08.90
OBNLM [29]	28.10	21.90	19.40	17.10
FSR [34]	29.60	24.10	20.80	18.20
INLM[36]	27.32	23.34	20.88	17.75
Proposed Method	29.30	26.65	23.85	22.80

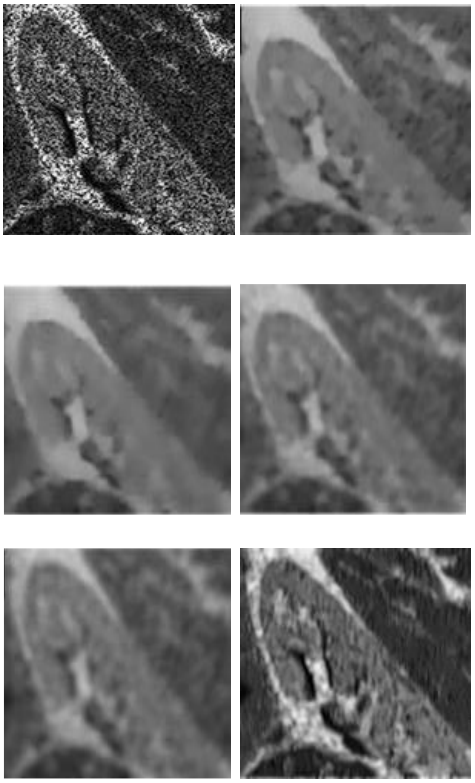


Figure 7: Noisy real ultrasound image with $\sigma^2 = 0.6$, results of obnlm [29] and inlm [36] (left column: from top to bottom); results of srad[14], fsr [34] and proposed method(right column: from top to bottom).

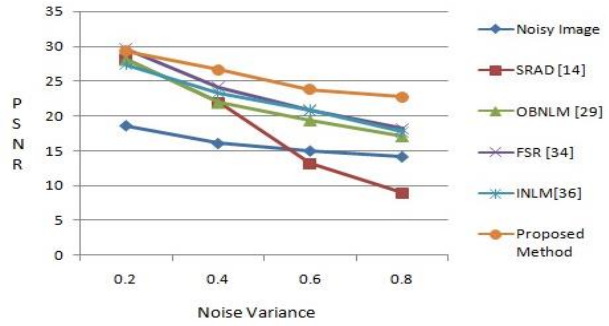


Figure 8: Graphical representation of PSNRs shown in table 3

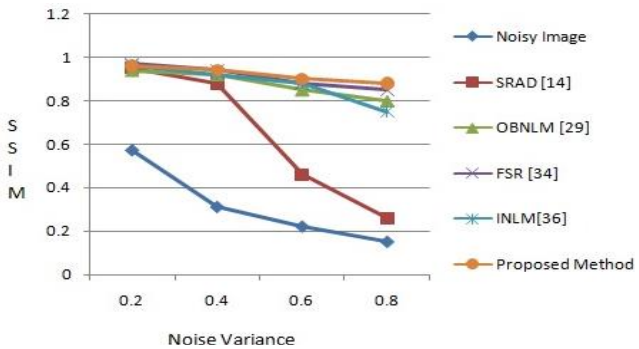


Figure 9: Graphical representation of SSIMs shown in table 4

4.5 Discussions

The performance analysis of the proposed two-phase method shown above demonstrates the effectiveness of ultrasound de-speckling in terms of better noise removal, acceptable edge preservation, easy implementation, and affordable computation requirements under the speckle noise variance level of 0.8. The proposed method does not require tuning of control parameters as used in existing anisotropic diffusion and NLM filtering. Because of this computational advantage, the proposed method is simple to implement and has a shorter execution time. However, the proposed method has a great deal of future scope. The main research issue in this work is to explore the properties of diffusion equations available widely in the literature to deal with discrete non-negative real image vector spaces for enhancing low gray level image pixels. The behaviour of the NLM filter can also be examined by the use of stochastic distances for a particular distribution in place of Euclidean distances.

The optimized mechanism of weight estimation in accordance with non-local self-similarity for a particular pixel requires parameter tuning through exhaustive experiments. This needs to be explored with the help of some optimization techniques for better performance. The real noisy images, like ultrasound and scenes from real SAR systems, need effective speckle reduction along with adequate information preservation, where this proposed method can be used as a better noise limiting tool. However, the selection of the sizes of the local patch and search area in the NLM filter and the properties of anisotropic diffusion need to be explored in future work for better compactness with a variety of images.

5 Conclusion

In this work, a hybrid framework for de-speckling low-contrast ultrasound images is presented. The anisotropic diffusion technique is used in conjunction with the KS distance-based non-local mean filtering mechanism. The proposed method is a two-phase method where prior contrast improvement of the ultrasound image is done by modified anisotropic diffusion, which is then followed by non-local mean filtering for de-speckling. The experiments are conducted on varieties of synthetic and real ultrasound images. The results show that the proposed method not only significantly reduces speckle noise but also preserves fine details and low grey level edges in the image, as well as an overall improvement in contrast. The above features of the proposed method make it useful in real-time medical applications to assist medical experts in the analysis and interpretation of ultrasonic images.

The proposed method can be further improved by investigating parameter optimization for better edge preservation and removal of high-level speckle noise in more complex ultrasound images. The exploration of the properties of anisotropic diffusion and researching the

mechanism of non-local mean filtering are possible research issues in this work.

References

[1] Goyal, B., Agrawal, S., &Sohi, B. S. (2018). Noise Issues Prevailing in Various Types of Medical Images. *Biomedical & Pharmacology Journal*, vol. 11, no. 3, pp. 12 – 27, 2018

[2] Tur, M., Chin, K. C., & Goodman, J. W. (1982). When is speckle noise multiplicative? *Applied optics*, *21*(7), 1157-1159.

[3] R. F. Wagner, S. W. smith, J. M. Sandrik and H. Lopez (1983), "Statistics of speckle in ultrasound B-scans", *IEEE Transactions on Sonics and Ultrasonics*, Vol 30(3), pp. 156 – 163

[4] Xin Wang (1992) Adaptive Multistage Median Filter. IEEE Transactions on Signal Processing, 1992, 40(4):1015-1017

[5] Kaur J., Kaur J., and Kaur M.: ' Survey of despeckling techniques for medical ultrasound images', *Int. Journal of Computer Technology and Applications*, 2011, **2**, pp. 1003– 1007

[6] HakanGüray Senel, Richard
Alan Peters, Benoit Dawant (2002), Topological median filter, *IEEE Transactions on Image Processing*, 11 (2)

[7] F. Jin, P. Fieguth, L. Winger, E. Jernigan (2003), Adaptive Wiener filtering of noisy images and image sequences, *IEEE Conference on Image Processing (ICIP 2003)*

[8] S. Sudha, G.R. Suresh, R. Sukanesh (2009), Speckle noise reduction in ultrasound images by wavelet thresholding based on weighted variance, *International Journal of Computer Theory and Engineering* 1(1):1793-8201

[9] Lee, J. S. (1980). Digital image enhancement and noise filtering by use of local statistics. *IEEE transactions on pattern analysis and machine intelligence*, (2), 165-168.

[10] Frost, V. S., Stiles, J. A., Shanmugan, K. S., &Holtzman, J. C. (1982). A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Transactions on pattern analysis and machine intelligence*, (2), 157-166.

[11] Kuan, D.T., Sawchuk, A. A., Strand, T. C., &Chavel, P. (1987). Adaptive restoration of images with speckle. IEEE Transactions on Acoustics, Speech, and Signal Processing, 35(3), 373-383.

[12] Lopes, A., Touzi, R., &Nezry, E. (1990). Adaptive speckle filters and scene heterogeneity. IEEE transactions on Geoscience and Remote Sensing, 28(6), 992-1000.

[13] Perona, P., & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on pattern analysis and machine intelligence*, *12*(7), 629-639.

[14] Yu, Y., & Acton, S. T. (2002). Speckle reducing anisotropic diffusion. *IEEE Transactions on image processing*, *11*(11), 1260-1270.

[15] Yongjian Yu and S. T. Acton, "Edge detection in ultrasound Imagery using the instantaneous coefficient of variation", *IEEE Transaction on image processing*, Vol 13, No. 12, pp. 1640 – 1655, 2004.

[16] Aja-Fernández, S., &Alberola-López, C. (2006). On the estimation of the coefficient of variation for anisotropic diffusion speckle filtering. *IEEE Transactions on Image Processing*, *15*(9), 2694-2701.

[17] Puvanathasan, P., &Bizheva, K. (2009). Interval type-II fuzzy anisotropic diffusion algorithm for speckle noise reduction in optical coherence tomography images. *Optics express*, *17*(2), 733-746.

[18] J. Wu and c. Tang, "PDA based random valued impulse noise removal based on new class of controlling functions", Vol. 20, No. 9, pp. 2428 – 2438, 2011.

[19] Fabbrini, L., Greco, M., Messina, M., &Pinelli, G. (2014). Improved edge enhancing diffusion filter for speckle-corrupted images. *IEEE Geoscience and Remote Sensing Letters*, *11*(1), 99-103.

[20] Ramos-Llordén, G., Vegas-Sánchez-Ferrero, G., Martin-Fernandez, M., Alberola-López, C., &Aja-Fernández, S. (2014). Anisotropic diffusion filter with memory based on speckle statistics for ultrasound images. *IEEE transactions on Image Processing*, *24*(1), 345-358.

[21] Zhou, Z., Guo, Z., Dong, G., Sun, J., Zhang, D., & Wu, B. (2015b). A doubly degenerate diffusion model based on the gray level indicator for multiplicative noise removal. *IEEE Transactions on Image Processing*, *24*(1), 249-260.

[22] Zhou, Z., Guo, Z., Dong, G., J., Zhang, D., & Wu, B. (2018c). A nonlinear diffusion equation based model for ultrasound speckle noise removal. *Journal of Nonlinear Science*, *28*(2), 443-470.

[23] Hu, Z., & Tang, J. (2016, September). Cluster driven anisotropic diffusion for speckle reduction in ultrasound images. In *2016 IEEE International Conference on Image Processing (ICIP)* (pp. 2325-2329). IEEE.

[24] Mishra, D., Chaudhury, S., Sarkar, M., Soin, A. S., & Sharma, V. (2018). Edge probability and pixel relativity-based speckle reducing anisotropic diffusion. *IEEE Transactions on Image Processing*, *27*(2), 649-664.

[25] Gao, M., Kang, B., Feng, X., Zhang, W., & Zhang, W. (2019). Anisotropic Diffusion Based Multiplicative Speckle Noise Removal. *Sensors*, *19*(14), 3164.

[26] Xu, H. H., Gong, Y. C., Xia, X. Y., Li, D., Yan, Z. Z., Shi, J., & Zhang, Q. (2019). Gabor-based anisotropic diffusion with lattice Boltzmann method for medical ultrasound de-speckling. *Mathematical biosciences and engineering: MBE*, *16*(6), 7546-7561.

[27] Krissian, K., Westin, C. F., Kikinis, R., &Vosburgh, K. G. (2007). Oriented speckle reducing anisotropic

diffusion. *IEEE Transactions on Image Processing*, *16*(5), 1412-1424.

[28] Buades, B. Coll and J.-M. Morel,"A review on image de-noising algorithms, with a new one", *Mathematical Modeling and Simulation*, Vol 4, No. 2, pp. 490 – 530, 2005

[29] Coupé, P., Hellier, P., Kervrann, C., &Barillot, C. (2009). Nonlocal means-based speckle filtering for ultrasound images. *IEEE transactions on image processing*, *18*(10), 2221-2229.

[30] C. A. Deledalle, V. Duval and J. Salmon,"Non-local Methods with Shape Adaptive Patches (NLM - SAP)", Journal of Mathematical Imaging and Vision, Vol. 43, pp. 103 – 120, 2011.

[31] Baselice, F. (2017). Ultrasound image de-speckling based on statistical similarity. *Ultrasound in Medicine & Biology*, 43(9), 2065-2078.

[32] Singh, K., Ranade, S. K., & Singh, C. (2017). A hybrid algorithm for speckle noise reduction of ultrasound images. *Computer methods and programs in biomedicine*, 148, 55-69.

[33] Shahdoosti, H. R., &Rahemi, Z. (2018). A maximum likelihood filter using non-local information for de-speckling of ultrasound images. *Machine Vision and Applications*, 29(4), 689-702.

[34] Nadeem, M., Hussain, A., &Munir, A. (2019). Fuzzy logic based computational model for speckle noise removal in ultrasound images. *Multimedia Tools and Applications*, 78(13), 18531-18548.

[35] Eskicioglu, A. M., & Fisher, P. S. (1995). Image quality measures and their performance. *IEEE Transactions on Communications*, *43*(12), 2959-2965.

[36] F. Mei, D. Zhang, Y. Yang. (2020). Improved non-local self-similarity measures for effective speckle noise reduction in ultrasound images. *Computer Methods and Programs in Biomedicine*, Vol. 196, 105670.

[37] H.Yan, P.Zhao, Z.Du, Y.Xu, P. Liu (2021). Frequency division de-noising algorithm based on VIF adaptive 2D-VMD ultrasound image. *Plos one*, Vol. 16, No. 3, e0248146, pp. 1-22.

[38] Niveditta Thakur, Nafis Uddin Khan and Sunildatt Sharma (2022). An efficient fuzzy inference system based approximated anisotropic diffusion for image de-noising. *Cluster Computing*, 25(43), 4303 – 4323.

[39] Niveditta Thakur, Nafis Uddin Khan and Sunildatt Sharma (2021). A Review on Performance Analysis of PDE Based Anisotropic Diffusion Approaches for Image Enhancement. *Informatica*, 45, 89 – 102.

[40] N. Singh and A. K. Bhandari (2022). Noise Aware L2-LP Decomposition-Based Enhancement in Extremely Low Light Conditions With Web Application. *IEEE Transactions on Consumer Electronics*, Vol. 68, No. 2, pp. 161 – 169.

[41] C. Yu and L. Z. Hou (2022). Realization of a Real-Time Image De-noising System for Dashboard Camera Applications. *IEEE Transactions on Consumer Electronics*, Vol. 68, No. 2, pp. 181 – 190.

[42] P. A. A. Penna and N. D. A. Mascarenhas (2019). SAR Speckle Nonlocal Filtering With Statistical Modeling of Haar Wavelet Coefficients and Stochastic Distances. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 57, No. 9, pp. 7194-7208.

[43] S. Tan, X. Zhang, H. Wang, L. Yu, Y. Du, J. Yin and B. Wu (2022). A CNN-Based Self-Supervised Synthetic Aperture Radar Image De-noising Approach. *IEEE Transactions on Geosciene and Remote Sensing*, Vol. 60, pp. 1-15, Art no. 5213615.

[44] A. G. Mullissa, D. Marcos, D. Tuia, M. Herold and J. Reiche (2022). deSpeckNet: Generalizing Deep Learning-Based SAR Image De-speckling. *IEEE Transactions on Geosciene and Remote Sensing*, Vol. 60, pp. 1-15, Art no. 5200315.

[45] S. S. Rajput, A. Singh, K. V. Arya and J. Jiang (2018). Noise robust face hallucination algorithm using local content prior based error shrunk nearest neighbors representation. *Signal Processing*, Vol. 147, pp. 233 – 246.

[46] S. S. Rajput and K.V. Arya (2019). Noise robust face hallucination via outlier regularized least square and neighbor representation. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, Vol. 1, Issue 4, pp. 252 – 263.

[47] S. S. Rajput, K.V. Arya and Vinay Singh (2018). Robust face super-resolution via iterative sparsity and locality-constrained representation. *Information Sciences*, Vol. 463, pp. 227 – 244.

[48] Nafis uddin Khan and K. V. Arya (2018). Two Stage Image De-noising and Edge Enhancement by SVD on Large Scale Heterogeneous Anisotropic Diffused Image Data. *Multimedia Tools and Applications*, Volume 77, pp. 22543 – 22566.

[49] Nafis uddin Khan, K. V. Arya and Manisha Pattanaik (2014). Edge Preservation of Impulse Noise Filtered Images by Improved Anisotropic Diffusion. *Multimedia Tools and Applications*, Volume 73, pp. 573-597.

[50] Nafis uddin Khan, K. V. Arya and Manisha Pattanaik (2013). Histogram Statistics based Variance Controlled Adaptive Threshold in Anisotropic Diffusion for Low Contrast Image Enhancement. *Signal Processing*, Volume 93, Issue 6, pp. 1684-1693.

[51] Nafis uddin Khan, K. V. Arya and Manisha Pattanaik (2012). A Robust PDE based Image De-noising Method. *International Journal of Computer and Electrical Engineering*, Volume 4, No. 3, pp. 303-306.

[52] Nafis uddin Khan, K. V. Arya and Manisha Pattanaik (2011). Image Enhancement and De-noising by Diffusion based Singular Value Decomposition. *International Journal of Computer Applications*, Volume 32, No. 8, pp. 14-22.

[53] Nafis uddin Khan, K. V. Arya and Manisha Pattanaik (2010). An Efficient Image Noise Removal and Enhancement Method. *Proceedings of 2010 IEEE International Conference on Systems, Man &*

Cybernetics (SMC 2010), October 10-13, 2010, Istanbul, Turkey, pp. 3735-3740.

Optimization of Personalized Recommendation Strategy for E-commerce Platform Based on Artificial Intelligence

Fan Yang
Henan Polytechnic Institute, Economic and trade Institute, China
E-mail: fanyang65@163.com

Keywords: artificial intelligence, e-commerce platform, personalized recommendation, collaborative filtering, optimization algorithm

Received: January 4, 2022

This article addresses the problems of "resource overload" and "information confusion" in the current e-commerce platform. This paper proposes a personalized recommendation strategy for e-commerce platform based on artificial intelligence considering the collaborative filtering method as the basic algorithm. The article proposes an optimized strategy using artificial intelligence to obtain satisfactory results. Based on this proposed personalized recommendation model, users are clustered by using ontology context information, further considering the influence of user preference and user trust relationship on similarity calculation. This method can alleviate the problems of data sparsity and cold start to a certain extent, effectively improve the recommendation quality. It further increases the diversity of recommendation results, and meet the needs of users and enterprises. Through the change of parameter α under different data sets, when α ∈ (1.84,1.88), the accuracy and recall rate of recommendation results remain at a high level. The personalized recommendation method can be applied to various situations such as social network friend recommendation and e-commerce platform commodity recommendation. The proposed work has a wide range of applications, especially for enterprises that master the user's rich dimensional situation information. This method has a prominent recommendation effect with the help of detailed analysis of the user's complex situation.

Povzetek: Članek obravnava težave preobremenitve z viri na obstoječih e-trgovinskih platformah. Predlaga se strategija personaliziranega priporočanja, ki temelji na umetni inteligenci in algoritmu kolaborativnega filtriranja.

1 Introduction

Artificial intelligence (AI) is a frontier science and interdisciplinary subject. At present, there is no unified theoretical system or even a unified definition in the world. Generally, it can be considered that artificial intelligence is a new technical science to study and develop theories, methods, technologies and application systems for simulating, extending and expanding human intelligence. As a branch of computer science, its main research contents include robot, language recognition, image recognition, natural language processing and expert system. Artificial intelligence is a simulation of human consciousness and thinking process. Although it is not human intelligence, it can simulate human thinking process and even surpass human wisdom in the future [1]. Artificial intelligence has gone through a long development process since its birth. We can sort out the context of the development of artificial intelligence from the two dimensions of time and technology. From the technical dimension, some scholars divide the development of artificial intelligence into three stages: computational intelligence, perceptual intelligence and cognitive intelligence. Computational intelligence is the primary stage of artificial intelligence. It mainly refers to the intelligence presented by the computer through its super large memory and supercomputing function. In this stage, logic is emphasized, but knowledge is ignored. Perceptual intelligence means that computers can simulate human perception with the support of big data, deep learning algorithms and other technologies, and then complete some tasks that originally need to be completed by humans, such as language recognition, image recognition, AR / VR and other technologies. Cognitive intelligence is not only the ultimate goal of the development of artificial intelligence, but also a key research field that people are committed to breaking through at present. It aims to enable computers to think and reason like people [2]. Since the concept of personalized recommendation was first proposed in the 1990s, it has quickly become a hot topic in academia, industry and other fields, and has maintained a high research heat. The core of personalized recommendation system is machine learning user interest. It is an advanced business intelligence platform based on massive data mining to help e-commerce websites provide fully personalized decision support and information services for their customers [3]. In essence, personalized recommendation system is to replace users to evaluate products they have never seen, automatically complete the process of personalized selection, help users explore interest, stimulate purchase desire and meet their personalized needs. At present, personalized recommendation system is widely used, especially in the

field of e-commerce. Businesses analyze their interests and hobbies according to the browsing, clicking, collection, purchase and other behaviors of e-commerce users, and recommend products that they may be interested in (such as daily necessities, books, audio-visual products, etc.). For businesses, the needs of users are usually unclear and vague. If the goods that meet the fuzzy needs of users can be recommended to users, the potential needs of users can be transformed into real needs, so as to improve product sales [4].

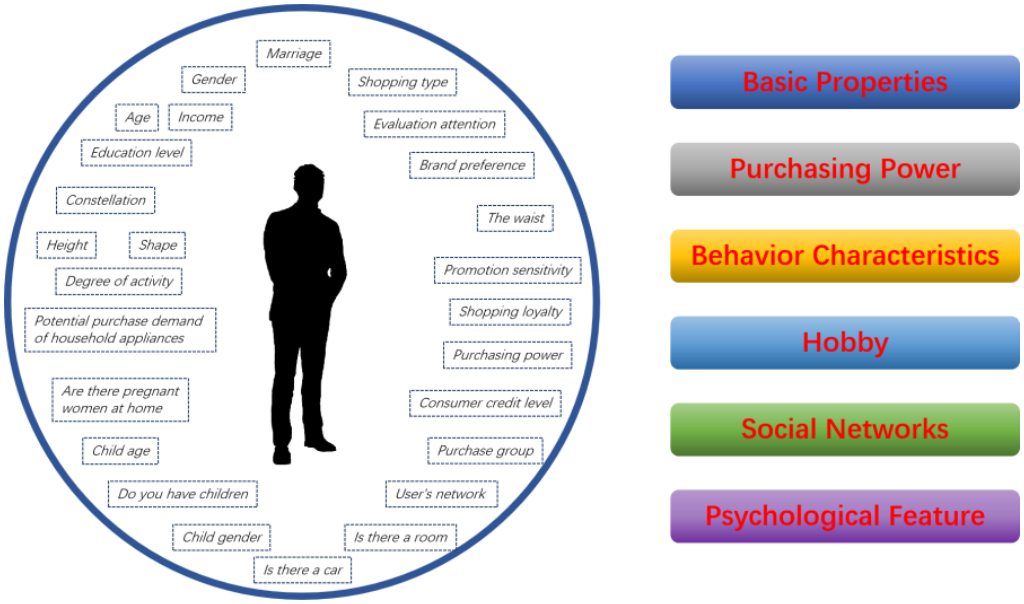


Figure 1: Concept diagram of personalized recommendation of e-commerce platform.

The core of personalized recommendation system is personalized recommendation method, which determines the quality of recommendation service. Figure 1 shows the concept of personalized recommendation of e-commerce platform.

This article contributes in addressing the problems of "resource overload" and "information confusion" in the current e-commerce platform. A personalized recommendation strategy for e-commerce platform based on artificial intelligence is proposed, considering the collaborative filtering method as the basic algorithm. The article proposes an optimized strategy utilizing the personalized recommendation model in which users are clustered by using ontology context information. This method can alleviate the problems of data sparsity and cold start to a certain extent, effectively improve the recommendation quality. It further increases the diversity of recommendation results, and meet the needs of users and enterprises. The proposed work has a wide range of applications, especially for enterprises that master the user's rich dimensional situation information.

The rest of this article is structured as: section 2 presents the literature review followed by the explanation of collaborative filtering method in section 3. Section 4 presents the results of the experimentations performed and the conclusion is provided in section 5.

2 Related work

In this section various state-of-the-art work in the field of personalized recommendation based on Artificial Intelligence are discussed.

Meng said that personalized recommendation methods originated from data mining technology. Scholars' research on personalized recommendation methods is more to improve and innovate them from the technical level [5]. In the era of e-commerce, enterprises provide a large number of goods to consumers through self-built e-commerce platforms or third-party e-commerce platforms. Dong and Zhou said that from the perspective of consumers, although rich goods increase consumers' selectivity. However, usually consumers cannot understand all the goods at a glance through the screen, nor can they directly check the quality of the goods. It takes a lot of time to fully understand these goods [6]. Therefore, Sanda *et al.* proposed that consumers need an e-shopping assistant, which can recommend products that consumers may be interested in according to their own interests [7]. Wei believe that personalized recommendation methods can help both e-commerce enterprises and traditional enterprises to quickly and accurately find the potential needs of users in an increasingly fierce competitive environment and meet the differentiated characteristics of users [8].

Aftalion and Bonnans said that the emergence of personalized recommendation can also solve the problems of "resource overload" and "information confusion" faced by users, and help individual users quickly and accurately obtain useful content from a large amount of information. In conclusion, the research on personalized recommendation is of great necessity [9]. Wei and Meng believe that with the popularization and deepening of e-commerce application in people's daily life, consumers put forward higher requirements on how

to more conveniently and quickly obtain the information resources they need, and the recommendation effect of traditional personalized recommendation methods in some scenarios is not satisfactory [10]. Lv said that at the same time, the concept of ubiquitous computing and the rise of ubiquitous commerce, a new business model, have made scholars fully realize that users' environmental information will have an important impact on their consumption decisions [11]. Therefore, Alhamid et al. said that personalized recommendation research based on social network context will be paid more and more attention [12]. Rana et al. said that in the e-commerce environment, especially in the mobile commerce environment, users' mobility, diversity of needs and dependence on context are particularly prominent [13]. Zhao said that the emergence of personalized recommendation service is to analyze and model consumers' historical behavior data with the help of certain methods or tools, evaluate consumers' preferences, and recommend products that target users may like [14]. Therefore, the explicit preference theory shows that the personalized recommendation method is reasonable and scientific. An approach through wavelet frames on the micropolar fluid flow is presented which is considered for high mass transfer [15]. The vibration on laminated skew sandwich plates is studied through finite element [16]. In another study numerical simulation based on space time fractional equation are evaluated [17]. The work can further be experimented with the integration of other Artificial Intelligence approaches and Machine learning as studied from several studies [18-20].

3 Collaborative filtering method

The memory-based method is also known as the user based collaborative filtering method. Its recommendation calculation depends on the user's evaluation of the commodity rather than the content information of the commodity itself. Its logical order is to first find other users with similar interests with the target user, and then use the scoring information of the commodity to make relevant prediction [21]. The main idea is to determine the nearest neighbor set by calculating the similarity between users, then predict the preference of the target user according to the score of the nearest neighbor on the commodity, and recommend the first several items with the highest prediction score to the user. For example, in an e-commerce website, the system predicts the possible interests and hobbies of the target user according to the scoring information of all over hit products and the similarity between users, and displays these products in the user's personal account [22]. It is assumed that there is a set of users $U = \{u_1, u_2, \ldots, u_m\}$, a collection of goods (objects) $O = \{o_1, o_2, \ldots, o_n\}$, $R_{i,j}$ indicates that the selected target user i scores the unselected product j, which needs to be obtained through the scoring prediction of the product by his similar users [23]. If U' represents the user set with the highest similarity with the target user i, the functional expression of the prediction score $R_{i,j}$ is as follows:

$$R_{i,j} = \frac{1}{m}\sum_{u \in U'} R_{u,j} \tag{1}$$

Where m is the number of users in U', and the prediction calculation is based on the simplest weighted average. In addition, there are other improvements based on weighted calculation, such as considering the influence of time factors to form a new expression:

$$R_{i,j} = \frac{1}{m}\sum_{u \in U'} \left(R_{u,j} \times Q_{u,j}\right) \tag{2}$$

Where Q represents a time series, and $Q_{u,j}$ represents the time point at which user U' selects or evaluates commodity j. Compared with formula 1, formula 2 incorporates the influence of time factors to improve the accuracy of recommendation [24].

In addition to the weighted prediction score, the most important thing is the similarity measurement between users. Similarity measure is to calculate the degree of similarity between individuals. The smaller the value of similarity measure, the smaller the similarity between individuals. The larger the value of similarity, the greater the individual difference. At present, there are many methods, among which the thre most commonly used methods are cosine similarity, Pearson correlation and jacquard correlation coefficient [25]. In fact, the value range of correlation is - 1 to 1. A positive value indicates that there is a positive correlation between the two vectors (variables), i.e., changes in the same direction, and a negative value indicates that there is a negative correlation between the two vectors (variables), i.e., changes in the opposite direction. In personalized recommendation, similarity generally only considers the positive relationship, and the corresponding similarity value range is 0 to 1. 0 means no similarity and 1 means complete similarity. Cosine similarity uses the cosine of the angle between two vectors in vector space as a measure of the difference between two individuals. The closer the cosine value is to 1, the closer the included angle is to 0 degrees, which means that the two vectors are more similar. Cosine similarity can be used in vector comparison in any dimension, especially in high-dimensional positive space. For example, in information retrieval, each term has different degrees. A document is represented by a weighted feature vector, and the calculation of the weight depends on the frequency of the term in the document [26]. Cosine similarity can give the similarity of two documents in terms of their topics. The expression formula of cosine similarity is:

$$Sim(i, f) = \frac{\sum_{j \in O(i) \cap O(f)} R_{ij} \cdot R_{fj}}{\sqrt{\sum_{j \in O(i) \cap O(f)} R_{ij}^2} \cdot \sqrt{\sum_{j \in O(i) \cap O(f)} R_{fj}^2}} \tag{3}$$

Where $Sim(i, f)$ represents the similarity between user i and user f (the value range of similarity is 0 to 1 and is a continuous value, 1 represents that the two users are completely similar, 0 represents that they are completely dissimilar). $O(i)$ and $O(f)$ represent the set

of goods selected by user i and user f respectively. R_{ij} and R_{fj} represent the scoring (evaluation) of product j by user i and user f respectively [27]. The advantage of cosine similarity is that it distinguishes differences in direction and is not sensitive to absolute values. Users' ratings of content are more used to distinguish similarities and differences of interest. It is precisely because cosine similarity is not sensitive to numerical values, that is, the difference of user rating scale is not considered, which also reduces the accuracy of user similarity calculation results. The emergence of Pearson correlation is to solve the problem of the difference of user scoring scale. The calculation results are corrected by subtracting the average score of the user for all items from each element in the user scoring vector [28].

Pearson correlation is a kind of linear correlation, which is a statistic used to reflect the linear correlation degree of two variables. In the standard formula, R is used to represent the correlation coefficient between vectors, n is used to represent the total number of samples, and X, Y, X' and Y' are used to represent the observed values and mean values of the elements in the two vectors respectively. R describes the degree of linear correlation between two vectors. The greater the absolute value of R, the stronger the correlation. The smaller the absolute value, the weaker the correlation. The expression formula of Pearson correlation is [29]:

$$Sim(i,f)=\frac{\sum_{j\in O(i)\cap O(f)}\left(R_{ij}-\overline{R_i}\right)\cdot\left(R_{fj}-\overline{R_f}\right)}{\sqrt{\sum_{j\in O(i)\cap O(f)}\left(R_{ij}-\overline{R_i}\right)^2}\cdot\sqrt{\sum_{j\in O(i)\cap O(f)}\left(R_{fj}-\overline{R_f}\right)^2}} \quad (4)$$

Where $Sim(i,f)$ represents the similarity between user I and user f, $O(i)$ and $O(f)$ represent the set of products selected by user i and user f respectively, R_{ij} and R_{fj} represent the scoring (evaluation) of product j by user i and user f respectively, and $\overline{R_i}$ and $\overline{R_f}$ represent the average score of all products by user i and user f respectively [30]. Jackard correlation coefficient was originally used to measure the similarity between two sets, which was defined as the ratio of the intersection of sets to the union of sets. The expression formula of jacquard correlation coefficient is:

$$Jaccard(A,B)=\frac{|A\cap B|}{|A\cup B|} \quad (5)$$

Where $Jaccard(A,B)$ represents the similarity between sets A and B. If sets A and B are empty, $Jaccard(A,B)=1$ is defined. Obviously, $0\le Jaccard(A,B)\le 1$, the larger the value, the more similar the two data sets [31]. Jackard correlation coefficient is actually more suitable to describe the similarity measure between vectors with discrete dimension characteristics. For the user similarity calculation in the collaborative filtering method, discrete data such as user score information are used, which is very suitable to use the $Jaccard$ correlation coefficient [32].

4 Results and analysis

This section presents the result analysis obtained for from the proposed model of personalized recommendation based on Artificial Intelligence.

Parameter α is involved in collaborative filtering recommendation model, and its optimal value needs to be determined through experiments. The optimal value range of α should be between 1.6 and 2. A positive value of parameter α indicates that unpopular goods contribute more to the similarity between the two users and popular goods contribute less. In order to quickly find the optimal value of parameter α, we still use the idea of binary search in the iterative process. Considering that the parameter value will take two decimal places, the spacing between each value in the iteration is set to 0.01, that is, the step size is 0.01 [33, 34]. The above strategy can effectively reduce the computational complexity and memory consumption. Figure 2 and 3 show the change of test index ranking score of the improved collaborative filtering recommendation method with the change of parameter α value under different data sets (recommendation list length L = 50). It can be seen from the figure that when α = 1.86, the ranking score under the two data sets reaches the minimum, that is, the products potentially liked by users are ranked in the high position. The results show that the unpopular goods contribute more to the similarity calculation between two users. This method can find the unique interests of users, improve the accuracy of recommendation results and increase the diversity of recommendation results. Of course, in practical application, users can also adjust the value of parameter α according to their own needs [35, 36].

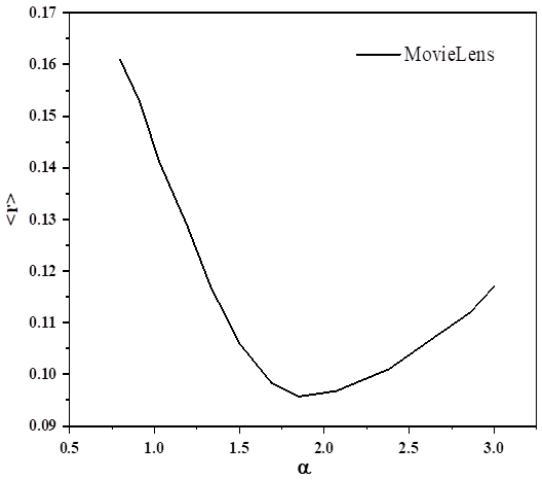


Figure 2: Performance influence curve of different values of parameter α in movie lens dataset on the recommended method.

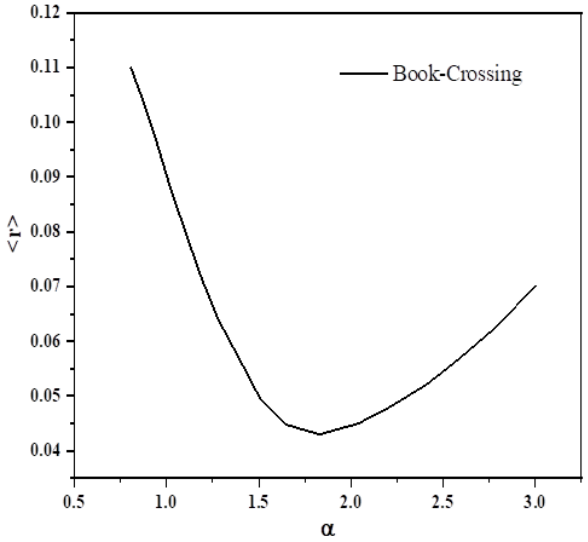


Figure 3: Performance influence curve of different values of parameter α in Book Crossing dataset on the recommended method.

When the data sets movie lens and Book Crossing and the length L of recommendation list are 50, when the parameter value is 1.86, the ranking score ($< R >$) reaches the minimum. The current experimental results come from the average of 5 random segmentation of data according to 80% test set and 20% training set. Accuracy and recall are often used to measure the effectiveness and efficiency of recommended methods.

When the length of the recommendation list is increased, the number of hits of potentially favorite products by users can be increased, thus improving the recall rate, but it may reduce the accuracy of recommendation (facts have proved that increasing the recall rate will basically reduce the accuracy rate). Therefore, it is necessary to find a balance between accuracy and recall, so as to have high accuracy and high recall at the same time. Under an appropriate recommendation list length, the accuracy reflects the precision of the recommended method, and the recall reflects the recall of the recommended method. Generally speaking, the length of the recommendation list shall not exceed 100, that is, the number of recommended products shall not exceed 100. Figures 4 and 5 show that under the movie lens data set, corresponding to the length of the recommendation list L = 50, the standard accuracy and recall rate of the test index change with the change of the parameter α value. It can be seen from the figure that when the parameter α = 1.86, the accuracy and recall of the recommended results reach the highest. In addition, when α ∈ (1.84,1.88), the accuracy and recall of the recommended results remain at a high level.

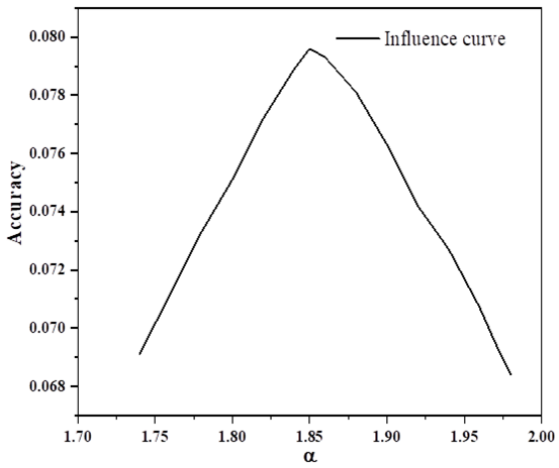


Figure 4: Influence of different values of parameter α on the accuracy of the recommended method.

Under the condition that the data set movie lens and the length of recommendation list L = 50, when the parameter value is 1.86, both indicators reach the highest value. The current experimental results come from the average of 5 random segmentation of data according to 80% test set and 20% training set.

Figures 6 and 7 show that under the book crossing data set, corresponding to the length of the recommendation list L = 50, the standard accuracy and recall of the test index change with the change of the parameter E value. It can also be seen from the figure that when the parameter α = 1.86, the accuracy and recall of the recommended results reach the highest at the same time. In addition, when α ∈ (1.84,1.88), the accuracy and recall of the recommended results remain at a high level.

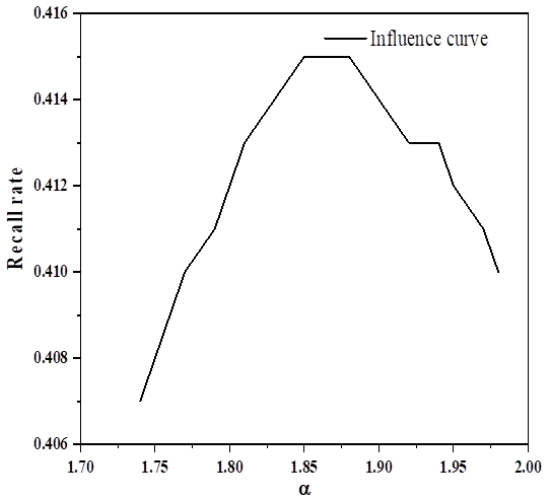


Figure 5: Effect of different values of parameter α on the recall rate of the recommended method.

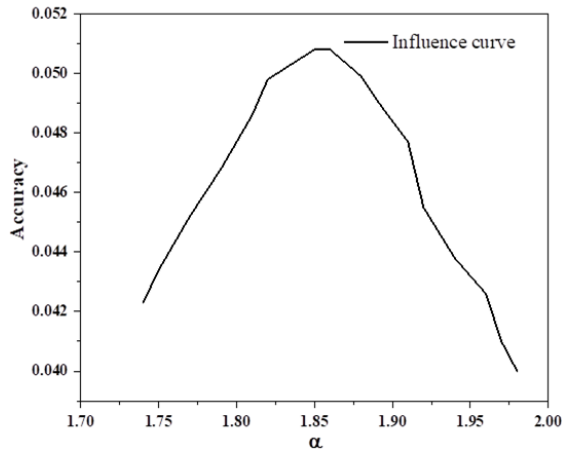


Figure 6: Influence of different values of parameter α on the accuracy of the recommended method.

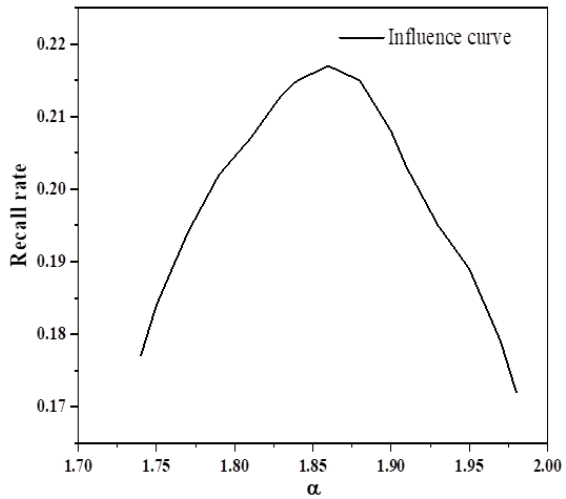


Figure 7: Effect of different values of parameter α on the recall rate of the recommended method.

When the data sets movie lens and Book Crossing and the length of recommendation list L are 50, when the parameter value is 1.86, both indicators reach the highest value. The current experimental results come from the average of 5 random segmentation of data according to 80% test set and 20% training set.

Through the test of the above three indicators (mainly measured by accuracy indicators), we iteratively obtain that the optimal value of the parameters in the model is 1.86. At this value, our recommended method achieves the optimal accuracy and the highest recall rate.

5 Conclusions

The development of e-commerce, especially the rapid development of mobile commerce, has brought greater challenges and opportunities to personalized recommendation services. The research on Personalized Recommendation theory and technology has attracted great attention of global academic circles. The research on Personalized Recommendation Based on context perception and personalized recommendation based on social relations in e-commerce environment has become a research hotspot. By comparing the parameters under different data sets α, the optimal value of the parameters in the model of 1.86 is achieved. Although some research results have been obtained, there are still some challenges in both theoretical research and practical application, especially for personalized recommendation in the field of e-commerce. There are few methods to integrate ontology situation, context situation and social relationship situation, lack of effective treatment of user interest drift, and lack of targeted recommendation strategy design for enterprises on demand. Therefore, it has become an important task to study a variety of personalized recommendation methods to meet the needs of different enterprises.

References

[1] Zhang, Z., Xu, G., Zhang, P., & Wang, Y. (2017). Personalized recommendation algorithm for social networks based on comprehensive trust. *Applied Intelligence*, *47*(3), 659-669. *https://doi.org/10.1007/s10489-017-0928-x*

[2] Fiorini, C. (2017). Optimization of running strategies according to the physiological parameters for a two-runners model. *Bulletin of mathematical biology*, *79*(1), 143-162. *https://doi.org/10.1007/s11538-016-0230-9*

[3] Li, H., Zhang, S., Shi, J., & Hu, Y. (2019). Research and design of intelligent learning system based on recommendation technology. *Mechatronic Systems and Control*, *47*(1), 43-49. *10.2316/J.2019.201-2968*

[4] Chen, C., Yang, J., Lu, M., Wang, T., Zheng, Z., Chen, Y., & Rudoff, A. (2021). Optimizing in-memory database engine for AI-powered on-line decision augmentation using persistent memory. *Proceedings of the VLDB Endowment*, *14*(5), 799-812. *https://doi.org/10.14778/3446095.3446102*

[5] Meng, G. (2018). A Brief Study of Optimization Strategies for Operation Site Allocation Plans for Distributed Queries. *International Journal of Computer Applications & Information Technology*, *11*(1), 224-228.

[6] Dong, W., & Zhou, M. (2014). Gaussian classifier-based evolutionary strategy for multimodal optimization. *IEEE Transactions on Neural Networks and Learning Systems*, *25*(6), 1200-1216. *10.1109/TNNLS.2014.2298402*

[7] Sanda, N. B., Sunusi, M., Hamisu, H. S., Wudil, B. S., Sule, H., & Abdullahi, A. M. (2018). Biological invasion of tomato leaf miner, Tuta absoluta (Meyrick) in Nigeria: problems and management strategies optimization: a review. *Asian Journal of Agricultural and Horticultural Research*, *1*(4), 1-14. *10.9734/AJAHR/2018/41959*

[8] Wei, X. L. L. (2018). Personalized Recommendation Strategy and Algorithm Optimization on Cloud Computing Platform. *International Journal of Performability Engineering*, *14*(10), 2492. *10.23940/ijpe.18.10.p25.24922503*

[9] Aftalion, A., & Bonnans, J. F. (2014). Optimization of running strategies based on anaerobic energy and variations of velocity. *SIAM Journal on Applied Mathematics*, *74*(5), 1615-1636. *https://doi.org/10.1137/130932697*

[10] Wei, J., & Meng, F. (2019). Personalized information recommendation based on synonymy tag optimization. *Cluster Computing*, *22*(3), 5467-5478. *https://doi.org/10.1007/s10586-017-1306-5*

[11] Lv, Y. (2017). Personalized recommendation model based on incremental learning with continuous discrete attribute optimization. Revista de la Facultad de Ingenieria, 32(2), 842-849.

[12] Alhamid, M. F., Rawashdeh, M., Dong, H., Hossain, M. A., & El Saddik, A. (2016). Exploring latent preferences for context-aware personalized recommendation systems. *IEEE Transactions on Human-Machine Systems*, *46*(4), 615-623. *10.1109/THMS.2015.2509965*

[13] Rana, I. A., Aslam, S., Sarfraz, M. S., & Shoaib, U. (2018). Analysis of Query Optimization Components in Distributed Database. *Indian Journal of Science and Technology*, *11*(18). *10.17485/ijst/2018/v11i18/122267*

[14] Zhao, Q. (2017). The application of personalized recommendation system in b2c e-commerce network platforms. Revista de la Facultad de Ingenieria, 32(14), 84-89.

[15] Kumbinarasaiah, S., & Raghunatha, K. R. (2021). A novel approach on micropolar fluid flow in a porous channel with high mass transfer via wavelet frames. *Nonlinear Engineering*, *10*(1), 39-45. *https://doi.org/10.1515/nleng-2021-0004*

[16] Dhotre, P. K., & Srinivasa, C. V. (2021). On free vibration of laminated skew sandwich plates: A finite element analysis. *Nonlinear Engineering*, *10*(1), 66-76. *https://doi.org/10.1515/nleng-2021-0006*

[17] Pedram, L., & Rostamy, D. (2021). Numerical simulations of stochastic conformable space–time fractional Kortewegde Vries and Benjamin–Bona–Mahony equations. *Nonlinear Engineering*, *10*(1), 77-90. *https://doi.org/10.1515/nleng-2021-0007*

[18] Wang, H., Hao, L., Sharma, A., & Kukkar, A. (2022). Automatic control of computer application data processing system based on artificial intelligence. *Journal of Intelligent Systems*, *31*(1), 177-192. *https://doi.org/10.1515/jisys-2022-0007*

[19] Sun, L., Gupta, R. K., & Sharma, A. (2022). Review and potential for artificial intelligence in healthcare. *International Journal of System Assurance Engineering and Management*, *13*(1), 54-62. *https://doi.org/10.1007/s13198-021-01221-9*

[20] Cai, Y., & Sharma, A. (2021). Swarm intelligence optimization: an exploration and application of machine learning technology. *Journal of Intelligent Systems*, *30*(1), 460-469. *https://doi.org/10.1515/jisys-2020-0084*

[21] Gungor, I., Emiroglu, B. G., Cinar, A. C., & Kiran, M. S. (2020). Integration search strategies in tree seed algorithm for high dimensional function optimization. *International journal of machine learning and cybernetics*, *11*(2), 249-267. *https://doi.org/10.1007/s13042-019-00970-1*

[22] Teshale, S. M., & Lachman, M. E. (2016). Managing daily happiness: The relationship between selection, optimization, and compensation strategies and well-being in adulthood. *Psychology and aging*, *31*(7), 687. *https://doi.org/10.1037/pag0000132*

[23] Biegler, L. T. (2017). Integrated optimization strategies for dynamic process operations. *Theoretical Foundations of Chemical Engineering*, *51*(6), 910-927. *https://doi.org/10.1134/S004057951706001X*

[24] Alam, S., Tawseef, M., Khan, F., Fattah, A. A., & Kabir, M. R. (2016). Differential evolution with alternating strategies: a novel algorithm for numeric function optimization. *Communications on Applied Electronics (CAE)–ISSN: 2394-4714 Foundation of Computer Science FCS, New York, USA Volume 4*, (2), 12-16. *10.5120/cae2016652030*

[25] Tadele, S., & Emana, G. (2018). Determination of the economic threshold level of tomato leaf miner, Tuta absoluta Meyrick (Lepidoptera: Gelechiidae) on tomato plant under glasshouse conditions. *Journal of Horticulture and Forestry*, *10*(2), 9-16. *10.5897/JHF2018.0522*

[26] Wang, S., Ali, S., Yue, T., & Liaaen, M. (2017). Integrating weight assignment strategies with NSGA-II for supporting user preference multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, *22*(3), 378-393. *10.1109/TEVC.2017.2778560*

[27] Zheng, Q., Xu, X., Martin, G. J., & Kentish, S. E. (2018). Critical review of strategies for CO2 delivery to large-scale microalgae cultures. *Chinese journal of chemical engineering*, *26*(11), 2219-2228. *https://doi.org/10.1016/j.cjche.2018.07.013*

[28] Li, S. (2020). Structure optimization of e-commerce platform based on artificial intelligence and blockchain technology. *Wireless Communications and Mobile Computing*, *2020*. *https://doi.org/10.1155/2020/8825825*

[29] Feng, Z. (2020). Constructing rural e-commerce logistics model based on ant colony algorithm and artificial intelligence method. *Soft Computing*, *24*(11), 7937-7946. *https://doi.org/10.1007/s00500-019-04046-8*

[30] Zeng, A., Yu, H., Da, Q., Zhan, Y., & Miao, C. (2020, April). Accelerating Ranking in E-Commerce Search Engines through Contextual Factor Selection. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No.

08, pp. 13212-13219). *https://doi.org/10.1609/aaai.v34i08.7026*

[31] Zhu, Y. (2021). Network public opinion prediction and control based on edge computing and artificial intelligence new paradigm. *Wireless Communications and Mobile Computing*, *2021*. *https://doi.org/10.1155/2021/5566647*

[32] Geng, T., Lin, X., & Nair, H. S. (2020, April). Online evaluation of audiences for targeted advertising via bandit experiments. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 08, pp. 13273-13279). *https://doi.org/10.1609/aaai.v34i08.7036*

[33] Renrui, L. (2022). Discussion on the Application of Artificial Intelligence in e-Commerce. *Journal of Electronics and Information Science*, *7*(1), 55-59. *10.23977/jeis.2022.070110*

[34] Sharma, R., Raju, C. S., Animasaun, I. L., Santhosh, H. B., & Mishra, M. K. (2021). Insight into the significance of Joule dissipation, thermal jump and partial slip: dynamics of unsteady ethelene glycol conveying graphene nanoparticles through porous medium. *Nonlinear Engineering*, *10*(1), 16-27. *https://doi.org/10.1515/nleng-2021-0002*

[35] Dhanalakshmi, A., Hui, X., Roopini, R., & Supriya, R. (2020). Technological Advancements in E-Commerce and Customer Relationship Management. *International Journal of Engineering and Management Research (IJEMR)*, *10*(6), 9-20. *10.31033/ijemr.10.6.2*

[36] Chen, Y., Zhang, W., Dong, L., Cengiz, K., & Sharma, A. (2021). Study on vibration and noise influence for optimization of garden mower. *Nonlinear Engineering*, *10*(1), 428-435. *https://doi.org/10.1515/nleng-2021-0034*

Logistics Distribution Route Optimization Based on Improved Particle Swarm Optimization

Hai Zhao*[1], Ashutosh Sharma[2]
[1]Anhui Industry & Trade Vocational Technical College, Huainan City, Anhui Province, 232007, China
[2]Southern Federal University, Russia
Emails: Haizhao8@126.com, sharmaashutosh1326@gmail.com
*Corresponding author

Keywords: particle swarm optimization, mountain climbing algorithm, logistics distribution, route optimization, hybrid PSO scheme

Received: February 15, 2022

This article improves the logistics distribution route and improves the distribution as well as transportation efficiency. The article combines the features of logistics dissemination along with mathematical designing of dissemination automobile routing issue. The mountain climbing procedure with strong local search ability is introduced into the particle swarm optimization (PSO) procedure to improve the offered approach. Two mountain climbing schemes are offered in this article, and two different hybrid (PSO) procedures are constructed. The experimental outcomes reveals the performance of Hybrid PSO scheme 1 and hybrid PSO scheme 2 offered in this paper which are better than that of standard PSO. Hybrid PSO scheme 2 offers best potential in efficiently solving the routing issue of logistics dissemination automobile. After the issue scale grows, the optimization advantages of Hybrid PSO scheme 2 are fully displayed. It was observed from the experimental analysis that using hybrid PSO scheme 2 to solve the logistics dissemination automobile routing issue can greatly shorten the dissemination mileage.

Povzetek: Članek izboljšuje logistične distribucijske poti s hibridnim pristopom rojev delcev (PSO), kar znatno skrajša razdaljo distribucije in izboljša učinkovitost prevoza.

1 Introduction

Dissemination is a task derived from the transportation link in the logistics system. It is an important link in the logistics system. In the logistics movement, delivery is actually the transportation of goods. Therefore, transportation is often used to represent delivery. Transportation costs account for the highest proportion of all logistics costs. Generally, the social logistics costs are calculated through comprehensive analysis, in which the goods accounts for about, and the goods of some products is even higher than the production cost [1]. The survey displays that the transportation cost of automobiles in China is times that of Europe and the United States. The empty driving rate of transportation automobiles in China is about 37%, of which the empty driving rate of automobiles in automobile logistics enterprises is as high as, and there are some issues such as return empty driving, waste of resources and high transportation cost. It is not difficult to see that the potential for saving transportation costs is very large [2, 3]. Using scientific methods to examine a reasonable dissemination route is an important work in dissemination activities. Reasonable selection of dissemination routes is of great significance to enterprises and society. For enterprises, optimizing the dissemination route can improve the dissemination efficiency, make the best use of the dissemination automobiles, reduce the dissemination cost as much as possible, deliver the goods to clients on time and quickly, greatly improve customer satisfaction, and help enterprises improve efficiency [4, 5]. For the society, it can save transportation automobiles, alleviate traffic tension, reduce transportation pollution such as noise and exhaust emissions, and contribute to the protection of ecological balance and the creation of a better home.

PSO is a growing computing technology based on the intelligence of swarm. Similar to GA (genetic algorithm), it is a population-based optimization tool. The system initializes a set of random results and searches for the optimal value through iteration [6]. However, there is no crossover and mutation of GA, but elements search for the optimal elements in the result space. In the system, each alternative result is called an "element". Several elements coexist and collaborate in optimization to approximate the bird swarm to find food. Each element "flies" to a better situation in the issue space according to its own "knowledge" and the best "knowledge" of neighboring element swarm to search for the optimal result [7, 8]. Figure 1 displays the optimization diagram of a reasonable logistics and dissemination route.

This article basically aims at the goal of providing decision support for logistics dissemination enterprises. The paper analyzes the logistics dissemination and automobile routing issue in detail, and establishes the mathematical design of logistics dissemination automobile routing issue. The main contribution of this work is to answer the lack of local search ability of (PSO) procedure, mountain climbing procedure with strong local search ability. The article further utilizes

(PSO) procedure to improve the offered approach. Two mountain climbing schemes are offered in this work, and two different hybrid (PSO) procedures are constructed in order to improve the logistics dissemination route and improve the dissemination efficiency.

The rest of this article is arranged as: section 2 describes literature review and the research method expressing the mathematical model of logistics dissemination automobile routing and model of hybrid (PSO) procedure is provided in section 3. Section 4 presents the experimantation and discussion of the experimental observations and conclusion is provided in section 5.

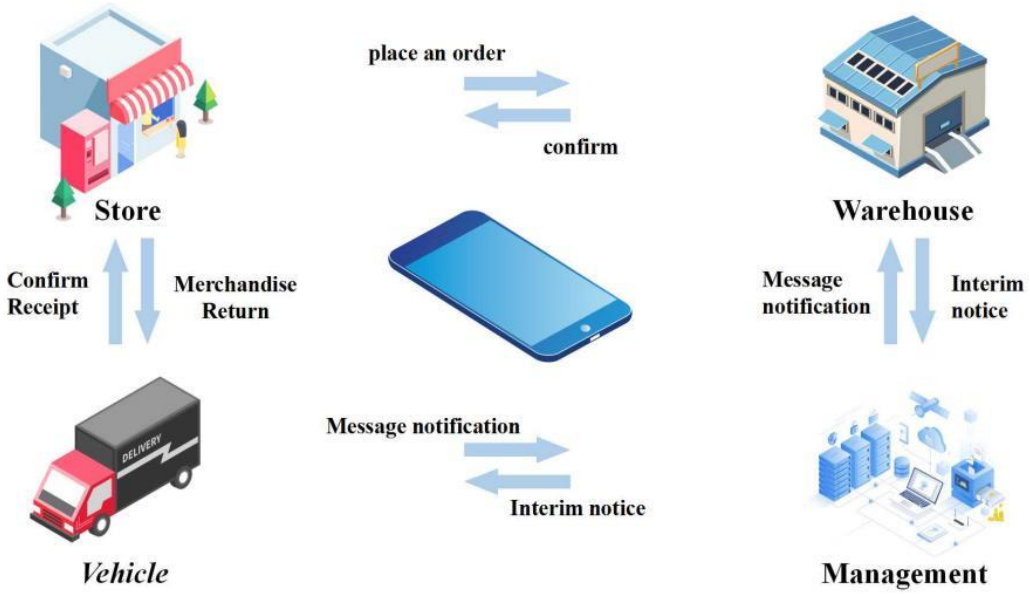


Figure 1: Optimization of the logistics dissemination route based on the improved element assembly procedure.

2 Related work

In terms of logistics dissemination route optimization, Muhammad, Muhammad *et al.* [9] used the improved GA to study the material dissemination issue, and the established dissemination mathematical design is also relatively simple. Wang *et al.* [10] used the improved saving method to study the optimization of material dissemination route, and did not consider the influence of cargo timeliness when establishing the objective task. Yuan *et al.* [11] used ant colony procedure to study the dissemination issue, focusing on the improvement of ant colony procedure and less research on the dissemination design.

As for the research of (PSO) procedure on route optimization, Moosavian and Lence [12] used (PSO) procedure to study the logistics dissemination issue, focusing on the improvement of the procedure, but the established dissemination mathematical design is relatively simple. Lagos *et al.* [13] applied (PSO) procedure to automobile routing issue and automobile routing issue with time window respectively. The research displays that (PSO) procedure can solve the issue quickly and effectively. Wan *et al.* [14] used the improved (PSO) procedure to solve the automobile routing issue. The comparison with GA and double population GA displays that the improved (PSO) procedure is an effective method to solve the automobile routing issue. Li *et al.* [15] applied the local (PSO) procedure to the non-fully loaded automobile routing issue. The research displays that the procedure improves the achievement rate of searching the optimal route and can solve the non-fully loaded automobile routing issue more effectively.

It can be seen from the relevant literature that the current research focuses on two aspects: the mathematical design and result method of logistics dissemination route optimization. The research on the mathematical design of dissemination route mainly focuses on how to more comprehensively and reasonably reflect the dissemination issue, and the commonly used result method is intelligent optimization procedure. However, most intelligent optimization procedures are easy to fall into local optimization and low search efficiency. Therefore, the improved design of intelligent optimization procedure and the reasonable establishment of mathematical design are the key to examine the advantages and disadvantages of route optimization methods.

3 Research methodology

This section includes the description of research methodology consisting of mathematical design of logistics dissemination automobile routing and design of hybrid (PSO) procedure.

3.1 Establish the mathematical design of logistics dissemination automobile routing

The automobile routing issue of logistics dissemination can be described as: several dissemination automobiles are used to deliver goods from a dissemination center to several clients. The location and cargo request of each customer are certain, and the load volume of each dissemination automobile is certain [16, 17]. It is required to reasonably arrange the automobile dissemination route, improve the objective task, and addresses the following assumptions:

1. The sum of the requests of each customer on each dissemination route shall not cross the load volume of the dissemination automobile.
2. An automobile can only select single route, but it can serve several clients.
3. A client has and can only have one automobile (car) to serve him.
4. Car starts from the dissemination center, sends the loaded goods to the corresponding clients along a driving route, and then returns to its own dissemination center.

Before setting the driving route, evaluate the quantity of automobiles used. It is based on the analysis that the more complex the loading (unloading) of goods and the more restrictions, the smaller the actual cargo volume of the automobile [18, 19]. The formula mentioned in Equation 1 is implemented to examine the quantity of automobiles K required:

$$K = \left[\frac{\sum_{i=1}^{L} g_i}{\alpha q}\right] + 1 \tag{1}$$

Where $[\sum_{i=1}^{L} g_i]$ means that the integer not greater than the quantity in parentheses, $\alpha \in (0,1)$ can be corrected according to the quantity of restrictions. Generally, the more restrictions, the smaller, and vice versa. The α general value is 0.85.

C_{ij} represents the transportation cost from customer i to customer j, such as time, distance, cost, etc. The quantity of the dissemination center is 0, the quantity of each customer is I (i = 1, ..., L), the quantity of each automobile is K (k = 1, ..., K), the cargo request of the i^{th} client is g_i, and the volume of the dissemination automobile is q.

Define the resulting values:

$$x_{ijk} = \begin{cases} 1 & Car\ k\ moves\ from\ i\ to\ j \\ 0 & otherwise \end{cases} \tag{2}$$

$$y_{ik} = \begin{cases} 1 & Car\ k\ serves\ clients\ i \\ 0 & otherwise \end{cases} \tag{3}$$

The goal of establishing the design is to minimize the total transportation cost. The transportation cost is directly proportional to the driving route of the automobile [20]. The smaller the operating route, the less the fuel ingestion of the automobile, the less the working time of the operator, and of course, the minimum the total conveyance cost. The following is a calculated design with the shortest driving route as the objective task for the logistics dissemination automobile routing issue:

$$min\ z = \sum_{i=0}^{L}\sum_{j=0}^{L}\sum_{k=0}^{K} c_{ij}x_{ijk} \tag{4}$$

$$\sum_{i=1}^{L} g_i y_{ik} \le q; \forall k \tag{5}$$

$$\sum_{k=1}^{K} y_{ik} = \begin{cases} 1 & i = 1,2,\ldots,L \\ K & i = 0 \end{cases} \tag{6}$$

$$\sum_{i=0}^{L} x_{ijk} = y_{ik} \quad j = 0,1,\ldots,L; \forall k \tag{7}$$

$$\sum_{j=0}^{L} x_{ijk} = y_{ik} \quad i = 0,1,\ldots,L; \forall k \tag{8}$$

$$x_{ijk} = 0\ or\ 1 \quad i,j = 0,1,\ldots,L; \forall k \tag{9}$$

$$y_{ik} = 0\ or\ 1 \quad i = 0,1,\ldots,L; \forall k \tag{10}$$

In the offered design:

Equation (4) is the objective task.

Equation (5) is the automobile volume constraint, and the total goods loaded by single automobile shall not cross its extreme carrying volume.

Equation (6) means that each customer has only one automobile to serve, and all tasks are completed by K automobiles.

Equation (7) indicates that there is and only one automobile arriving at a customer.

Equation (8) indicates that there is and only one automobile leaving a customer.

Equations (9) and (10) are integer restrictions.

3.2 Design of hybrid (PSO) procedure

Standard PSO process implementation process is as follows:

Step 1: Start the process with the element swarm, take the real quantity between 1 ~ (K + L-1) randomly for each dimension of each element situation vector x, and take the real quantity between - (K + L-2) ~ (K + L-2) randomly for each dimension of each velocity vector V, and set the parameters ω, C1, C2, R.

Step 2: Converts the situation vector of each element into the form of general route.

Step 3: Calculate the fitness rate of each element according to formula (6), and take the initial evaluation value as the individual extreme value $pbest_{i1}$, find the global extreme value gbest1.

Step 4: For each element, obtain the velocity V according to formula (7), and calculate the next assembly situation x according to formula (8). When calculating V and X, if it cross the range, take the value according to the boundary and convert x into the form of regular route.

Step 5: Calculate the fitness rate of each element according to formula (9) and compare it with $pbest_{id}$, $gbest_d$ and. If the fitness value is smaller, update $pbest_{id}$ or $gbest_d$.

Step 6: If the termination situations are not meet, step 4 is returned.

Mountain climbing scheme 1: for the global extremum in every generation assembly formed through standard PSO, the mountain climbing operation is implemented through domain search. This paper is realized by exchanging the situations of any two dimensions in the element situation vector. The specific operations are as follows:

1. For the global extremum in each generation, any two dimensions in the situation vector are randomly selected to exchange their situations.
2. Judge whether the fitness value becomes smaller after changing the situation. If it becomes smaller, replace the global extreme value with the individual after changing the situation.
3. Repeat 1 and 2 until a certain quantity of exchanges are reached.

The hybrid (PSO) procedure using mountain climbing scheme I is called hybrid PSO scheme I.

Mountain climbing scheme 2: in each iteration, the mountain climbing operation is implemented for each element through domain search. It is also realized by exchanging the situations of any two dimensions in the element situation vector. The specific operations are as follows:

1. For each element, any two dimensions in the situation vector are randomly selected to exchange their situations.
2. Judge whether the fitness value becomes smaller after changing the situation. If it becomes smaller, replace the individual extreme value of the element with the individual after changing the situation.
3. Repeat 1 and 2 until a certain quantity of exchanges are reached.
4. Compare the individual extreme value obtained after mountain climbing with the global extreme value. If the fitness value becomes smaller, update the global extreme value.

The hybrid (PSO) procedure using mountain climbing scheme II is called hybrid PSO scheme II.

Hybrid PSO scheme 1 and hybrid PSO scheme 2 are obtained by introducing mountain climbing procedure on the basis of standard PSO. The difference between the two procedures is that the time of introducing the mountain climbing procedure is different. The former is to climb the global extremum in each generation of population formed by standard PSO through domain search, and the latter is to climb each element through domain search in each iteration.

The purpose of introducing mountain climbing operation into (PSO) procedure is two: one is to enhance the local search ability of (PSO) procedure. The second is to prevent premature convergence without finding the optimal result. The local search ability of Hybrid PSO scheme 1 is greater than that of standard pso} Hybrid PSO scheme 2, which is fully proved in the later example analysis.

4 Results and analysis

This section describes outcome analysis obtained for from the offered design of route optimization based on PSO.

4.1 Example 1 analysis

The issue of example 1 is a dissemination system with 8 chain stocks and 1 Dissemination Center. The quantity of automobiles used for dissemination in the dissemination center is 2, and the automobile volume is eight tons. The distance among chain stocks (km) and its request (ton) are tabulated in Table 1. The dissemination center quantity is 0. It is compulsory to position appropriate driving routes to minimize the total transportation mileage.

The optimal total route observed through experimentation test is: 6-7-4-0-1-3-5-8-2, the corresponding driving route is:

Automobile 1: dissemination center - chain stock 6 - chain stock 7 - chain stock 4 - dissemination center

Automobile 2: dissemination center - chain stock 1 - chain stock 3 - chain stock 5 - chain stock 8 - chain stock 2 - dissemination center.

The optimized total driving area is 67.5 kms.

Due to the introduction of mountain climbing procedure, the convergence speed of hybrid procedure is much better than that of standard PSO procedure. The changes of the optimal element fitness values of standard PSO procedure, mixed PSO scheme I and mixed PSO scheme II with the quantity of iterations are shown in Figure 2, Figure 3and Figure 4 respectively. The observations through this Figure are that among the three procedures of standard PSO, mixed PSO scheme 1 and mixed PSO scheme 2, the convergence speed of mixed PSO scheme 2 is the fastest, the convergence speed of mixed PSO scheme is the second, and the convergence speed of standard PSO is the slowest.

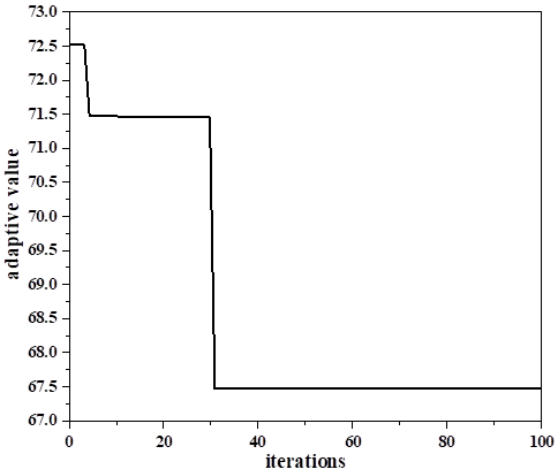


Figure 2: Growth of standard PSO optimal result.

Table 1: Distance and request between chain stocks

Chain stock Quantity	Requirement	Distance								
		0	1	2	3	4	5	6	7	8
0	0	0	4	6	7.5	9	20	10	16	8
1	1	4	0	6.5	4	10	5	7.5	11	10
2	2	6	6.5	0	7.5	10	10	7.5	7.5	7.5
3	1	7.5	4	7.5	0	10	5	9	9	15
4	2	9	10	10	10	0	10	7.5	7.5	10
5	1	20	5	10	5	10	0	7	9	7.5
6	4	10	7.5	7.5	9	7.5	7	0	7	10
7	2	16	11	7.5	9	7.5	9	7	0	10
8	2	8	10	7.5	15	10	7.5	10	10	0

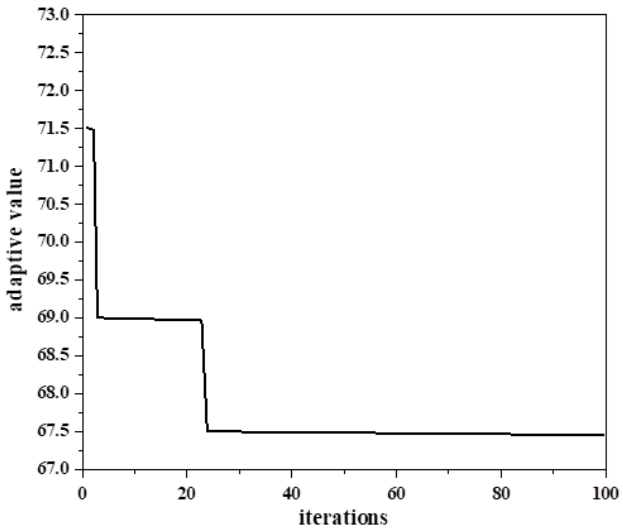


Figure 3: Growth of optimal result of Hybrid PSO scheme I.

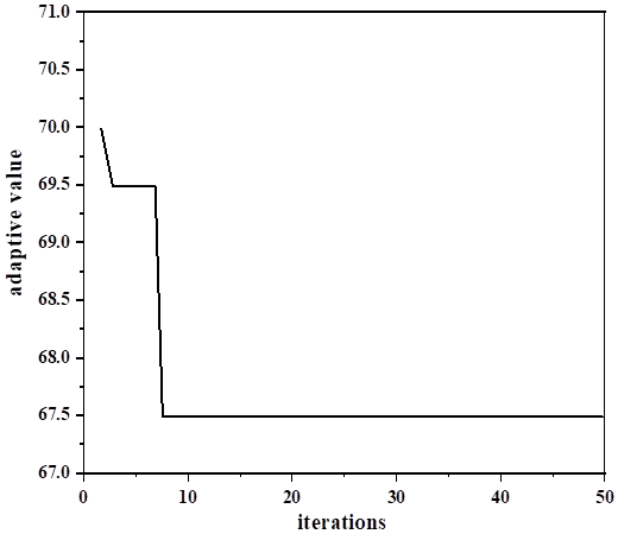


Figure 4: Growth of the second optimal result of the mixed PSO scheme.

The three procedures of standard PSO, mixed PSO scheme I and mixed PSO scheme II all carry out 20 operations, and the operation outcomes are that the standard PSO 20 operations and 5 operations reach the optimum. In the mixed PSO scheme 1, 20 operations and 8 operations reach the optimum, and the mixed PSO scheme 2 20 operations and 17 operations reach the optimum. It can be seen from the test outcomes that the stick PSO scheme 1 is better than the standard PSO under the same quantity of iterations. The reason is that the hybrid PSO scheme introduces the mountain climbing operation to the global extremum in every generation, which improves the local examination capability of the normal PSO. The quantity of iterations of Hybrid PSO scheme 2 is only half of that of standard PSO and hybrid PSO scheme 1, but the operation outcome is obviously better than that of standard PSO and hybrid PSO scheme 1. The reason is that hybrid PSO scheme 2 introduces mountain climbing operation for each element, which greatly increases the intensity of local search and makes up for the defect that (PSO) procedure is easy to fall into local optimization.

4.2 Example 2 analysis

There are only 8 dissemination points in example 1. In order to further test the performance of Hybrid PSO scheme 1 and hybrid PSO scheme 2, a dissemination system with issue scale of 20 is randomly generated in this paper. The coordinates of the dissemination center are (45 kms, 45 kms), the location coordinates of clients are (x km, y km), X and y are actual quantities among 0 and 100, and the cargo request is a random quantity between 0 and 2. See Table 2 and 3 for the coordinates of 20 clients and their cargo request. The load volume of automobiles in the dissemination center is 8t, and the quantity of automobiles in the dissemination center is 3. It is required to reasonably arrange the dissemination route of automobiles to minimize the dissemination mileage. For ease, the distance among clients and the distance between clients and dissemination center are straight-line distance. The example includes 20 clients, and the total quantity of clients is up to 2.433 x 1018. Due to time restrictions, this issue cannot be realized by exhaustive method. The graphical representation of customer coordinates and their request is presented in Figure5.

Table 2: Customer coordinates and their request

Client quantity	1	2	3	4	5	6	7	8	9	10
Abscissa	42	57	41	70	96	91	62	72	76	26
Ordinate	14	32	99	47	58	88	79	10	8	54
Requirement	0.3	0.4	1.2	1.5	0.8	1.3	1.1	0.6	1.2	0.4

Table 3: Customer coordinates and their request (Continued)

Client quantity	11	12	13	14	15	16	17	18	19	20
Abscissa	55	93	45	28	78	10	16	11	97	56
Ordinate	39	28	74	96	9	27	71	55	31	94
Requirement	0.9	1.3	0.7	1.9	1.7	1.1	1.5	1.6	1.2	1.5

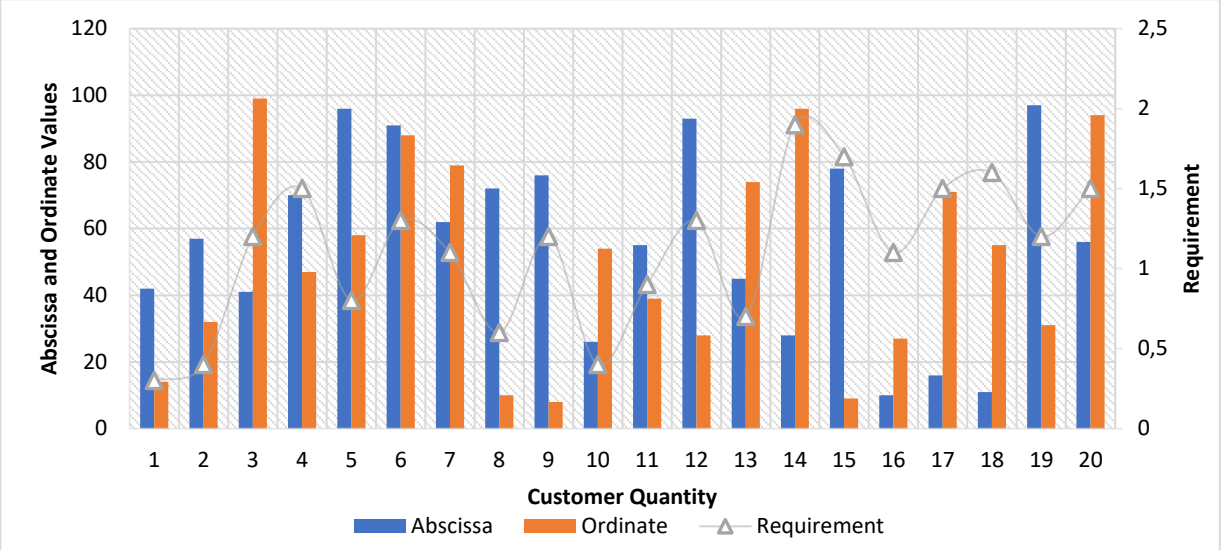


Figure 5: Graphical Representation of customer coordinates and their request.

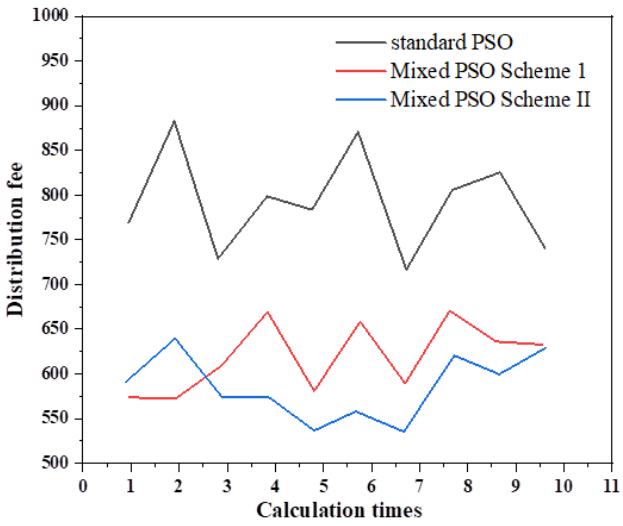


Figure 6: Comparison of outcomes of standard PSO, mixed PSO scheme I and mixed PSO scheme II.

The example includes 20 clients, and the total quantity of clients is up to 2.433 x 1018. Due to time restrictions, this issue cannot be realized by exhaustive method. The graphical representation of customer coordinates and their request is presented in Figure 5.

The three procedures of standard PSO, mixed PSO scheme I and mixed PSO scheme II all carry out 10 operations, and the operation outcomes are shown in Figure 6. It is observed that the optimal value obtained by standard PSO is 717.79. The average value of 10 results is 793.28. The optimal value of Hybrid PSO scheme 1 is 575.68, and the average value of 10 times is 620.94. The optimal value of Hybrid PSO scheme 2 is 535.90, and the average value of 10 results is only 588.52.

4.3 Discussion

It can be seen from the figure that after the issue scale increases, the performance of Hybrid PSO scheme 1 and hybrid PSO scheme 2 is also better than that of standard PSO method. Due to the introduction of mountain climbing operation for each element, the performance of Hybrid PSO scheme 2 is much better than that of standard PSO and hybrid PSO scheme 1. High quality results are obtained after 10 times of result. Using hybrid PSO scheme 2 to solve the logistics dissemination automobile routing issue can greatly shorten the dissemination mileage. This method is an effective method to solve the logistics dissemination automobile routing issue. Through the analysis of calculation outcomes of example 1 and example 2, the following conclusions can be drawn:

The local search ability of Hybrid PSO scheme 1 is greater than that of standard PSO} Hybrid PSO scheme 2. Due to the mountain climbing operation of each element, the local search ability of Hybrid PSO scheme 1 is greater than that of Hybrid PSO scheme 1 and standard PSO. The introduction of mountain climbing procedure makes up for the defect that (PSO) procedure is easy to fall into local optimization. Due to the introduction of mountain climbing procedure, the convergence speed of hybrid procedure is much better than that of standard PSO procedure. Among the three procedures of standard PSO, mixed PSO scheme 1 and mixed PSO scheme 2, the convergence speed of mixed PSO scheme 2 is the fastest, the convergence speed of mixed PSO scheme is the second, and the convergence speed of standard PSO is the slowest.

When the population size is roughly the same, the hybrid PSO scheme 2 is also better than the standard GA and the two population GA. When solving the automobile routing issue of logistics dissemination, the performance of Hybrid PSO scheme I and hybrid PSO scheme II are better than standard PSO, and the performance of Hybrid PSO scheme II is the best. Hybrid PSO scheme 2 can quickly and effectively solve the logistics dissemination automobile routing issue. Using this method to solve the logistics dissemination automobile routing issue can greatly shorten the

dissemination mileage. The procedure is simple and easy to program. It is a useful and realistic optimization technique to resolve the logistics dissemination automobile routing issue.

5 Conclusion

In this paper, the mathematical design of logistics dissemination automobile routing issue is established. Aiming at the lack of local search ability of (PSO) procedure, the mountain climbing procedure with strong local search ability is introduced into (PSO) procedure to improve the offered procedure. The hybrid (PSO) procedure is applied to the automobile routing issue of logistics dissemination. Combined with the features of logistics dissemination, two mountain climbing schemes are offered, and two different hybrid (PSO) procedures are constructed. For the global extremum in each generation, the climbing operation is introduced to form stick PSO scheme 1, and the climbing operation for each element forms Hybrid PSO scheme 2. The standard PSO, hybrid PSO scheme 1 and hybrid PSO scheme 2 are used to solve the logistics dissemination automobile routing issue. The example analysis displays that the hybrid PSO scheme 1 and hybrid PSO scheme 2 offered in this paper have better performance than the standard PSOA method in solving the automobile routing issue. It can efficiently resolve the logistics dissemination automobile routing issue. After the issue scale rises, the optimization advantages of Hybrid PSO scheme 2 are fully displayed. The article reveals that hybrid PSO scheme 2 is effective to solve the logistics dissemination automobile routing issue and can greatly shorten the dissemination mileage.

References

[1] Li, H., Yang, D., Su, W., Lü, J., & Yu, X. (2018). An overall distribution particle swarm optimization MPPT algorithm for photovoltaic system under partial shading. *IEEE Transactions on Industrial Electronics*, *66*(1), 265-275. *10.1109/TIE.2018.2829668*

[2] Cao, Y., Zhang, H., Li, W., Zhou, M., Zhang, Y., & Chaovalitwongse, W. A. (2018). Comprehensive learning particle swarm optimization algorithm with local search for multimodal functions. *IEEE Transactions on Evolutionary Computation*, *23*(4), 718-731. *10.1109/TEVC.2018.2885075*

[3] Memari, A., Ahmad, R., Rahim, A., Abdul, R., & Hassan, A. (2018). Optimizing a Just-In-Time logistics network problem under fuzzy supply and request: two parameter-tuned metaheuristics algorithms. *Neural Computing and Applications*, *30*(10), 3221-3233. *https://doi.org/10.1007/s00521-017-2920-0*

[4] Son, P. V. H., Duy, N. H. C., & Dat, P. T. (2021). Optimization of construction material cost through logistics planning model of dragonfly algorithm—particle swarm optimization. *KSCE Journal of Civil Engineering*, *25*(7), 2350-2359. *https://doi.org/10.1007/s12205-021-1427-5*

[5] Ha, M. P., Nazari-Heris, M., Mohammadi-Ivatloo, B., & Seyedi, H. (2020). A hybrid genetic particle swarm optimization for distributed generation allocation in power distribution networks. *Energy*, *209*, 118218. *https://doi.org/10.1016/j.energy.2020.118218*

[6] Liu, S., Liang, M., & Hu, X. (2018). Particle swarm optimization inversion of magnetic data: Field examples from iron ore deposits in China. *Geophysics*, *83*(4), J43-J59. *https://doi.org/10.1190/geo2017-0456.1*

[7] Ahmadian, A., Elkamel, A., & Mazouz, A. (2019). An improved hybrid particle swarm optimization and tabu search algorithm for expansion planning of large dimension electric distribution network. *Energies*, *12*(16), 3052. *https://doi.org/10.3390/en12163052*

[8] Ding, J., Wang, Q., Zhang, Q., Ye, Q., & Ma, Y. (2019). A hybrid particle swarm optimization-cuckoo search algorithm and its engineering applications. *Mathematical Problems in Engineering*, *2019*. *https://doi.org/10.1155/2019/5213759*

[9] Muhammad, M. H., Mahmoud, K. R., Hameed, M. F. O., & Obayya, S. S. A. (2018). Optimization of highly efficient random grating thin-film solar cell using modified gravitational search algorithm and particle swarm optimization algorithm. *Journal of Nanophotonics*, *12*(1), 016016. *https://doi.org/10.1117/1.JNP.12.016016*

[10] Wang, Y., Assogba, K., Fan, J., Xu, M., Liu, Y., & Wang, H. (2019). Multi-depot green vehicle routing problem with shared transportation resource: Integration of time-dependent speed and piecewise penalty cost. *Journal of Cleaner Production*, *232*, 12-29. *https://doi.org/10.1016/j.jclepro.2019.05.344*

[11] Yuan, F., Lv, K., Tang, B., Wang, Y., Yang, W., Qin, S., & Ding, C. (2021). Optimization design of oil-immersed iron core reactor based on the particle swarm algorithm and thermal network model. *Mathematical problems in Engineering*, *2021*. *https://doi.org/10.1155/2021/6642620*

[12] Moosavian, N., & Lence, B. (2019). Testing evolutionary algorithms for optimization of water distribution networks. *Canadian Journal of Civil Engineering*, *46*(5), 391-402. *https://doi.org/10.1139/cjce-2018-0137*

[13] Lagos, C., Guerrero, G., Cabrera, E., Moltedo, A., Johnson, F., & Paredes, F. (2018). An improved particle swarm optimization algorithm for the VRP with simultaneous pickup and delivery and time windows. *IEEE Latin America Transactions*, *16*(6), 1732-1740. *10.1109/TLA.2018.8444393*

[14] Wan, M., Gu, G., Qian, W., Ren, K., Chen, Q., & Maldague, X. (2018). Particle swarm optimization-based local entropy weighted histogram equalization for infrared image enhancement. *Infrared Physics & Technology*, *91*,

164-181.
https://doi.org/10.1016/j.infrared.2018.04.003

[15] Li, S., Zhang, Q., Zhang, Z., Zhao, Q., & Liang, L. (2021). Improved subgroup method coupled with particle swarm optimization algorithm for intra-pellet non-uniform temperature distribution problem. *Annals of Nuclear Energy*, *153*, 108070. *https://doi.org/10.1016/j.anucene.2020.108070*

[16] Sun, S. H., Yu, T. T., Nguyen, T. T., Atroshchenko, E., & Bui, T. Q. (2018). Structural shape optimization by IGABEM and particle swarm optimization algorithm. *Engineering Analysis with Boundary Elements*, *88*, 26-40. *https://doi.org/10.1016/j.enganabound.2017.12.007*

[17] Ceylan, O. (2021). Multi-verse optimization algorithm-and salp swarm optimization algorithm-based optimization of multilevel inverters. *Neural Computing and Applications*, *33*(6), 1935-1950. *https://doi.org/10.1007/s00521-020-05062-8*

[18] Wang, Y., Assogba, K., Liu, Y., Ma, X., Xu, M., & Wang, Y. (2018). Two-echelon location-routing optimization with time windows based on customer clustering. *Expert Systems with Applications*, *104*, 244-260. *https://doi.org/10.1016/j.eswa.2018.03.018*

[19] Silva, L. I., Belati, E. A., Gerez, C., & Silva Junior, I. C. (2021). Reduced search space combined with particle swarm optimization for distribution system reconfiguration. *Electrical Engineering*, *103*(2), 1127-1139. *https://doi.org/10.1007/s00202-020-01150-z*

[20] Chen, J., & Shi, J. (2019). A multi-compartment vehicle routing problem with time windows for urban distribution–A comparison study on particle swarm optimization algorithms. *Computers & Industrial Engineering*, *133*, 95-106. *https://doi.org/10.1016/j.cie.2019.05.008*

Improved Genetic Algorithm in Multi-objective Cargo Logistics Loading and Distribution

Zhilin He
Chengdu Agriculture college, Economic trade department, Chengdu, Sichuan, 611130, China
E-mail: zhilinhe3@163.com

Keywords: improved genetic algorithm, material distribution, path planning, optimization objectives

Received: February 2, 2022

To solve the problem of material distribution path planning in a production workshop, this paper proposes research on multi-objective cargo logistics loading and distribution based on an improved genetic algorithm. This paper improves the genetic algorithm to solve the problem (P), that is, the evolution model based on the genetic algorithm draws lessons from the coding mode of the genetic algorithm, and uses the row insertion method to obtain the initial population. The improved genetic algorithm is better than the traditional genetic algorithm. The rapid development of railway transportation towards high speed, high density, and heavy load has led to even higher requirements for the safety of railway signal equipment. The safety of railway signal equipment is an important part of ensuring railway traffic safety, thus, it is very necessary to study a system that can diagnose the fault of railway signal equipment according to the actual situation. This article utilizes the genetic algorithm of artificial intelligence for investigating the loading and distribution of logistics in transportation. It is demonstrated that genetic algorithm integration is an effective method to improve the performance of logistic distribution model. The convergence speed of the improved genetic algorithm is fast, and it shows a stable upward trend with the increase of the number of iterations.

Povzetek: Predlagan je izboljšan genetski algoritem za učinkovito logistično natovarjanje in distribucijo, kar izboljšuje uspešnost in hitrost logistične distribucije.

1 Introduction

With the rapid development of science and technology and the acceleration of the process of global economic integration, the logistics activities in the daily business activities of enterprises play a more and more important role in the global economic development, and their impact on all aspects of global economic activities is becoming more and more obvious [1]. Especially with the development of e-commerce websites, logistics has gradually become an important competitive field for enterprises [2]. Logistics distribution refers to integrated and systematic management that coordinates with the daily business activities of enterprises such as production, procurement, supply, and sales, and integrates the corresponding logistics activities such as distribution, handling, loading and unloading, storage, transportation, packaging, and information transmission. Its purpose is to provide customers with the best service at the lowest possible cost, to improve the overall economic benefits of the enterprise and enhance the overall competitive level of the enterprise [3].

At present, in China's logistics industry, important decision-making problems such as distribution center location, transportation route, inventory control and cargo assembly scheme are still in the state of semi manual decision-making, and the technical support of the whole logistics activities lags behind the global average level [4]. Like the vast majority of different businesses, sharing is a major problem for logistics factories now, from Uber-style ways to deal with last-mile conveyance, to more operations specialist organizations and organizations at big business level, cooperation is re-imagined in the entire SLN (sharing logistics network). By and large, various capabilities in SLN can be shared, from request assortment, pressing, arranging, to ship, stockpiling and conveyance, qualified conventional suppliers are allowed to go along with it and import their own accessible operations assets into asset pool for summon, and the coordinated suppliers perform brought together asset evening out to fulfil irregular strategies needs from different clients. Contrasted and the ordinary frameworks independently set-up, SLN with enormous clients prompts further developing individual asset productivity, defending by and large asset work, reducing transport expense, and saving conveyance time.

For instance, asset evening out in SLN can further develop responsibility by changing the rundown of operations task start times and re-establishing the priority connection among undertakings intra-stage, and furthermore can keep up with the power in asset productivity considering benefit assignment and administration criticism between stage. One might say that multi-stage asset evening out among customary suppliers is particularly essential to asset the board in SLN, which is the obvious elements of Industry 3.5. Multi-stage asset evening out issue in sharing planned operations network is a multi-objective improvement issue, which is firmly non-deterministic polynomial hard in open circle climate.

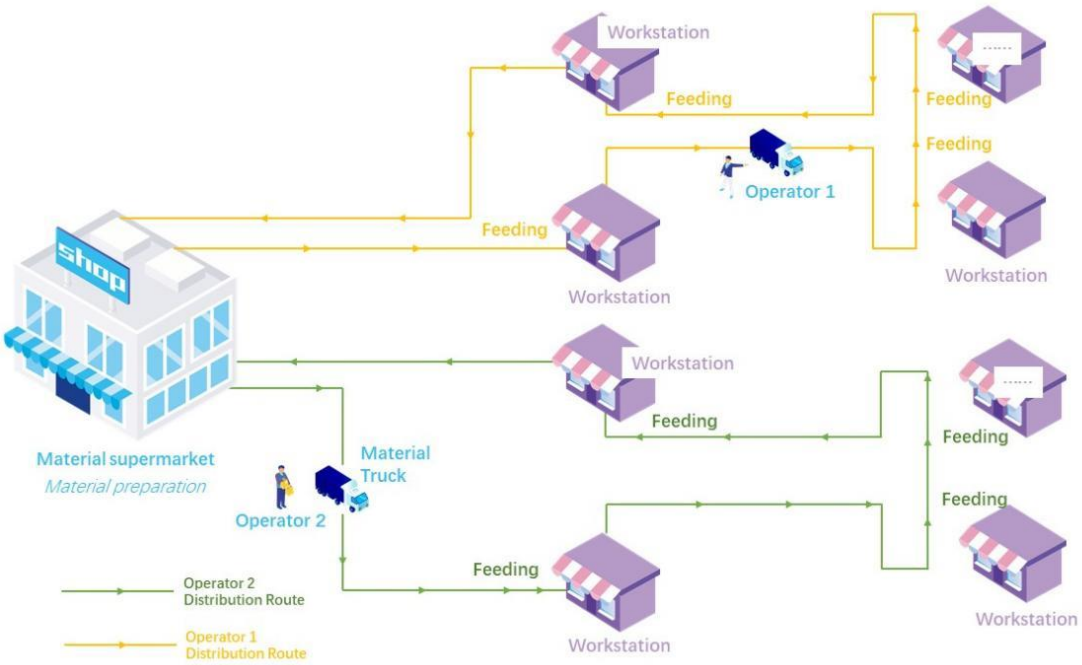


Figure 1: Material distribution process

Especially in the e-commerce environment with huge transaction volumes and extremely fast transaction speed, the contradiction between information flow and logistics has seriously affected the service quality of enterprises. Therefore, through the research on the optimization of distribution path, we can clarify the work details of each link of logistics distribution in enterprise business activities, and further improve the relevant technical system, to improve the overall intelligent level of the logistics industry. It is assumed that each operator's initial position is in the material supermarket. When the operator completes all batching tasks, he returns to the material supermarket. Dispatch the minimum number of operators to load the material truck according to the assigned task and send it to each workstation in turn until all materials are delivered and returned to the material supermarket. The material distribution process is shown in Figure 1 [5].

The rest of this article is organized as: Section 2 presents the related works in various domains. Section 3 consists of multi-objective mathematical model and its solution using genetic algorithm. Results and analysis are discussed in section 4 followed by concluding remarks in section 5.

2 Related work

Aiming at this research problem, Yun and Li changed the constraints in the distribution path and analyzed the impact of each index on the path [6]. Feng *et al.* proposed an optimization model aiming at the shortest distribution path and the least distribution times [7]. Sheng *et al.* used a hybrid particle swarm optimization algorithm to solve the material distribution model aiming at material transportation cost, transportation time, and line inventory [8]. Abbasi *et al.* proposed a material distribution method focusing on stations, and established a material flow path optimization model with the optimization objective of minimizing the total distribution time [9]. Wang *et al.* mainly establish models for VRP problems in two processes of auxiliary materials and replenishment and use a genetic algorithm to solve them with the goal of path optimization [10]. Saric *et al.* studied the just-in-time distribution of materials under the condition of limited distribution vehicles and solved it with an improved genetic algorithm and three-stage heuristic algorithm [11]. Mazhari *et al.* expounded on the problems existing in the material distribution process of the automobile assembly line and put forward the material distribution path optimization model with a time window with the goal of the shortest distribution distance [12]. Zan *et al.* proposed an optimization model to minimize distribution cost and solved the model by using dynamic programming and simulated annealing genetic algorithm [13]. Minaei *et al.* proposed an M- VRPTW research model based on the VRPTW problem [14]. Grisales *et al.* proposed a VRPTW mathematical model with time window constraints based on multi-supply points and place dependence and further summarized the research progress of VRPTW [15]. Based on the current research, this paper proposes the research of multi-objective cargo logistics loading and distribution based on an improved genetic algorithm.

A technique which considers wavelet frames for micropolar fluid flow is used for high mass transfer [16]. The vibration over the sandwich plates of laminated skew is studied through finite element [17]. The numerical simulation based on space time fractional equation are evaluated [18]. The proposed work can further be extended by using integration approaches of Artificial Intelligence and Machine learning as studied

from several studies [19-21]. This paper improves the genetic algorithm to solve the problem (P), that is, the evolution model based on the genetic algorithm draws lessons from the coding mode of the genetic algorithm, and uses the row insertion method to obtain the initial population. In the crossover operation, the narrow gene similarity is used to distinguish the chromosome similarity, and the double variation rate is added to the mutation operation in the evolution process. MATLAB is used to realize the algorithm, and compared with the traditional genetic algorithm to verify the feasibility and effectiveness of the proposed model and algorithm.

3 Mathematical model establishment

This section includes the description of model establishment including variables and providing solution based on genetic algorithm.

3.1 Decision variables

$$x_{ijk} = \begin{cases} 1, \text{Moving tool from station I to station J;} \\ 0, \text{Otherwise;} \end{cases} \quad (1)$$

$$y_{ik} = \begin{cases} 1, \text{Station I is serviced by handling tool K;} \\ 0, \text{Otherwise;} \end{cases} \quad (2)$$

3.2 Multi-objective mathematical model

According to the problems, the following mathematical models are constructed:

A. Optimization objectives

Minimize total distance traveled:

$$d = \min \sum_{k=1}^{K} \sum_{i=0}^{N} \sum_{j=0}^{N} d_{ij} x_{ijk} \quad (3)$$

Minimize vehicle usage:

$$v = \min \sum_{i=0}^{N} \sum_{k=1}^{K} x_{0jk} \quad (4)$$

Average satisfaction of the largest chemical industry:

$$u = \max \frac{1}{N} \sum_{i=1}^{N} u_i(t_i) \quad (5)$$

Maximize vehicle capacity:

$$l = \max \frac{\sum_{i=1}^{N} q_i}{KQ} \quad (6)$$

B. Constraints

$$ET_i \le t_i \le LT_i, i \in \{1,2,...,N\} \quad (7)$$

$$\sum_{i=1}^{N} q_{iQ} y_{ik} \le Q, k \in \{1,2,...,K\} \quad (8)$$

$$\sum_{k=1}^{K} y_{ik} = 1, i \in \{1,2,...,N\} \quad (9)$$

$$\sum_{I=0}^{N} x_{ijk} = y_{jk}, k \in \{1,2,...,K\} \quad (10)$$

$$\sum_{j=0}^{N} x_{ijk} = y_{jk}, k \in \{1,2,...,K\} \quad (11)$$

$$\sum_{i=0}^{N} x_{i0k} = 1, k \in \{1,2,...,K\} \quad (12)$$

$$\sum_{j=0}^{N} x_{0ik} = 1, k \in \{1,2,...,K\} \quad (13)$$

$$\sum_{i=1}^{R} \sum_{j=1}^{R} x_{ijk} \le |R| - 1, R \in \{1,2,...,n\} \quad (14)$$

Where n is the number of jobs; K is the number of distribution tools; Q is the carrying capacity; Q_i is the demand of station i; D_{ij} is the distance from stations i and j; D is the service time at station i; T_{ij} is the travel time from station i to station j; W_i is the waiting time at station i; t_i is the time to start service station i. Equation (3) ~ equation (6) represents the objective function, which is the driving distance of distribution tools, number of used tools, station satisfaction, and carrying rate in turn; Equation (7) means that the service shall not be started outside the time window; Equation (8) represents the carrying capacity constraint; Equation (9) indicates that each station has and only has one distribution tool service; Equations (10) to (11) represent the relationship between two variables x_{ijk} and y_{ik}; Equations (12) to (13) indicate that the distribution tools leave the warehouse and finally drive back to the warehouse; Equation (14) is a branch elimination constraint [22].

3.3 Solving the model with an improved genetic algorithm

The material distribution path planning problem in the production workshop is an NP problem. At present, many heuristic algorithms have their defects in solving multi-objective and multi-constraint optimization problems, such as slow convergence, premature convergence, unbalanced global optimization ability, and local search ability. Therefore, this paper improves the genetic algorithm to solve the problem (P), that is, based on the evolution model of the genetic algorithm, draws

lessons from the coding mode of the genetic algorithm, and uses the row insertion method to obtain the initial population. In the crossover operation, the narrow gene similarity is used to distinguish the similarity of chromosomes, and the double mutation rate is added to the mutation operation in the evolution process [23].

3.4 Construct chromosomes to produce the initial population

The quality of initial population generation determines the starting point of algorithm search. The high-quality population can accelerate the convergence of the algorithm and improve the solution efficiency [24]. When there is no feasible insertion position in the current distribution tool path, a new distribution tool is opened, and this paper improves it on this basis, as described in step 4.

Definition 1: The optimal insertion position is in the set of feasible insertion positions of the current path, and the initial feasible population generation steps are as follows:

Step 1: Randomly arrange N stations to obtain the station sequence set $X = \{x_1, x_2, \dots, x_N\}$, then assign it to the vehicle from left to right, and finally initialize the distribution path number Car_Code=0;

Step 2: Set car_Code = Car_Code + 1, the newly opened delivery tool, and its path number is Car_Code, arrange the leftmost station in the current x to the new path, and delete the station from X;

Step 3: Judge whether X is empty. If yes, go to step 5; Otherwise, go to step 4.

Step 4: Take the leftmost station in the current set X and judge whether there is a feasible insertion position in the current distribution tool path. If it exists, insert it into the best insertion position of the current distribution tool path, and turn to step 3; If it does not exist, judge whether it exists from small to large according to the path number of the distribution tool. If it exists, insert it into the best insertion position of the path, and turn to step 3, otherwise turn to step 2.

Step 5: Calculate the best start service time of the generated feasible chromosome to maximize the average station satisfaction of the chromosome.

Step 6: Repeat steps 2 to 5, and stop when the number of feasible chromosomes reaches the specified number. Among them, the feasible insertion location and the best service start time are determined as follows:

i. Determination of feasible insertion position

If the feasible distribution path of an established distribution tool is B and the start service time at the station x_i is t_{xi}, the maximum delay time for the distribution tool to arrive at the station is:

$$\max\_ pone(x_i) = \begin{cases} LT_{xi} - t_{xi} + W_{xi}, i = p \\ \min\left\{ LT_{xi} - t_{xi} + W_{xi}, \\ \max\_ pone(x_{i+1}) \right\}, 1 \le i \le p-1 \end{cases} \tag{15}$$

If a station x_s is inserted after the i^{th} station in the feasible path, the path is still feasible, then condition 1 capacity constraints:

$$\sum_{i=1}^{p} q_{xi} + q_s \le Q \tag{16}$$

Condition 2-time constraints, the time when the delivery tool reaches the station s shall not be later than the lower limit of the time window of the station, that is, when $t_{xi} + T_{xi} + T_{xi,s} \le LT_s$ and $i = 0$, $t_{xi} = T_{xi} = 0$, $T_{x1,s} = T_{0,s}$. When $i \le P - 1$ the delivery tool must meet the following requirements when arriving at the station X_{i+1}:

$$\max\{t_{xi} + T_{xi} + T_{xi,s}, ET_s\} + T_s + T_{s,xi+1} \le t_{xi+1} + \max\_ pone(x_{i+1}) - W_{xi+1} \tag{17}$$

ii. Determination of the best service start time

In the feasible path obtained by the above operation, the comprehensive satisfaction of the station is not the best. Based on the research on the determination method of the best start service time, this paper adjusts the time from the last station of the distribution tool path, omits the determination of the non-push able part, and improves the calculation of the maximum deferrable amount.

The calculation steps of the best service start time are as follows:

Step 1: Assume that the distribution path of a distribution tool is $(x_1, x_2, \dots, x_p)$.

Step 2: Divide the delivery tool path into several segments $(x_i, x_{i+1}, x_{i+2}, \dots, x_j)$, meet condition $W_{xi+1} = W_{xi+2} = \cdots = W_{xj} = 0$, when $i \ne 1$, $W_{xi} \ne 0; W_{xj} < p, W_{xj+1} \ne 0$.

Step 3: Take the last segment of the distribution path and set it as $(x_1, x_2, \dots, x_m)$, Adjust the service start time of the station $x_m, x_{m-1}, \dots, x_2, x_1$ in sequence; Under the condition of meeting the time constraint, the maximum delay to station $x_s (s = 1, 2, \dots, m)$ is $t_0 = \max\{t' | \sum_{n=s}^{m} u_{xn}'(t_{xn} + t') \ge 0\}$.

Step 4: Take the previous segment of the adjusted segment in the current distribution tool path and adjust it according to the method in step 3; If the segment forms a new segment with the latter segment in the current distribution tool path, readjust the newly formed segment; Otherwise, go to step 5.

Step 5: Continue to repeat step 4 until the service start time of all stations in the distribution tool path is determined.

3.5 Fitness function

The fitness of each individual is calculated according to the optimization target value and weight parameter. The solving function is as follows:

$$f(h_k)=\alpha_1\frac{d_k}{d_{\max}}+\alpha_2\frac{v_k}{v_{\max}}+\alpha_3\left(1-\frac{u_k}{u_{\max}}\right)+\alpha_4\left(1-\frac{l_k}{l_{\max}}\right)$$
$$\sum_{i=1}^{4}\alpha_i=1, a_i\geq 0, i=1,2,3,4 \quad (18)$$

Where, h_k represents chromosome K; d_k, v_k, u_k, l_k represent the driving distance of chromosome h, the total number of distribution tools used, the average satisfaction of stations, and the carrying rate of distribution tools; $d_{max}, v_{max}, u_{max}, l_{max}$ represents the maximum travel distance, the maximum number of distribution tools used, the maximum average satisfaction of stations, and the maximum carrying rate of chromosomes in the current population; $\alpha_i(i=1,2,3,4)$ is the weight coefficient [25].

4 Results and analysis

This section illustrates the result and analysis of proposed system using genetic algorithm. The distance between stations is measured along with the driving schedule which is presented in this section.

In material distribution in the production workshop, there are 9 stations and a warehouse, and the carrying capacity of the distribution tool is 12 units. The service time, reservation time, and demand of each station are shown in Table 1, and the driving time and distance between each station are shown in Tables 2 and 3; Improve the basic parameters of the genetic algorithm and take the population size pop_Size = 100, number of iterations Max_Gen = 200, selection probability $P_x = 0.8$, crossover probability $P_c = 0.8$, double mutation probability Local_Pm = 0.1 and Global_Pm=0.2.

Table 1: Information about each station

Station number	1	2	3	4	5	6	7	8	9
Service time/min	3.1	5.1	7.8	5.2	4.2	6.9	7.6	6.2	5.5
Reservation time / min	33,37,39	5.7.4.10	26,27,30	18,20,22	15,18,21	31,34,36	9,11,12	22,25,27	29,33,35
Demand/ unit	3.8	5.2	2.7	2.8	4.4	4.8	5.7	2.7	4.7

Table 2: Distance between stations and Driving Schedule (stations 1-4)

Station number		Driving time/min				
		Warehouse	Station	Station	Station	Station
		0	1	2	3	4
Driving distance/m	0	0	10	5	9	11
	1	23.43	0	6	5	7
	2	27.12	24.10	0	9	7
	3	12.29	16.54	15	0	6
	4	26.49	36.59	18.49	21.07	0
	5	11.59	33.31	30.21	20	21.56
	6	7.11	29.30	32.42	18. 65	28.69
	7	7.79	26.79	34.39	19.51	33.49
	8	21.89	25.31	43.59	28.89	47.79
	9	20.11	7.89	29.79	18.35	39.80

According to the above algorithm, Matlab simulation is used to calculate under different weight settings. The results are shown in Table 4.

It can be seen from Table 4 that the setting of optimization target weight parameters has a great impact on the experimental results [26]. d. L is closely related to v, that is, d increases with the increase of v, and l decreases with the increase of 0; u is opposite to d, v and l, that is, the increase of u is at the cost of the increase of d and v and the decrease of L [21].

Table 3: Distance between stations and Driving Schedule (stations 5-9)

Station number		Driving time/min				
		Station 5	Station 6	Station 7	Station 8	Station 9
	0	5	11	7	10	10
	1	8	6	8	11	10
	2	9	12	7	11	11
	3	6	11	6	6	9
Driving distance/m	4	8	6	11	9	10
	5	0	9	10	12	9
	6	9.59	0	10	10	9
	7	14.69	6.15	0	5	6
	8	31.39	22.39	17.19	10	10
	9	31.69	25.79	21.89	17.51	0

Table 4: Simulation results

Weight setting				Experimental result			
α_1	α_2	α_3	α_4	d	v	u	l
0.7	0.1	0.1	0.1	209.31	5	0.849	0.621
0.1	0.7	0.1	0.1	227.53	4	0.679	0.767
0.1	0.1	0.7	0.1	243.31	6	0.951	0.521
0.1	0.1	0.1	0.7	231.21	4	0.761	0.776

It can be seen from Table 4 that the setting of optimization target weight parameters has a great impact on the experimental results [20]. d. L is closely related to v, that is, d increases with the increase of v, and l decreases with the increase of 0; u is opposite to d, v and l, that is, the increase of u is at the cost of the increase of d and v and the decrease of L [27].

To prove the effectiveness of the improved genetic algorithm proposed in this paper, the improved genetic algorithm takes the weights $\alpha_1 = 1$, $\alpha_2 = 1$ and $\alpha_3 = 1$ respectively, which is compared with the corresponding single objective solution of the traditional genetic algorithm ($\alpha_4 = 1$ is not analyzed here because l is inversely proportional to v). The traditional genetic algorithm is used to calculate this example. The initial population size is 100, the selection probability and

crossover probability are 0.8, and the mutation probability is 0.1 [28].

As shown in Figure 2, when $\alpha_1 = 1$, the improved genetic algorithm shows a stable downward trend after 30 generations and converges to 55 generations [29]. However, the convergence speed of the traditional genetic algorithm is very slow in the middle and late stages, and it does not begin to converge until generation 126. As shown in Figure 3, when $\alpha_2 = 1$, the improved genetic algorithm has no fluctuation. From the whole image, the two have a downward trend, that is, the connecting line between the starting point and the convergence point. The slope of the improved genetic algorithm is significantly greater than that of the traditional genetic algorithm [30].

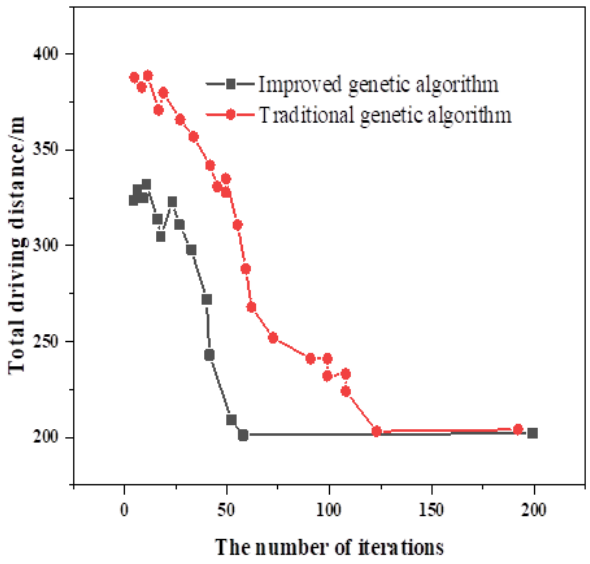


Figure 2: Comparison of travel distance convergence of distribution tools.

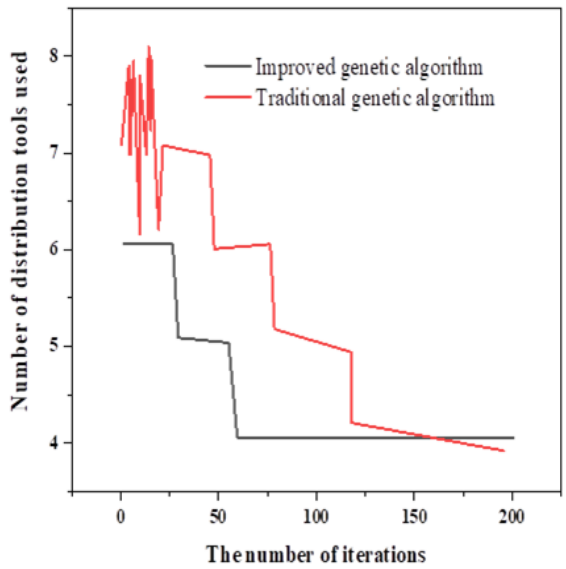


Figure 3: Convergence comparison of usage number of distribution tools.

As shown in Figure 4, when $\alpha_3 = 1$, the convergence speed of the improved genetic algorithm is also fast, showing a stable upward trend with the increase in the number of iterations [31]. To sum up, from the convergence diagrams in Figures 2, 3, and 4, it is obvious that the improved genetic algorithm is superior to the traditional genetic algorithm [32].

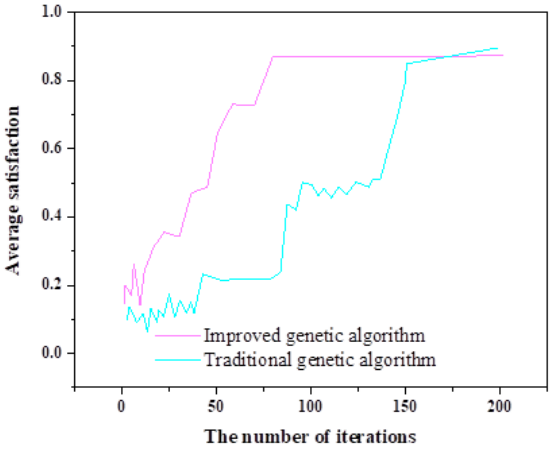


Figure 4: Convergence comparison of station average satisfaction.

In the crossover operation, the narrow gene similarity is used to distinguish the chromosome similarity, and the double variation rate is added to the mutation operation in the evolution process. The basic parameters of the genetic algorithm are improved and the population size pop is taken_ size = 100, number of iterations Max_ gen = 200, the selection probability P_x = 0.8, crossover probability $P_c = 0.8$, double mutation probability Local_ Pm = 0.1 and Global_ Pm=0.2. Matlab simulation is used to calculate under different weight settings. When $\alpha_1 = 1$, the improved genetic algorithm shows a stable downward trend after 30 generations and converges to 55 generations; However, the convergence speed of the traditional genetic algorithm is very slow in the middle and late stages, and it does not begin to converge until generation 126. When $\alpha_2 = 1$, the improved genetic algorithm has no fluctuation. From the whole image, we can see the downward trend between the two, that is, the connection between the starting point and the convergence point. The slope of the improved genetic algorithm is significantly greater than that of the traditional genetic algorithm. When $\alpha_3 = 1$, the convergence speed of the improved genetic algorithm is fast, and it shows a stable upward trend with the increase in the number of iterations.

5 Conclusions

In this paper, based on the research of multi-objective cargo logistics loading and distribution based on an improved genetic algorithm, a multi-objective mathematical programming model under the condition of fuzzy station reservation time is constructed, and an improved genetic algorithm is proposed to solve the model. In the algorithm design, the feasible insertion method is used to generate the initial population to speed

up the convergence speed. The concept of narrow gene similarity is proposed to avoid the cross operation of individuals with high similarity. At the same time, the double mutation mechanism is added to the mutation process to control the search range and convergence speed of the global solution space. Finally, the algorithm is implemented by MATLAB and compared with the traditional genetic algorithm to verify the feasibility and effectiveness of the proposed model and algorithm. The next research will consider the dynamic changes in workstation material requirements. In future work this research, delicate time window should to be thought of as because of the impediment of difficult time window in this paper, which can recognize the distinction among calculations. Moreover, the double effect of vulnerability from resource supply and client request is one more worry in proposed system.

References

[1] Wang, X., Chang, J., Meng, X., & Wang, Y. (2017). Research on multi-objective operation based on improved NSGA-II for lower Yellow River. *J. Hydraul. Eng*, *48*, 135-145. *10.13243/j.cnki.slxb.20151361*

[2] Geng, K., Ye, C., & Liu, L. (2020). Research on multi-objective hybrid flow shop scheduling problem with dual resource constraints using improved memetic algorithm. *IEEE Access*, *8*, 104527-104542. 10.1109/ACCESS.2020.2999680

[3] Mei, K., Li, D., Yoon, S. W., & Ryu, J. H. (2016). Multi-objective optimization of collation delay and makespan in mail-order pharmacy automated distribution system. *The International Journal of Advanced Manufacturing Technology*, *83*(1), 475-488. https://doi.org/10.1007/s00170-015-7555-7

[4] Arabzad, S. M., Ghorbani, M., & Tavakkoli-Moghaddam, R. (2015). An evolutionary algorithm for a new multi-objective location-inventory model in a distribution network with transportation modes and third-party logistics providers. *International Journal of Production Research*, *53*(4), 1038-1050. https://doi.org/10.1080/00207543.2014.938836

[5] Zhang, L. X. (2019). Research on similarity measurement of video motion images based on improved genetic algorithm in paper industry. *Paper Asia*, *2*(1), 135-138. 10.1155/2022/3936958

[6] Yun, Y., & Li, Y. (2016). An Improved Nonlinear Multi-Objective Optimization Problem Based on Genetic Algorithm. *International Journal of Hybrid Information Technology*, *9*(7), 361-372. (2016), pp. 361-372 http://dx.doi.org/10.14257/ijhit.2016.9.7.33

[7] Feng, Q., Yang, F., Xu, X., Zhang, B., Ding, Y., & Liu, Q. (2019). Multi-objective optimization genetic algorithm for multi-point light focusing in wavefront shaping. *Optics Express*, *27*(25), 36459-36473. https://doi.org/10.1364/OE.27.036459

[8] Sheng, W., Liu, K. Y., Liu, Y., Meng, X., & Li, Y. (2014). Optimal placement and sizing of distributed generation via an improved nondominated sorting genetic algorithm II. *IEEE Transactions on power Delivery*, *30*(2), 569-578. 10.1109/TPWRD.2014.2325938

[9] Abbasi, M., Mohammadi Pasand, E., & Khosravi, M. R. (2020). Workload allocation in iot-fog-cloud architecture using a multi-objective genetic algorithm. *Journal of Grid Computing*, *18*(1), 43-56. https://doi.org/10.1007/s10723-020-09507-1

[10] Wang, C., Feng, Y., Ye, J., Luo, B., & Liu, K. (2017). Multi-objective parameters optimization of centrifugal slurry pump based on RBF neural network and NSGA-genetic algorithm. *Transactions of the Chinese Society of Agricultural Engineering*, *33*(10), 109-115.

[11] Šarić, M., Hivziefendić, J., & Konjić, T. (2019). Multi-objective DG allocation in a radial power distribution network for power loss reduction, voltage profile improvement and investment deferral. *Elektrotehniski Vestnik*, *86*(5), 253-258.

[12] Mazhari, S. M., Monsef, H., & Romero, R. (2015). A multi-objective distribution system expansion planning incorporating customer choice on reliability. *IEEE Transactions on Power Systems*, *31*(2), 1330-1340. 10.1109/TPWRS.2015.2430278

[13] Zan, T. T. T., Gupta, P., Wang, M., Dauwels, J., & Ukil, A. (2018). Multi-objective optimal sensor placement for low-pressure gas distribution networks. *IEEE Sensors Journal*, *18*(16), 6660-6668. 10.1109/JSEN.2018.2850847

[14] Minaei, A., Sabzkouhi, A. M., Haghighi, A., & Creaco, E. (2020). Developments in multi-objective dynamic optimization algorithm for design of water distribution mains. *Water Resources Management*, *34*(9), 2699-2716. https://doi.org/10.1007/s11269-020-02559-8

[15] Grisales, L. F., Grajales, A., Montoya, O. D., Hincapie, R. A., Granada, M., & Castro, C. A. (2017). Optimal location, sizing and operation of energy storage in distribution systems using multi-objective approach. *IEEE Latin America Transactions*, *15*(6), 1084-1090. 10.1109/TLA.2017.7932696

[16] Kumbinarasaiah, S., & Raghunatha, K. R. (2021). A novel approach on micropolar fluid flow in a porous channel with high mass transfer via wavelet frames. *Nonlinear Engineering*, *10*(1), 39-45. *https://doi.org/10.1515/nleng-2021-0004*

[17] Dhotre, P. K., & Srinivasa, C. V. (2021). On free vibration of laminated skew sandwich plates: A finite element analysis. *Nonlinear Engineering*, *10*(1), 66-76. *https://doi.org/10.1515/nleng-2021-0006*

[18] Pedram, L., & Rostamy, D. (2021). Numerical simulations of stochastic conformable space–time

fractional Kortewegde Vries and Benjamin–Bona–Mahony equations. *Nonlinear Engineering*, *10*(1), 77-90.
https://doi.org/10.1515/nleng-2021-0007

[19] Wang, H., Hao, L., Sharma, A., & Kukkar, A. (2022). Automatic control of computer application data processing system based on artificial intelligence. *Journal of Intelligent Systems*, *31*(1), 177-192.
https://doi.org/10.1515/jisys-2022-0007

[20] Sun, L., Gupta, R. K., & Sharma, A. (2022). Review and potential for artificial intelligence in healthcare. *International Journal of System Assurance Engineering and Management*, *13*(1), 54-62.
https://doi.org/10.1007/s13198-021-01221-9

[21] Cai, Y., & Sharma, A. (2021). Swarm intelligence optimization: an exploration and application of machine learning technology. *Journal of Intelligent Systems*, *30*(1), 460-469.
https://doi.org/10.1515/jisys-2020-0084

[22] Kaur, R., & Kinger, S. (2014). Enhanced genetic algorithm-based task scheduling in cloud computing. *International Journal of Computer Applications*, *101*(14), 1-6.
10.5120/17752-8653

[23] Abdelaziz, M. (2017). Distribution network reconfiguration using a genetic algorithm with varying population size. *Electric Power Systems Research*, *142*, 9-11.
https://doi.org/10.1016/j.epsr.2016.08.026

[24] Wang, H. K., Chien, C. F., & Gen, M. (2015). An algorithm of multi-subpopulation parameters with hybrid estimation of distribution for semiconductor scheduling with constrained waiting time. *IEEE Transactions on Semiconductor Manufacturing*, *28*(3), 353-366.
10.1109/TSM.2015.2439054

[25] Sun, Y., Yen, G. G., & Yi, Z. (2018). Improved regularity model-based EDA for many-objective optimization. *IEEE Transactions on Evolutionary Computation*, *22*(5), 662-678.
10.1109/TEVC.2018.2794319

[26] Ebrahimi, S., Ghorbani, B., & Vijayaraghavan, K. (2017). Optimization of catalyst distribution along PEMFC channel through a numerical two-phase model and genetic algorithm. *Renewable Energy*, *113*, 846-854.
https://doi.org/10.1016/j.renene.2017.06.067

[27] Sankary, N., & Ostfeld, A. (2017). Scaled multiobjective optimization of an intensive early warning system for water distribution system security. *Journal of Hydraulic Engineering*, *143*(9), 04017025.
https://doi.org/10.1061/(ASCE)HY.1943-7900.0001317

[28] Manjunath Patel, G. C., Krishna, P., & Parappagoudar, M. B. (2016). Modelling and multi-objective optimisation of squeeze casting process using regression analysis and genetic algorithm. *Australian Journal of Mechanical Engineering*, *14*(3), 182-198.
https://doi.org/10.1080/14484846.2015.1093231

[29] Santos, M. C., Machado, A. R., Barrozo, M. A. S., Jackson, M. J., & Ezugwu, E. O. (2015). Multi-objective optimization of cutting conditions when turning aluminum alloys (1350-O and 7075-T6 grades) using genetic algorithm. *The International Journal of Advanced Manufacturing Technology*, *76*(5), 1123-1138.
https://doi.org/10.1007/s00170-014-6314-5

[30] Liang, C., Li, M., Lu, B., Gu, T., Jo, J., & Ding, Y. (2017). Dynamic configuration of QC allocating problem based on multi-objective genetic algorithm. *Journal of Intelligent Manufacturing*, *28*(3), 847-855.
https://doi.org/10.1007/s10845-015-1035-7

[31] Nazarahari, M., Khanmirza, E., & Doostie, S. (2019). Multi-objective multi-robot path planning in continuous environment using an enhanced genetic algorithm. *Expert Systems with Applications*, *115*, 106-120.
https://doi.org/10.1016/j.eswa.2018.08.008

[32] Guo, Q. (2018). Research on the application of improved shuffled frog leaping algorithm in mechanical fault diagnosis. *Academic Journal of Manufacturing Engineering*, *16*(2).

Potential Impact of Climate Change on Groundwater Level Declination in Bangladesh: A Mathematical Modeling Computation

Md. Anarul Islam, Md. Haider Ali Biswas
Mathematics Discipline, Science Engineering and Technology School, Khulna University, Khulna-9208, Bangladesh
E-mail: anarul1996@gmail.com, mhabiswas@yahoo.com

Keywords: groundwater, mathematical modeling, equilibrium point, stability analysis and numerical simulations

Received: November 4, 2020

Groundwater is one of the most important natural resources for the human being due to its ecological diversity. It has become a vastly vital and dependable source of water in all climatic regions together with each urban and rural areas of developed and developing countries like Bangladesh. Groundwater level declines for many reasons, some of which occurs due to natural phenomena and others are caused by human's activities and it has been declining since the introduction of deep tube wells and shallow tube wells after 1970s. Excessive demand of water, evaporation, pollution, deforestation, poor storage, low rainfall, urbanization, frequent pumping of water are the important causes of groundwater level declination in Bangladesh as well as all over the world. Taking these potential effects of climatic change into account, we formulate a mathematical model of groundwater level declination with the help of a system of nonlinear ordinary differential equations (ODEs). The model has been analyzed by finding the existence of equilibrium points and also the conditions of stability and instability near the equilibrium points have been derived by using the stability theory of non-linear differential equations with Lyapunov function and phase portrait analysis. Finally, the numerical simulations have been performed to illustrate the effect of pollution, deforestation, frequent pumping of water and evaporation on the groundwater level in support of analytical findings. Our study shows that, groundwater level decreases significantly due to over pumping, pollution, deforestation and global warming.

Povzetek: Padajoča raven podzemne vode je povezana s človekovimi aktivnostmi in naravnimi pojavi. V študiji je bil uporabljen matematični model za analizo vpliva onesnaževanja, sečnje gozdov in prekomernega črpanja podzemne vode.

1 Introduction

Groundwater is an essential part of public water supply and food production in Bangladesh and also significant natural resource that constitutes 95 percent of the freshwater [27]. About three billion people depend directly upon for drinking water and 40% of the world's food is produced from irrigated agriculture that relies mainly on groundwater. Moreover, in the rural area, farmers are used freshwater about 70% for the purpose of cultivation and by 2050, nine billion people will depend upon an approximated 50% increase in irrigated agricultural production [22]. In Bangladesh, 45,400 deep tubewells, 1,533,920 shallow tubewells and 170,470 low lift pumps are employed presently to provide water for cultivation and using groundwater around 79% of the total cultivated area is irrigated [27]. Nowadays, groundwater acts as the mainstay for agriculture in the southern region of Asia like India, Sri Lanka, Pakistan, Bangladesh, and the Northern China. The water reserved in the ground can be compared to money kept in a bank account. If we pick off money at a swift rate than savings new money, we will ultimately start having account-supply problems. Over pumping of water from the ground level rapidly than it is replenished for the lengthy period causes' similar problems [26].

Thus, groundwater level declination is a serious threat to the environment. It is primarily caused by sustained groundwater pumping. Moreover, due to poor reserving system, lots of rain water is wasted and the scientific development in procuring rain water is very poor. In case of saltwater contamination in the deep within the ground, a vast amount of usable water is reduced. Rainfall has diminished considerably over the years and it is the cause of large-scale deforestation and also rigorous climatic changes. These add to the reduction in water resources as well as groundwater resources. For agricultural use ground water is pumped more frequently, this can be done by fixing a tube well. Most of the industrial waste water are dumped to the water sources. This can be considered another reason for declination of water resources. Due to global warming and change in the climate, huge amounts of surface water and ground water is being evaporated, these also accumulate to the reduction in water resources as well as groundwater resources [51].

The influential consequences of groundwater declination are considered as drying up of tubewells, reduction of water in rivers, lakes i.e. in the land surface, degradation of water quality, increased the costs of pumping due to interrelation between ground and surface

water. The rivers, streams and lakes connected to groundwater can also have their supply declined while groundwater is overused [22].

Mathematical biology has become a well-recognized area of environmental research which is the most enjoyable applications of basic sciences [15, 42]. The aim of the mathematical biology is to represent a real phenomenon by mathematics and develop a model on the basis of conservation principle of biological processes, exploitation techniques by using the tools of mathematics. It has both practical and theoretical applications in the field of biotechnological research, biomedicine, ecology and so on [8, 31, 36]. Mathematical simulations and modeling play an incredible role for providing quantitative insight into different field of science. It has already contributed to a better understanding of the mechanisms in various field nowadays. Mathematical modeling has gotten attention because modeling and simulation of any physical phenomena allows us for rapid assessment. So, it is mainly used to describe the real phenomena which lead to design better prediction, management and control strategies [9].

Groundwater declination problem has been discussed by many researchers with different view of points. Declination rate and trend line analysis of groundwater level underneath Dhaka and Gazipur city in Bangladesh is described in [44]. A scientific study is discussed for the management and development of groundwater by using integrated remote sensing with geographic information systems [1]. The quality assessment of groundwater for the usefulness of drinking water is analyzed in [34]. A Mathematical model was proposed and inspected to study the impacts of predicted climate mutation on recharge and groundwater levels [49], whereas the effects of climate change on vegetation is characterized in [5] that is the important factor to store groundwater for long period. Another model based on the causes and quantification was also approached to describe the declination of groundwater level and aquifer dewatering in Dhaka metropolitan area in Bangladesh [230]. However, some articles [2-4, 7, 10-13, 15, 19-20, 21, 24, 29, 32, 38-48, 50, 52, 55-57] are referred for more details in the explanation of global climate change in Bangladesh as well as all over the world, interactions among atmospheric, surface and groundwater with the significance factors for depletion of ground layer and development of mathematical model on groundwater level declination.

In this paper, we would like to propose a new mathematical model on groundwater level declination. The goal of this study is to describe the impacts of climate change in the environment like deforestation, pollution, evaporation and frequent pumping of water on groundwater level. We formulate a mathematical model and study the existence and stability of the model with nonlinear Lyapunov function. Lastly, numerical simulations will be performed to show the effectiveness of the analytical analysis of groundwater model.

The manuscript is organized as follows. In section 2.1, we describe the present scenery of groundwater level declination perspective to Bangladesh; materials and methods of the study are given in section 2.2. We propose a new mathematical model for groundwater level declination in section 3.1. We evaluate the boundedness and positivity of the solution of the model, in subsections 3.2.1 and 3.2.2. and 3.3 we calculated the equilibrium analysis, in 3.4, we discuss the local stability and global stability of the model at both of the equilibrium points. The numerical simulation and sensitivity analysis of the groundwater model is then presented in section 4. In section 5, we observe the dynamical behavior of the model by plotting phase-plane diagram and we conclude the study with some recommendations for preventive steps to reduce the declination of groundwater level in section 6.

2.1 Groundwater declination perspective to bangladesh

Groundwater is an imperishable, renewable and potential natural resource in Bangladesh. In the wet or monsoon season (June–September), groundwater sustains it's dynamism by recharging, through rainfall and flooding and seceding during the dull season (March–May) due to withdrawal for various application, especially for cultivation. Agricultural yield of Bangladesh has successively increased over the last few decades while we have reasons to be happy about to fact that we must ensure this is not just a fleeting surge but a sustainable propensity. Professor E. M. Bean from University of Guelph observed that the region irrigated by groundwater in Bangladesh increased from 4 to 70 percent from 1972 to 1999 [19]. This research also investigated that due to the dependency of surface water and groundwater to one another, barrages built upstream of rivers entering Bangladesh block down the natural flow rates and it affects the groundwater recharge in the areas of downstream [6]. The average annual rainfall of the country is at least 2000 mm per year in most of the districts in Bangladesh with the exception of the comparatively dry western area of Rajshahi, where the average rainfall is approximated 1650 mm. The northeastern part of Bangladesh receives the greatest average precipitation, sometimes over 4000 mm per year and near about 80% of Bangladesh's rain falls during the monsoon season [56]. Before the extensive evocation of groundwater, water tables in the mostly unconfined aquifers of Bangladesh were generally shallow with a feeble seasonal vacillating propensity from the mid-2000s and onwards. Groundwater level were near or very close to the ground surface in some places after the monsoon period and at the end of the dry season, groundwater tables declined due to evaporation, evapotranspiration and inter-basin flow out of the aquifers. Average depth of groundwater level in Bangladesh is given from 2000 to 2018 in Figure 1.

Figure 1: Groundwater level has declined during dry season since 2000 to 2018 in Bangladesh [6].

2.2 Materials and methods

In this study, a mathematical model on groundwater level declination with the help of a system of nonlinear ordinary differential equations has been developed. We evaluate the biological feasibility and mathematical exactness of the model with boundedness and positivity analysis of the solution by using Lemma 1 and Lemma 2. We calculate the equilibrium analysis and study the local stability and global stability of the model at both of the equilibrium points by satisfying Routh-Hurwitz criterion through the sign of the eigen value from the characteristic equation and using Lyapunov stability theorem. Finally, numerical simulations have been performed in MATLAB programming language based on the parametric values collected from secondary data as shown in Table 1 which are estimated from research articles, existing literature, annual reports of government, and non-government institutions. To find the parametric values from collective data, different types of qualitative methods like rough order of magnitude (ROM) or conceptual estimation, system observation, least square method have been performed for securing data authentically.

3 Mathematical formulation of the model

In this section, we introduce the background of the model with some basic assumptions and formulate a mathematical model on groundwater level declination.

3.1 Basic assumption and description of the model

To construct a mathematical model on groundwater level, firstly we have to know the biological interactions and background among the state variables (i.e. atmosphere, surface and ground water). For this purpose, we discuss the water flow management with water cycle in the earth and assume some basic assumption as follows:

According to hydrologic cycle, water vaporizes to become the part of the atmosphere from the oceans and the land surface; then water vapor is deported and lifted in the atmosphere until it condenses and precipitates on the surface or the oceans. Through plantation and vegetation precipitated water can be intercepted for overland flow on the ground surface. Water then infiltrates into the ground as subsurface flow and discharge into streams as surface runoff. Finally, the intercepted water and surface runoff returns again to the atmosphere through evaporation directly from the soil and vegetation surface or transpired from plant leaves. In this way hydrologic cycle continues [28].

This study aims to show the effect of the important parameters used in the model like deforestation, pollution, evaporation and frequent pumping of water on groundwater level declination. The interaction of atmosphere, surface and ground water is shown in the Figure 2.

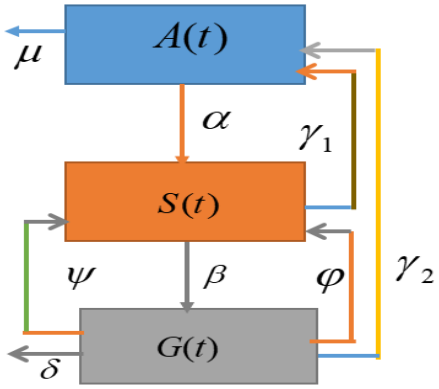


Figure 2: The schematic diagram with biological interactions of groundwater model.

Considering the basic assumptions and schematic diagram in **Figure 2**, we formulate the model with the following system of nonlinear differential equations:

$$\frac{dA}{dt} = \gamma_1 S(t) + \gamma_2 G(t) - \alpha A(t) - \mu A(t) \quad (1)$$

$$\frac{dS}{dt} = \alpha A(t) - \beta G(t)S(t) - \gamma_1 S(t) + \psi G(t)S(t) + \varphi G(t) \quad (2)$$

$$\frac{dG}{dt} = \beta G(t)S(t) - \gamma_2 G(t) - \psi G(t)S(t) - \varphi G(t) - \delta G(t) \quad (3)$$

with initial conditions $A(0) = A_0$, $S(0) = S_0$ and $G(0) = G_0$ and the total amount of water, $N(t) = A(t) + S(t) + G(t)$.

Here, $A(t)$, $S(t)$ and $G(t)$ are the three state variables that represent the levels of atmospheric water, surface water and groundwater respectively at time t.

In the model (1)-(3), we have considered the parameter α as precipitation rate from atmosphere to surface water and μ is the dissipation rate of atmospheric water. The infiltration rate from surface to groundwater be β whereas γ_1 and γ_2 are the evaporation rate from surface and groundwater to atmospheric water respectively. The surface water is polluted at the rate ψ,

moreover δ and φ be the deforestation rate and rate of frequent pumping of water from ground level.

3.2 Analysis of the model

We have to analyze the qualitative behavior of the solutions in the neighborhood of the equilibrium points. For the analysis of the model (1)-(3), a closed set has been considered as

$$\Omega = \left\{ \left(A(t), S(t), G(t) \right) \in \mathbb{R}_+^3 : 0 \le N(t) \le \frac{\eta}{\mu + \delta} \right\}$$ where, η is a constant.

3.2.1 Boundedness of the model

To prove the system mathematically and biologically well posed, the following lemma has to be satisfied.

Lemma 1: The region $\Omega = \left\{ \left(A(t), S(t), G(t) \right) \in \mathbb{R}_+^3 : 0 \le N(t) \le \frac{\eta}{\mu + \delta} \right\}$ is a positively invariant set for the model (1)-(3).

Proof: Since the total amount of water is $N(t)$, then $N(t) = A(t) + S(t) + G(t)$.

The rate of change of total amount of water is

$$\frac{dN}{dt} = \frac{dA}{dt} + \frac{dS}{dt} + \frac{dG}{dt}$$

$$= \gamma_1 S + \gamma_2 G - \alpha A - \mu A + \alpha A - \beta GS - \gamma_1 S + \psi GS + \varphi G + \beta GS - \gamma_2 G - \psi GS - \varphi G - \delta G - \mu G$$

$$= -\mu A - \mu S - \delta G - \mu G$$

$$= -\mu \left(A + S + G \right) - \delta G$$

$$= -\mu N - \delta \left(N - A - S \right)$$

After simplifying, we get the differential inequality as

$$\frac{dN}{dt} \le \eta - \mu N - \delta N$$

$$\Rightarrow \frac{dN}{dt} + \left(\mu + \delta \right) N \le \eta$$ (4)

By solving the differential equation (4), we obtain

$$N(t) \le \frac{\eta}{\mu + \delta} + \left(N_0 - \frac{\eta}{\mu + \delta} \right) e^{-(\mu + \delta)}$$

Taking limit as $t \to \infty$, we get $0 < N(t) \le \frac{\eta}{\mu + \delta}$ (5)

Thus, we conclude that the region Ω is the positively invariant set induced by the model (1)-(3) [33, 36]. Therefore, the model is both mathematically and biologically well-posed in the region Ω.

Hence, the **Lemma 1** is proved.

3.2.2 Positivity of the solution of the model

Since the state equations describes the rate of change of water, so it is required to prove that all the state variables used in the model (1)-(3) are positive.

Lemma 2: If $A(0) > 0, S(0) > 0$ and $G(0) > 0$ then the solutions $A(t), S(t), G(t)$ of the model (1) are all positive [46].

Proof: To prove the Lemma 2, we will use the system of differential equations (1)-(3).

The first equation of the model (1)-(3), can be written as

$$\frac{dA}{dt} \ge -\left(\alpha + \mu \right) A$$

$$\Rightarrow \frac{dA}{dt} + \left(\alpha + \mu \right) A \ge 0$$ (6)

The integrating factor (I.F) of (6) is given by

$$I.F = e^{\int (\alpha + \mu) dt} = e^{(\alpha + \mu) t}$$

Multiplying the I.F on the both side of (6) and integrating we get

$$e^{(\alpha + \mu) t} \frac{dA}{dt} + \left(\alpha + \mu \right) e^{(\alpha + \mu) t} \ge 0$$

$$\Rightarrow A(t) \ge c_1 e^{-(\alpha + \mu) t}$$ (7)

where c_1 is an integrating constant

To find the value of constant c_1 we apply the initial condition at $t = 0, A(0) = A_0$. we get $A_0 \ge c_1$.

Putting the value of c_1 in (7) we obtain

$$A(t) \ge A_0 e^{-(\alpha + \mu) t}$$. (8)

Since $A_0 > 0$ and very larger than α and μ.

Therefore $A(t) > 0$ for all $t \ge 0$.

Similarly, we obtain from the differential equations (2) and (3) that, $S(t) \ge 0$, and $G(t) \ge 0$ for all $t \ge 0$.

Hence the **Lemma 2** is proved.

3.3 Existence of the equilibrium

To find the equilibrium points of the model (1)-(3), we have to solve

$$\frac{dA}{dt} = \frac{dS}{dt} = \frac{dG}{dt} = 0$$

Then the system takes the following form:

$$\gamma_1 S + \gamma_2 G - \alpha A - \mu A = 0 \quad (9)$$

$$\alpha A - \beta GS - \gamma_1 S + \psi GS + \varphi G = 0 \quad (10)$$

$$\beta GS - \gamma_2 G - \psi GS - \varphi G - \delta G = 0 \quad (11)$$

For pollution free and without pumping from groundwater, we consider $G = G_0$ (i.e. initial value of $G(t)$).

Let $\bar{E}_1\left(\bar{A},\bar{S},\bar{G}\right)$ be the pollution free equilibrium point. Using $G=G_0$ in the system (9)-(11), and solving the system of algebraic equation, we get

$$\bar{A}=\frac{\gamma_1\bar{S}+\gamma_2 G_0}{\alpha+\mu} \quad \text{and} \quad \bar{S}=\frac{\gamma_2+\varphi+\delta}{\beta-\psi} \tag{12}$$

Hence, the pollution free equilibrium point is

$$\bar{E}_1\left(\bar{A},\bar{S},\bar{G}\right)=\left(\frac{\gamma_1 S+\gamma_2 G_0}{\alpha+\mu},\frac{\gamma_2+\varphi+\delta}{\beta-\psi},G_0\right) \tag{13}$$

Again, let $E_2^*\left(A^*,S^*,G^*\right)$ be the equilibrium point of the model (1)-(3) when all the activities are performed in the groundwater (i.e. the general equilibrium point), then we obtain by solving the system (9)-(11) as

$$A^*=\frac{\gamma_1 S^*+\gamma_2 G^*}{\alpha+\mu}, \quad S^*=\frac{\gamma_2+\varphi+\delta}{\beta-\psi} \quad \text{and}$$

$$G^*=\frac{S^*\left[\alpha\gamma_1-\gamma_1(\alpha+\mu)\right]}{\left(\beta S^*+\psi S^*+\varphi\right)(\alpha+\mu)-\alpha\gamma_2} \tag{14}$$

Thus, the equilibrium point of the system is

$$E_2^*\left(A^*,S^*,G^*\right)=\begin{pmatrix}\frac{\gamma_1 S^*+\gamma_2 G^*}{\alpha+\mu},\frac{\gamma_2+\varphi+\delta}{\beta-\psi},\\ \frac{S^*\left[\alpha\gamma_1-\gamma_1(\alpha+\mu)\right]}{\left(\beta S^*+\psi S^*+\varphi\right)(\alpha+\mu)-\alpha\gamma_2}\end{pmatrix} \tag{15}$$

where, $A^*=\dfrac{\alpha\delta^2\gamma_1+\alpha\delta\gamma_1\gamma_2+\alpha\delta\gamma_1\varphi}{\alpha(\beta-\psi)(\alpha\delta+\delta\mu+\gamma_2\mu)}$, $S^*=\dfrac{\delta+\gamma_2+\varphi}{\beta-\psi}$

and $G^*=\dfrac{\delta\gamma_1\mu+\gamma_1\gamma_2\mu+\gamma_1\mu\varphi}{\left(\alpha\delta\psi+\delta\mu\psi+\gamma_2\mu\psi-\alpha\beta\delta-\beta\delta\mu-\beta\gamma_2\mu\right)}$

3.4 Stability analysis

In this section, we have to perform stability analysis at the equilibrium point $\bar{E}_1$ and E_2^*.

3.4.1 Stability at $\bar{E}_1\left(\bar{A},\bar{S},\bar{G}\right)$

Theorem 1: The equilibrium point $\bar{E}_1\left(\bar{A},\bar{S},\bar{G}\right)$ of the model (1)-(3) is asymptotically stable if $a_1>0,\ a_3>0$ and $a_1a_2>a_3$, otherwise it is unstable [33, 53].

Proof: In order to prove Theorem 1, we first find Jacobian matrix of the model (1)-(3) [21].

The Jacobian matrix of the system (1)-(3) is given by

$J(A,S,G)=$

$$\begin{pmatrix}-\alpha-\mu & \gamma_1 & \gamma_2\\ \alpha & -\beta G-\gamma_1+\psi G & -\beta S+\psi S+\varphi\\ 0 & \beta G-\psi G & \beta S-\gamma_2-\psi S-\varphi-\delta\end{pmatrix} \tag{16}$$

At the equilibrium point $\bar{E}_1\left(\bar{A},\bar{S},\bar{G}\right)$, the Jacobian matrix (16) takes the following form

$J\left(\bar{A},\bar{S},\bar{G}\right)=$

$$\begin{pmatrix}-\alpha-\mu & \gamma_1 & \gamma_2\\ \alpha & -\beta G_0-\gamma_1+\psi G_0 & -\beta\bar{S}+\psi\bar{S}+\varphi\\ 0 & \beta G_0-\psi G_0 & \beta\bar{S}-\gamma_2-\psi\delta-\varphi-\delta\end{pmatrix} \tag{17}$$

The characteristic equation for the eigen value λ is given as

$$\left|J-\lambda I\right|=0$$

$$\begin{vmatrix}-\alpha-\mu-\lambda & \gamma_1 & \gamma_2\\ \alpha & -\beta G_0-\gamma_1+\psi G_0-\lambda & -\beta\bar{S}+\psi\bar{S}+\varphi\\ 0 & \beta G_0-\psi G_0 & \beta\bar{S}-\gamma_2-\psi\bar{S}-\varphi-\delta-\lambda\end{vmatrix}=0$$

$$\Rightarrow\left(-\alpha-\mu-\lambda\right)\left(-\beta G_0-\gamma_1+\psi G_0-\lambda\right)$$

$$\left(\beta\bar{S}-\gamma_2-\psi\bar{S}-\varphi-\delta-\lambda\right)-\left(-\alpha-\mu-\lambda\right)$$

$$\left(\beta G_0-\psi G_0\right)+\left(-\beta\bar{S}+\psi\bar{S}+\varphi\right)-\alpha\gamma_1$$

$$\left(\beta\bar{S}-\gamma_2-\psi\bar{S}-\varphi-\delta-\lambda\right)-\alpha\gamma_2\left(\beta G_0-\psi G_0\right)=0$$

$$\Rightarrow\left(A_1-\lambda\right)\left(A_2-\lambda\right)\left(A_3-\lambda\right)-\left(A_1-\lambda\right)A_4A_5$$

$$-A_6\left(A_3-\lambda\right)-A_7A_4=0 \tag{18}$$

where, $A_1=-\alpha-\mu,\ A_2=-\beta G_0-\gamma_1+\psi G_0,$

$$A_3=\beta\bar{S}-\gamma_2-\psi\bar{S}-\varphi-\delta,\ A_4=\beta G_0-\psi G_0,$$

$A_5=-\beta\bar{S}+\psi\bar{S}+\varphi$, and $A_6=\alpha\gamma_1,\ A_7=\alpha\gamma_2$

After simplifying the equation (18), we obtain

$$\lambda^3+\lambda^2\left\{-\left(A_1+A_2+A_3\right)\right\}+\lambda\left\{A_1\left(A_2+A_3\right)A_2+A_3\right\}$$

$$-\lambda A_4A_5-\lambda A_6+A_1A_4A_5-A_1A_2A_3+A_3A_6-A_4A_7=0$$

$$\Rightarrow\lambda^3+a_1\lambda^2+a_2\lambda+a_3=0 \tag{19}$$

where,

$a_1=\left\{-\left(A_1+A_2+A_3\right)\right\},\ a_2=\left\{A_1\left(A_2+A_3\right)A_2A_3\right\}$

$-A_4A_5-A_6$ and $a_3=A_1A_4A_5-A_1A_2A_3+A_3A_6-A_4A_7$.

Here, $a_1=\left\{-\left(A_1+A_2+A_3\right)\right\}$

$$=\alpha+\mu+\lambda+\beta G_0+\gamma_1-\psi G_0-\beta\bar{S}+\gamma_2+\psi\bar{S}+\varphi+\delta>0$$

similarly, $a_3>0$ and $a_1a_2>a_3$.

From the Routh-Hurwitz criterion [18], we know that all the eigen values of (19) have negative real roots if and only if $a_1>0,\ a_3>0$ and $a_1a_2>a_3$.

Hence, the equilibrium point $\bar{E}_1\left(\bar{A},\bar{S},\bar{G}\right)$ of the model (1)-(3) is asymptotically stable for $a_1>0,\ a_3>0$ and $a_1a_2>a_3$.

3.4.2 Stability at $E_2^*(A^*, S^*, G^*)$

Theorem 2: The equilibrium point $E_2^*(A^*, S^*, G^*)$ of the model (1)-(3) is asymptotically stable if $b_1 > 0,\ b_3 > 0$ and $b_1 b_2 > b_3$, otherwise it is unstable [46, 53].

Proof: The Jacobian matrix of the system (1)-(3) is given by

$$J(A,S,G) = \begin{pmatrix} -\alpha-\mu & \gamma_1 & \gamma_2 \\ \alpha & -\beta G-\gamma_1+\psi G & -\beta S+\psi S+\varphi \\ 0 & \beta G-\psi G & \beta S-\gamma_2-\psi S-\varphi-\delta \end{pmatrix}$$

At the equilibrium point $E_2^*(A^*, S^*, G^*)$, the Jacobian matrix takes the following form

$$J(A^*, S^*, G^*) = \begin{pmatrix} -\alpha-\mu & \gamma_1 & \gamma_2 \\ \alpha & -\beta G^*-\gamma_1+\psi G^* & -\beta S^*+\psi S^*+\varphi \\ 0 & \beta G-\psi G & \beta S^*-\gamma_2-\psi S^*-\varphi-\delta \end{pmatrix}$$

The characteristic equation for the eigen value λ is given as

$$|J-\lambda I| = 0$$

$$\begin{vmatrix} -\alpha-\mu-\lambda & \gamma_1 & \gamma_2 \\ \alpha & -\beta G^*-\gamma_1+\psi G^*-\lambda & -\beta S^*+\psi S^*+\varphi \\ 0 & \beta G-\psi G & \beta S^*-\gamma_2-\psi S^*-\varphi-\delta-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (-\alpha-\mu-\lambda)(-\beta G^*-\gamma_1+\psi G^*-\lambda)\begin{pmatrix} \beta S^*-\gamma_2-\psi S^* \\ -\varphi-\delta-\lambda \end{pmatrix}$$

$$-(-\alpha-\mu-\lambda)(\beta G-\psi G)\ (-\beta S^*+\psi S^*+\varphi)-\alpha\gamma_1$$

$$(\beta S^*-\gamma_2-\psi S^*-\varphi-\delta-\lambda)-\alpha\gamma_2(\beta G-\psi G)=0$$

$$\Rightarrow (B_1-\lambda)(B_2-\lambda)(B_3-\lambda)-(B_1-\lambda)B_4B_5 - B_6(B_3-\lambda)-B_7B_4=0 \quad (20)$$

where, $B_1 = -\alpha-\mu,\ B_2 = -\beta G^*-\gamma_1+\psi G^*,$

$B_3 = \beta S^*-\gamma_2-\psi S^*-\varphi-\delta,\ B_4 = \beta G-\psi G,$

$B_5 = -\beta S^*+\psi S^*+\varphi,\ B_6 = \alpha\gamma_1,$ and $B_7 = \alpha\gamma_2$

After simplifying the equation (20), we get

$$\lambda^3+\lambda^2\{-(B_1+B_2+B_3)\}+\lambda\{B_1(B_2+B_3)B_2B_3-B_4B_5-B_6\} + B_1B_4B_5-B_1B_2B_3+B_3B_6-B_4B_7=0$$

$$\Rightarrow \lambda^3+b_1\lambda^2+b_2\lambda+b_3=0 \quad (21)$$

where, $b_1 = \{-(B_1+B_2+B_3)\},$

$b_2 = \{B_1(B_2+B_3)B_2B_3-B_4B_5-B_6\}$

and $b_3 = B_1B_4B_5-B_1B_2B_3+B_3B_6-B_4B_7$

From the Routh-Hurwitz criterion, we know that all the eigen values of (21) have negative real roots if and only if $b_1 > 0,\ b_3 > 0$ and $b_1 b_2 > b_3$.

By substituting the value of B_1, B_2, B_3, B_4 and B_5, we obtain $b_1 > 0,\ b_3 > 0$ and $b_1 b_2 > b_3$.

Hence, the equilibrium point $E^*(A^*, S^*, G^*)$ of the model (1)-(3) is asymptotically stable if $b_1 > 0,\ b_3 > 0$ and $b_1 b_2 > b_3$, otherwise it is unstable.

3.4.3 Global Stability at the Equilibrium Point $E_2^*(A^*, S^*, G^*)$

In this section, we use the Lyapunov direct method [14, 16, 48] to establish sufficient conditions for the global asymptotic stability of the equilibrium point E_2^* in int $\mathbb{R}^3_+$.

Theorem 3: The equilibrium point E_2^* of the of the model (1)-(3) is globally asymptotically stable if $\frac{\alpha\varphi A^*}{\gamma_1 S^*} > \frac{\psi\delta}{\gamma_2 G^*}$ in the interior of the feasible region, otherwise it is unstable.

Proof: Theorem 3 can be proved based on the Lyapunov stability theorem. For that purpose, we **c**onsider the following nonlinear Lyapunov function,

$$V = A - A^* - A^*\ln\frac{A}{A^*} + S - S^* - S^*\ln\frac{S}{S^*} + G - G^* - G^*\ln\frac{G}{G^*} \quad (22)$$

Then V is C^1 on the interior of, $\mathbb{R}^3_+$ $E_2^*(A^*, S^*, G^*)$ is the equilibrium point.

The derivative of (22) along the solution curves of the model (1)-(3) is given by the expression:

$$\dot{V} = \dot{A} - \frac{A^*\dot{A}}{A} + \dot{S} - \frac{S^*\dot{S}}{S} + \dot{G} - \frac{G^*\dot{G}}{G}$$

$$= \gamma_1 S + \gamma_2 G - \alpha A - \mu A - \frac{A^*}{A}(\gamma_1 S + \gamma_2 G - \alpha A - \mu A) +$$

$$\alpha A - \beta GS - \gamma_1 S + \psi GS + \varphi G - \frac{S^*}{S}\begin{pmatrix} \alpha A - \beta GS - \gamma_1 S \\ +\psi GS + \varphi G \end{pmatrix}$$

$$+ \beta GS - \gamma_2 G - \psi GS - \varphi G - \delta G - \frac{G^*}{G}\begin{pmatrix} \beta GS - \gamma_2 G \\ -\psi GS - \varphi G - \delta G \end{pmatrix}$$

$$= \gamma_1 S\left(1-\frac{A^*}{A}\right) + \gamma_2 G\left(1-\frac{A^*}{A}\right) - (\alpha+\mu)A\left(1-\frac{A^*}{A}\right) +$$

$$\alpha A\left(1-\frac{S^*}{S}\right) - (\beta-\psi)GS\left(1-\frac{S^*}{S}\right)$$

$$-\gamma_1 S\left(1-\frac{S^*}{S}\right) + \varphi G\left(1-\frac{S^*}{S}\right) - (\gamma_2+\varphi+\delta)$$

$$G\left(1-\frac{G^*}{G}\right) + (\beta-\psi)GS\left(1-\frac{G^*}{G}\right) \quad (23)$$

At the equilibrium point $E_2^*(A^*, S^*, G^*)$, we have

$$\alpha+\mu=\frac{\gamma_1 S^*+\gamma_2 G^*}{A^*} \tag{24}$$

$$\beta-\psi=\frac{\alpha A^*-\gamma_1 S^*+\varphi G^*}{G^* S^*} \tag{25}$$

$$\gamma_2+\varphi+\delta=\frac{\beta G^* S^*-\psi G^* S^*}{G^*} \tag{26}$$

using the equations (24)-(26) in (23), we obtain

$$\dot{V}=\gamma_1 S\left(1-\frac{A^*}{A}\right)+\gamma_2 G\left(1-\frac{A^*}{A}\right)-\frac{\gamma_1 S^*+\gamma_2 G^*}{A^*}$$
$$A\left(1-\frac{A^*}{A}\right)+\alpha A\left(1-\frac{S^*}{S}\right)-\frac{\alpha A^*-\gamma_1 S^*+\varphi G^*}{G^* S^*}$$
$$GS\left(1-\frac{S^*}{S}\right)-\gamma_1 S\left(1-\frac{S^*}{S}\right)+\varphi G\left(1-\frac{S^*}{S}\right)-$$
$$\frac{\beta G^* S^*-\psi G^* S^*}{G^*} G\left(1-\frac{G^*}{G}\right)+\frac{\alpha A^*-\gamma_1 S^*+\varphi G^*}{G^* S^*}$$
$$GS\left(1-\frac{G^*}{G}\right)$$
$$=\gamma_1 S-\frac{\gamma_1 S A^*}{A}+\gamma_2 G-\frac{\gamma_2 G A^*}{A}-\frac{\left(\gamma_1 S^*+\gamma_2 G^*\right)A}{A^*}$$
$$+\left(\gamma_1 S^*+\gamma_2 G^*\right)+\alpha A-\frac{\alpha A S^*}{S}-\frac{\left(\alpha A^*-\gamma_1 S^*+\varphi G^*\right)GS}{G^* S^*}$$
$$+\frac{\left(\alpha A^*-\gamma_1 S^*+\varphi G^*\right)G}{G^*}-\gamma_1 S^*+\varphi G\frac{S^*}{S}-$$
$$\frac{\beta G^* S^*-\psi G^* S^*}{G^*} G+\beta G^* S^*-\psi G^* S^*+$$
$$\frac{\alpha A^*-\gamma_1 S^*+\varphi G^*}{G^* S^*} GS-\frac{\left(\alpha A^*-\gamma_1 S^*+\varphi G^*\right)S}{S^*}$$
$$=\gamma_1\left(S-S^*\right)-\frac{\gamma_1 S A^{*2}-\left(\gamma_1 S^*+\gamma_2 G^*\right)A^2}{AA^*}-\frac{\gamma_2 G A^*}{A}+$$
$$\gamma_1 S^*+\alpha A-\frac{\alpha A S^*}{S}-\frac{\left(\alpha A^*-\gamma_1 S^*+\varphi G^*\right)GS}{G^* S^*}$$
$$+\frac{\left(\alpha A^*-\gamma_1 S^*+\varphi G^*\right)G-\beta G G^* S^*-\psi G G^* S^*}{G^*}$$
$$+\gamma_2\left(G-G^*\right)+\beta G^* S^*-\psi G^* S^*+\frac{\alpha A^*-\gamma_1 S^*+\varphi G^*}{G^* S^*} GS$$
$$-\frac{\left(\alpha A^*-\gamma_1 S^*+\varphi G^*\right)S^2+\varphi GS^{*2}}{SS^*}$$
$$=\gamma_1\left(S-S^*\right)-\frac{\gamma_2 G A^*}{A}+\gamma_1 S^*-\frac{\left(\alpha A^*-\gamma_1 S^*+\varphi G^*\right)GS+}{G^* S^*}$$
$$\left(\left(\alpha A^*-\gamma_1 S^*+\varphi G^*\right)S^* G-\beta G G^* S^{*2}-\psi G G^* S^{*2}\right)\frac{1}{G^* S^{*2}}+$$
$$\gamma_2\left(G-G^*\right)+\beta G^* S^*-\psi G^* S^*-\frac{\alpha\varphi\gamma_1 A^* S^* S^2 G S^{*2}}{SS^*}\frac{AG^* S}{S^* G A^*}$$
$$-\frac{\gamma_1 S A^{*2}-\left(\gamma_1 S^*+\gamma_2 G^*\right)A^2}{AA^*}+\alpha A$$
$$-\frac{\alpha A S^*}{S}+\frac{\alpha\varphi\gamma_1 A^* S^* S^2 G S^{*2}}{SS^*}\left(4-\frac{A^*}{A}-\frac{S^* G}{G^* S}\right)$$
$$=\gamma_1\left(2-\frac{S^*}{S}-\frac{S}{S^*}\right)-\frac{\gamma_2 G A^*}{A}-\gamma_1 S^*-\alpha A\frac{\left(A^*+\gamma_1 S^*+\varphi G^*\right)}{G^* S^*}$$
$$\left(\frac{\alpha\varphi A^*}{\gamma_1 S^*}-\frac{\psi\delta}{\gamma_2 G^*}\right)-\frac{\gamma_1 S A^{*2}-\left(\gamma_1 S^*+\gamma_2 G^*\right)A^2}{AA^*}$$
$$+\gamma_2\left(G-G^*\right)-\beta G^* S^{**}+\frac{\alpha\varphi\gamma_1 A^* S^* S^2 G S^{*2}}{SS^*}$$
$$\left(4-\frac{A^*}{A}-\frac{S^* G}{G^* S}-\frac{AG^* S}{S^* G A^*}\right)$$

Since the arithmetic mean is greater than or equal to the geometric mean and geometric mean is greater than or equal to the harmonic mean, it follows that

$$2-\frac{S^*}{S}-\frac{S}{S^*}\le 0$$
$$4-\frac{A^*}{A}-\frac{S^* G}{G^* S}-\frac{AG^* S}{S^* G A^*}\le 0$$

Further, since all the model parameters are nonnegative, it follows that $\dot{V}\le 0$ for $\frac{\alpha\varphi A^*}{\gamma_1 S^*}>\frac{\psi\delta}{\gamma_2 G^*}$ with $\dot{V}=0$ if and only if $A=A^*, S=S^*$ and $G=G^*$ holds.

The largest compact invariant set in $\left\{(A,S,G)\in\mathbb{R}_+^3 : dV/dt=0\right\}$ is the singleton $\left\{E_2^*\right\}$, where E_2^* is the equilibrium point of the model. By using LaSalle's invariance principle [37, 54] then implies that $E_2^*\left(A^*,S^*,G^*\right)$ is globally asymptotically stable in the interior of $\mathbb{R}_+^3$ if $\frac{\alpha\varphi A^*}{\gamma_1 S^*}>\frac{\psi\delta}{\gamma_2 G^*}$, otherwise it is unstable.

4 Numerical analysis

In this section, we have analyzed the model with graphical analysis and sensitivity analysis.

4.1 Numerical simulations

We have solved the model numerically based on the respective parameters present in the system of equations

(1)-(3) to investigate the dynamical behavior of the model. The simulations are carried out by ode45 solver using MATLAB programming language. We use a set of suitable parameter values. The description of all the parameters with the estimated values used in the simulation is presented in Table 1. We have considered the initial condition $A_0 = 50\times10^9$, $S_0 = 110\times10^9$ and $G_0 = 60\times10^9$.

Firstly, we solve the model (1)-(3) considering the initial values and all other parameters that are estimated from [6, 56]. Also, we have performed the numerical simulations for time interval $t \in [0, 20]$ for 20 years. The value of the model parameters is given in Table 1.

Table 1: Values and explanation of parameters

Descriptions	Symbols	Values
Precipitation rate	α	0.5
Dissipation rate	μ	0.01
Infiltration rate	β	0.3005
Deforestation rate	δ	0.06
Pumping rate of water	φ	0.02
Pollution rate	ψ	0.300
Evaporation rate from surface water	γ_1	0.09
Evaporation rate from ground water level	γ_2	0.028

Our object is to study the effects of deforestation rate (δ), pumping rate of water (φ), pollution rate (ψ) and evaporation rate (γ_2) due to global warming on groundwater level as well as in surface water and atmospheric water. We have selected these parameters because they have a large impact on groundwater level declination. So, if it is possible to minimize the deforestation rate, pumping rate of water, pollution rate then the declination of groundwater level will be controlled. Considering these parameters into account, we have run the program for the state variables, atmospheric water $A(t)$, surface water $S(t)$ and groundwater level $G(t)$). The result of simulation of the combined class is presented in Figure 3.

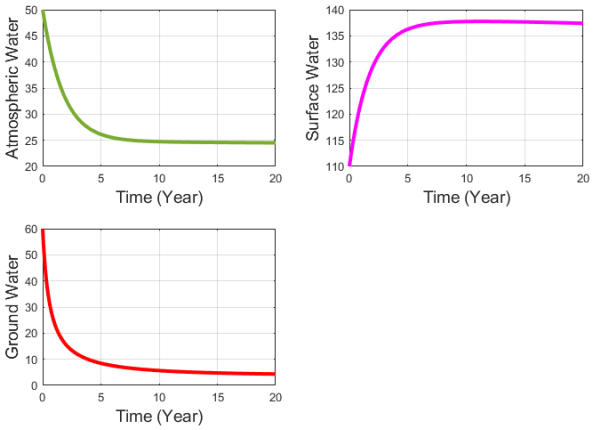


Figure 3: Numerical simulation for Groundwater model, with time (20 years).

From **Figure 3**, we observe that groundwater level is decreasing day by day due to natural phenomena, over demand, natural disaster and so on whereas surface water primarily increases but after sometimes it shows the stable situation. At the same time, atmospheric water is also decreasing day by day. In everyday life we would face the problems with freshwater shortage is sure to cause problems in many aspects of our lives. The activities that lead to groundwater declination come mostly from human being, but a portion of it also comes from changes in our climate (for examples, deforestation, drought, global warming etc.) and can speed up the process. So, the human population is the main responsible for water contamination.

Again, we run the program for two state variables, surface water and groundwater level keeping all the values of the parameters same as before. The result obtained in this case is represented in Figure 4.

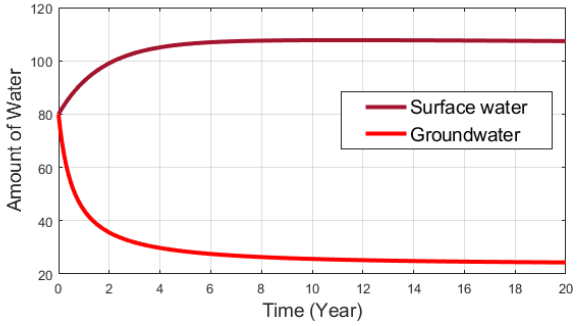


Figure 4: Dynamics of surface water and groundwater where groundwater decrease significantly as a result surface water is increasing day by day.

Figure 4 shows the state trajectories of the two compartments such as surface water and groundwater in the absence
of any control measures. We have observed that when no control measure is employed, groundwater level decreases from initial state. At the same time, surface water gradually increases from the initial position and after some years it reaches to the peak level. It is occurred due to evaporation of water from the ocean and land surfaces, because this water is temporarily considered as vapor in the atmosphere, and falls back to earth's surface as precipitation, then surface water is formed by the residue of precipitation and melted snow and then it will be stable if no others factor (for examples, drought and low streamflow or over streamflow) as employed in the atmosphere.

Now, we run the program keeping all other values of the parameters same as before. We have some small changed to pollution rate to show the effect of pollution rate on groundwater level declination, we consider the same initial values of the three state variables. The result obtained in this case is given in Figure 5.

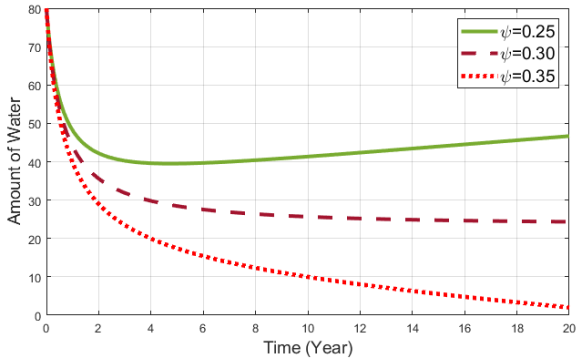


Figure 5: Variation of groundwater for different values of pollution rate where groundwater level is significantly decreased due to increase of pollution rate.

In Figure 5, we see the variation of groundwater level for different values of pollution rate (ψ) with time. It is easy to see that falling of groundwater level is increased surprisingly due to increase of pollution rate ψ. In other words, falling of groundwater level decreases for the lower value of pollution rate ψ. Groundwater protection specialists inspect and analyze that human are the main cause of water contamination, which is provoked in many ways: by the dumping of industrial waste; due to temperature rise, that cause the variation of water by reducing the oxygen in its composition; which causes of silt and bacteria to appear under the soil and therefore pollute groundwater and is declined significantly in day by day. Thus, groundwater level decreases as the increase of pollution rate and groundwater level increases as the decrease of pollution rate.

Again, we solve the model numerically for the case of groundwater to show the change in the groundwater level due to frequent pumping of water. Taking the values of all parameters into account as in Table 1 and considering the same initial values of the state variables as we did before. The result in this case is presented in Figure 6.

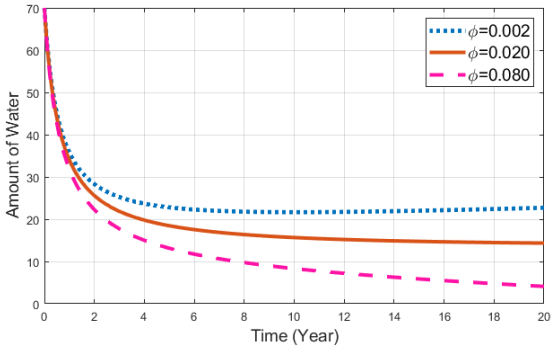


Figure 6: Variation of groundwater for different values of pumping rate were falling of groundwater level is significantly increased due to increase of pumping of water.

Figure 6 represents the variation of groundwater level for different values of pumping rate φ with time while pollution rate and other effects are fixed. We observe that falling of groundwater level is increased tremendously due to increase of pumping rate φ. The more we extract groundwater from below the earth's surface, then we have to go in the downward of porous soil and rock in order to get more fresh water. But it is the matter of very anxious that when we have emitted water from deeper within the Earth, we find that there is less water available. Consequently, we will have to use more resources to develop alternative methods to reach further into the groundwater level.

Next, we have solved the model numerically to show the change in the groundwater level due to evaporation rate for the parameters in Table. The result in this case is shown in Figure 7.

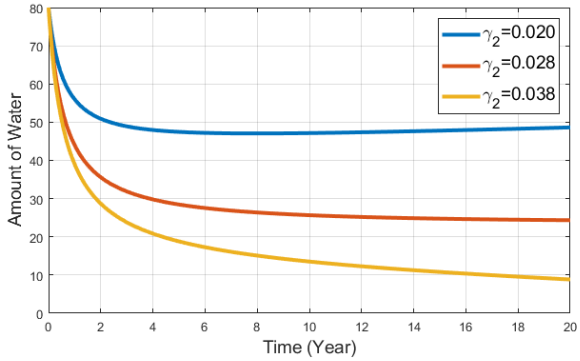


Figure 7: Variation of groundwater for different values of evaporation rate where falling of groundwater level is extensively increased due to global warming.

In **Figure 7**, one can see the variation of groundwater level for different values of evaporation rate (γ_2) over the time while deforestation rate, pollution rate and other effects are constant. It is easy to understand that falling of groundwater level is increased significantly due to increase of evaporation rate γ_2. That is, falling of

groundwater level decreases for the lower value of evaporation rate γ_2 and falling of groundwater level increases for the higher value of evaporation rate γ_2. Global warming can lead to longer periods of droughts, which directly affects availability and dependency on groundwater. Moreover, due to long periods of droughts there is a higher risk of declination on aquifers, especially in case of small and shallow aquifers as a result groundwater level is declined.

Lastly, we run the program keeping all other values of the parameters same as before and change in the deforestation rate to show the effect of deforestation on groundwater level. The result in this case is presented in Figure 8.

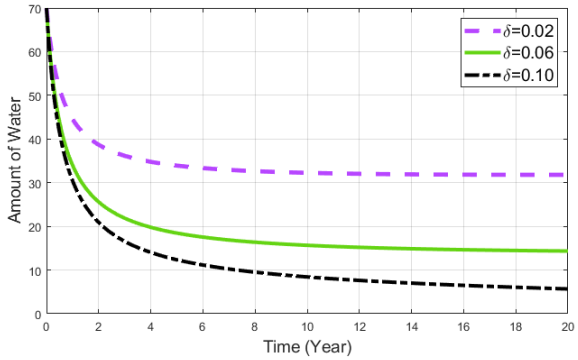


Figure 8: Groundwater level is tremendously decreased due to increase of deforestation rate from $\delta = 0.02$ to $\delta = 0.10$

Figure 8 represents the variation of groundwater level for different values of deforestation rate δ with time while evaporation rate, pumping rate and other effects are constant. We observe that groundwater level is decreased extensively due to increase of deforestation rate δ. In other words, falling of groundwater level decreases for the lower value of deforestation rate δ and falling of groundwater level increases for the higher value of deforestation rate δ. The diminution of trees and other plants can cause global warming, soil erosion, desertification, fewer crops, flooding, increased greenhouse gases in the environment. Thus, it has a significant impact on groundwater. The roots from vegetation and plantation are used to filter the various contamination from groundwater. Thus, by increasing deforestation i.e. by reducing plantation and vegetation, groundwater is faced a great threat in our country as well as all over the world.

4.2 Sensitivity analysis

To evaluate the proper management of groundwater and realizing the impact of climate change or global environmental crisis on groundwater level declination in Bangladesh as well as all over the world, it is mandatory to know the relative importance of each input parameter.

In Sensitivity analysis, we evaluate the uncertainty in the output of a mathematical simulation and which parameters and interactions have the most significant impact on the dynamical behavior of the system. It plays a significant role to the experimental designs and data assimilation of nonlinear compartmental model [8].

There are many ways to perform the sensitivity analysis. One of the ways is normalized forward sensitivity index of a variable with respect to a parameter. It is defined as the ratio of the relative change in the variable to the comparative change in the parameter [17]. When the variable is a differentiable function of the parameter, then the sensitivity index may be represented with partial derivatives.

If the normalized forward sensitivity index of G^* is differentiable with respect to a given parameter P, then it is denoted as $\gamma_P^{G^*}$ and defined as

$$\gamma_P^{G^*} = \frac{\partial G^*}{\partial P}\frac{P}{G^*}$$

For example, the sensitivity index of G^* with respect to β is

$$\gamma_\beta^{G^*} = \frac{\partial G^*}{\partial \beta}\frac{\beta}{G^*} = 0.2763$$

The values of the sensitivity indices for the corresponding parameters that we have used in the model are presented in Table 2.

Table 2: Sensitivity indices of different parameter values given in Table 1

Parameters	Values	Sensitivity index
β	0.3005	0.2763
δ	0.06	-0.7203
φ	0.02	-0.8462
ψ	0.300	-0.4860
γ_2	0.028	-0.0312

From Table 2, we notice that the most sensitive parameters to the groundwater level declination of the model (1)-(3) are infiltration rate (β), deforestation rate (δ), pumping rate of water (φ), and pollution rate (ψ). In practically, the sensitivity analysis predicts that an increase of the infiltration rate β will increase the groundwater level by 27.63%. Conversely, an increase of the value of frequently pumping rate of water φ will

decrease groundwater level by 84.62%. These phenomena also can be observed from the **Figures 9-12**.

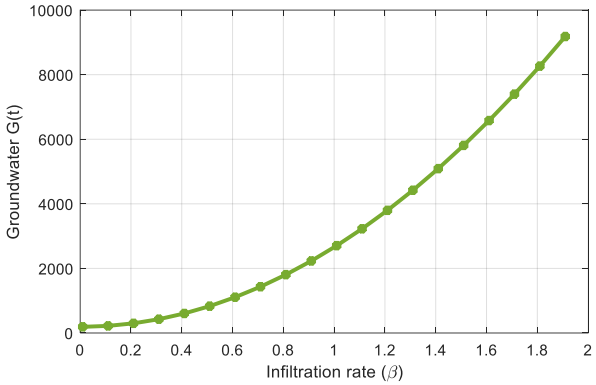


Figure 9: Sensitivity graph of groundwater level on the basis of infiltration rate (β).

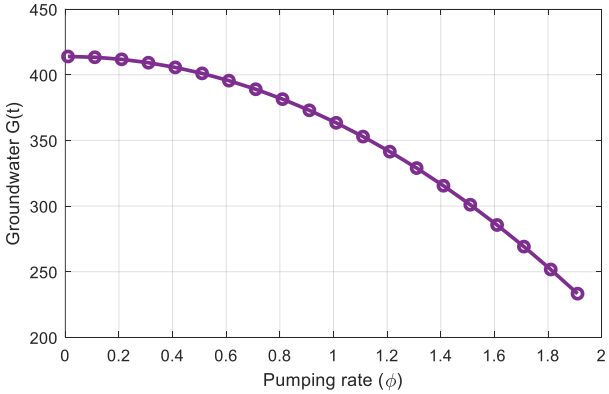


Figure 10. Sensitivity graph of groundwater level on the basis of pumping rate (φ)

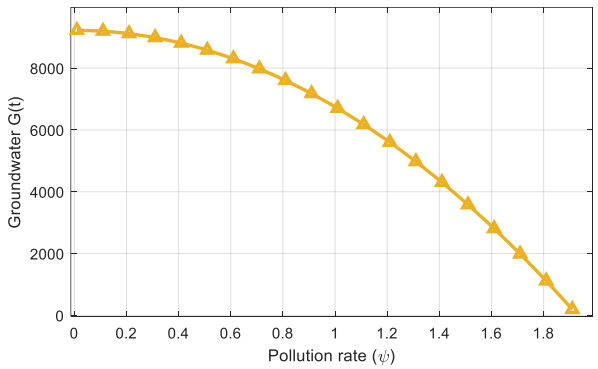


Figure 11: Sensitivity graph of groundwater level on the basis of pollution rate (ψ).

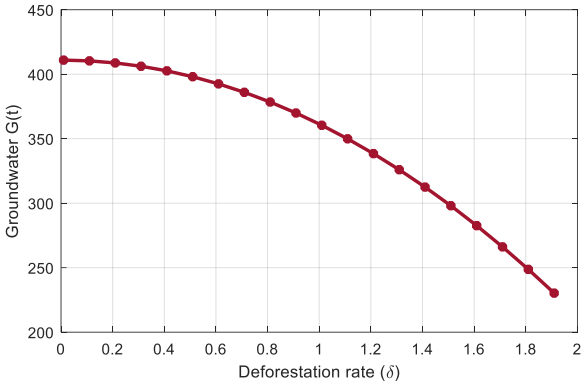


Figure 12: Sensitivity graph of groundwater level on the basis of deforestation rate (δ).

5 Phase portrait analysis

We now investigate the dynamical behavior of the model (1)-(3) by plotting phase-plane diagram based on the respective parameters present in the model. In these phase-planes, small arrows in Figures 9-11 show the direction field; the red dot represents the equilibrium point and dashed lines are the nullclines.

In this model, we have computed three graphs (see **Figures 13-15**) to show the stability of the model near equilibrium points. It is clear that the pollution free equilibrium point is asymptotically stable and the nature of equilibrium is node for the parameter values $\alpha = .05,\ \beta = 0.3005,\ \varphi = 0.02, \psi = 0.300$ and $\alpha = .05,\ \beta = 0.3005,\ \varphi = 0.06,\ \psi = 0.400$ but when the value of pollution rate and frequent pumping rate is continuously increased, then the solution curve of the groundwater level is extensively decreased and the nature of equilibrium is unstable saddle point.

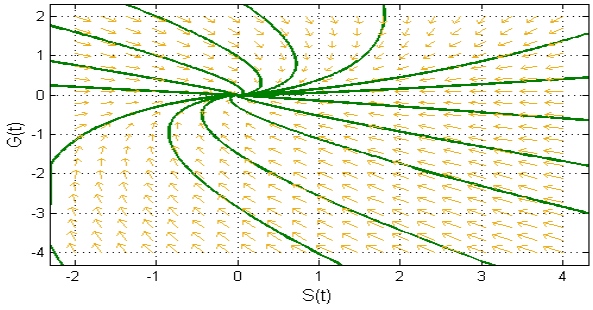


Figure 13: Phase plane for the model (1)-(3) with the parameter values $\alpha = .05,\ \beta = 0.3005,\ \varphi = 0.02$ and $\psi = 0.300$

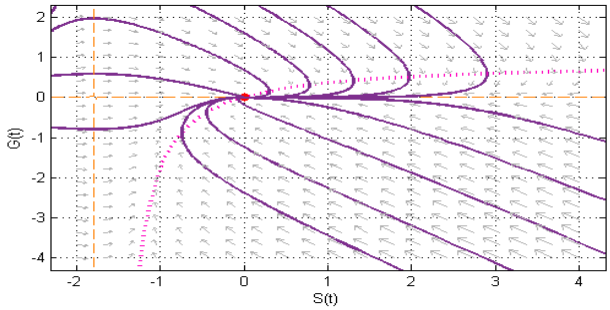


Figure 14: Phase plane and nullclines for the model (1)-(3) with the parameter values
$\alpha = .05,\ \beta = 0.3005,\ \varphi = 0.06$ and $\psi = 0.400$

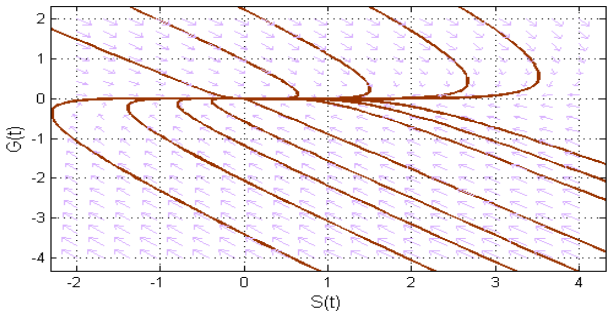


Figure 15: Phase plane for the model (1)-(3) with the parameter values
$\alpha = .05,\ \beta = 0.3005,\ \varphi = 0.500$ and $\psi = 0.300$

6 Conclusions

In this study, a mathematical model on groundwater level declination is presented with qualitative and quantitative analysis. We investigate the dynamical behavior of the model (1)-(3) by stability analysis and plotting phase-plane diagram based on the respective parameters present in the model. From the numerical simulations, it is clear from Figures 5 and 6 that, the groundwater level is significantly decreased due to increase of pollution rate while the falling of groundwater level is tremendously increased due to increase of over pumping through shallow tubewells, deep tubewells and low lift pumps. Because excessive pumping is the cause of lower groundwater table. Again, Groundwater and surface water are connected to each other. When groundwater is used superfluously, then lakes, streams and rivers that are connected to subsurface water can also have their supply diminished. We also observe from Figures 7 and 8 that, Groundwater level is tremendously decreased due to increase of deforestation rate, where falling of groundwater level is extensively increased for higher rate of evaporation that is caused of global warming.

Lastly, we demonstrate that, this model gives a latest picture of groundwater level management in Bangladesh as well as all over the world that we have to very careful and conscious to use water specifically groundwater and making sure that we have to reduce the misusage of it in our daily life. The proposed model can be helped for the researchers and planners who are associated with the research of groundwater level. It also may be helpful for the government to make and take decision regarding the prevention of groundwater level declination as well as may be increase the public awareness in case of using groundwater.

Acknowledgements:

The first author is greatly acknowledged the partial support of national science and technology (NST) fellowship, bearing ID: MSc-181205, Serial: 708, Merit: 267, No: 39.00.0000.012.002.04.19-06, Session: 2019-2020, which is provided by the Ministry of Science and Technology, Government of the People's Republic of Bangladesh.

Conflicts of Interest:

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

References:

[1] Adams TM, Tang AYS, Wiegand N. Spatial Data Models for Managing Subsurface Data. *Journal of Computation in Civil Engineering*, 7(3):260–277, 1993. https://doi.org/10.1061/(asce)08873801(1993)7:3(260)

[2] Anandalekshmi A, Panicker ST, Adarsh S et al. Modeling the Concurrent Impact of Extreme Rainfall and Reservoir Storage on Kerala Floods 2018: a Copula Approach. *Model. Earth Syst Environ*, 5(4):1283–1296, 2019. https://doi.org/10.1007/s40808-019-00635-6.

[3] Attaullah, Sohaib M. Mathematical Modeling and Numerical Simulation of HIV Infection Model. *Results in Applied Mathematics*, 7(100118): 1-11, 2020. https://doi.org/10.1016/j.rinam.2020.100118.

[4] Baalousha HM, Barth N, Ramasomanana FH et al. Groundwater Recharge Estimation and its Spatial Distribution in Arid Regions Using GIS: a Case Study from Qatar Karst Aquifer. *Model Earth Syst Environ*, 4(4): 1319–1329, 2018. https://doi.org/10.1007/s40808-018-0503-4.

[5] Bagherzadeh A, Hoseini AV, Totmaj LH. The Effects of Climate Change on Normalized Difference Vegetation Index (NDVI) in the Northeast of Iran. *Model Earth Syst Environ*, 6(2): 671–683, 2020. https://doi.org/10.1007/s40808-020-00724-x

[6] Bangladesh Water Development Board. https://www.hydrology.bwdb.gov.bd/index.php?pagetitle=ground_water_daily_data; [accessed 13 September 2020].

[7] Bear J, Verruijt A. Modeling Groundwater Flow and Pollution: Theory and Applications of Transport in Porous Media. *D Reidel Publishing Company*,

Holland, 2020.

[8] Biswas MHA, Islam MA, Akter S, Mondal S, Khatun MS, Samad SA, Paul AK and Khatun M R. Modelling the Effect of Self-Immunity and the Impacts of Asymptomatic and Symptomatic Individuals on COVID-19 Outbreak, *Computer Modeling in Engineering & Sciences*, 125(3): 1033–1060, 2020. https://doi.org/10.32604/cmes.2020.012792.

[9] Biswas MHA, Hossain MR, Mondal MK. Mathematical Modeling Applied to Sustainable Management of Marine Resources. *Procedia Engineering*, 194, 337-344, 2017. https://doi.org/10.1016/j.proeng.2017.08.154

[10] Biswas MHA, Rahman T, Haque N. Modeling the potential impacts of global climate change in Bangladesh: an optimal control approach. *J Fundam Appl Sci*, 8(1), 1–19, 2016. https://doi.org/10.4314/jfas.v8i1.1

[11] Brammer H, Ravenscroft P. Arsenic in Groundwater: A Threat to Sustainable Agriculture in South and South-east Asia. *Environment International*, 35(3), 647- 654, 2009. https://doi.org/10.1016/j.envint.2008.10.004

[12] Brenner S, Coxon G, Nicholas JKH, Freer J, Hartmann A. Process-based Modelling to Evaluate Simulated Groundwater Levels and Frequencies in a Chalk Catchment in South-western England. *Natural Hazards and Earth System Sciences*, 18, 445-461, 2018. https://doi.org/10.5194/nhess-18-445-2018

[13] Brouyere S, Carabin G, Dassargues A. Climate Change Impacts on Groundwater Resources: ModelledDeficits in a Chalky Aquifer, Geer Basin, Belgium. *Hydrogeology Journal*, 12, 123-134, 2004. https://doi.org/10.1007/s10040-003-0293-1

[14] Buonomo B, Vargas-De-León C. Global stability for an HIV-1 Infection Model Including an Eclipse Stage of Infected Cells. *Journal of Mathematical Analysis and Applications*, 385, 709-720, 2012. https://doi.org/10.1016/j.jmaa.2011.07.006

[15] Burges DN, Borrie MS. Modeling with Differential Equations. *Ellis Horwood Limited*, New York, USA, 1981.

[16] Cheng Y, Wang J, Yang X. On the Global Stability of a Generalized Cholera Epidemiological Model. *Journal of Biological Dynamics*, 6(2): 1088-1104, 2012. https://doi.org/10.1080/17513758.2012.728635.

[17] Chitnis, N., Hyman, J. M., Cushing, J. M. Determining Important Parameters in the Spread of Malaria through the Sensitivity Analysis of a Mathematical Model. Bulletin of Mathematical Biology, 70: 1272-1296, 2020. https://doi.org/10.1007/s11538-008-9299-0.

[18] De Jesus EX, Kaufman C. Routh-Hurwitz criterion in the examination of eigenvalues of a system of nonlinear ordinary differential equations. *Phys Rev A*, 35(12): 5288, 1987. https://doi.org/10.1103/physreva.35.5288

[19] Depletion of groundwater and Its Consequences. https://www.daily-sun.com/printversion/details/271861/Groundwater-Depletion-and-Its-Consequences; [accessed 11 September 2020].

[20] Dym CL. Principles of Mathematical Modeling. Second Edition, Elsevier Academic Press, California, USA, 2004.

[21] Fabio DN, Abba SI, Pham BQ, Islam ARMT, Talukdar S, Francesco G. Groundwater level forecasting in Northern Bangladesh using nonlinear autoregressive exogenous (NARX) and extreme learning machine (ELM) neural networks. *Arabian Journal of Geosciences*, 15(7): 1-20, 2022. https://doi.org/10.1007/s12517-022-09906-6

[22] Freshwater used in agriculture. https://blogs.worldbank.org/opendata/chart-globally-freshwater-agriculture; [accessed 13 September 2020].

[23] Gantmacher FR. The Theory of Matrices. Vol. 1, AMS Chelsea Publishing, Providence, RI, USA, 1988.

[24] Golla V. Delineation of groundwater potential zones in Sathyavedu area, Chittoor District (Andhra Pradesh), South India, using geospatial technologies. *Model Earth Syst Environ*, 6(2): 895–905, 2020. https://doi.org/10.1007/s40808-020-00726-9.

[25] Grauso S, Pasanisi F, Tebano C. Modeling the suspended sediment yield in Lesotho rivers. *Model Earth Syst Environ*, 6(2): 759–768, 2020. https://doi.org/10.1007/s40808-020-00738-5.

[26] Groundwater Decline and Depletion. https://www.usgs.gov/special-topic/water-scienceschool/science/groundwater-decline-and depletion; [Accessed 5 September 2020].

[27] Groundwater Management in Bangladesh. https://csisa.org/wpcontent/uploads/sites/2/2014/01/ Groundwater-management-in-Bangladesh; [Accessed 12 September 2020].

[28] Information on Earth's Water. https://www.ngwa.org/what-is-groundwater/About groundwater/information-on-earths-water; [Accessed 28 June 2020].

[29] Hasan K, Paul S, Chy TJ. et al. Analysis of groundwater table variability and trend using ordinary kriging: the case study of Sylhet, Bangladesh. *Applied Water Science*, 11(12): 1-14, 2021. https://doi.org/10.1007/s13201-021-01454-w.

[30] Hoque MA, Hoque MM, Ahmed KM. Declining groundwater Level and Aquifer Dewatering in Dhaka Metropolitan Area, Bangladesh: Causes and Quantification. *Hydrogeology*, 15(3): 1523-1534, 2007.

[31] Islam MA, Imran MBU, Biswas MHA. Modeling the Effects of Transmission Dynamics of Malaria: A Mathematical Approach on Healthcare. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Dubai, UAE, 2191-2202, 2020.

[32] Islam R, Islam MM, Islam MN et al. Climate change adaptation strategies: a prospect toward crop modelling and food security management. *Model Earth Syst Environ*, 6(2): 769–777, 2020. https://doi.org/10.1007/s40808-019-00708-6.

[33] Khalil HK. Nonlinear Systems. Third edition, Prentice Hall Inc., USA, 2002.

[34] Khan MJ, Shah BA, Nasir B. Groundwater quality assessment for drinking purpose: a case study from Sindh Industrial Trading Estate, Karachi, Pakistan. *Model Earth Syst Environ*, 6(1): 263–272, 2020. https://doi.org/10.1007/s40808-019-00676-x.

[35] Khatun MS, Biswas MHA. Optimal control strategies for preventing hepatitis B infection and reducing chronic liver cirrhosis incidence. *Infectious Disease Modelling*, 5: 91-110, 2020. https://doi.org/10.1016/j.idm.2019.12.006.

[36] Khatun MS, Biswas MHA. Modeling the Effect of Adoptive T cell Therapy for the Treatment of Leukemia. *Computational and Mathematical Method*, 2019. https://doi.org/10.1002/cmm4.1069.

[37] LaSalle JP. The Stability of Dynamical Systems. CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, USA, 1976.

[38] Mamo DK. Model the transmission dynamics of COVID-19 propagation with public health intervention, *Results in Applied Mathematics*, 7: (100123), 2020. https://doi.org/10.1016/j.rinam.2020.100123.

[39] Mandal S, Islam MS, Biswas MHA. Modeling the Potential Impact of Climate Change on Living Beings near Coastal Area. *Model Earth Syst Environ*, 2020. https://doi.org/10.1007/s40808-020-00897-5.

[40] Manuel W, Walter A, Nicole M, Raimund W. Modeling and simulation of curved fibers in dry spinning scenarios. *Results in Applied Mathematics*, 3: (100013): 1-10, 2019. https://doi.org/10.1016/j.rinam.2019.100013.

[41] Mridha MAH, Rahman SH. Prediction of groundwater trends for irrigation in Northern Bangladesh. *Water Supply*, 21(6): 2596–2614, 2021. https://doi.org/10.2166/ws.2021.076.

[42] Murray JD. Mathematical Biology: I. An Introduciton, Third Edition, Springer, New York, USA, 1989.

[43] National Water Policy, Ministry of Water Resources, Government of the People's Republic of Bangladesh, Dhaka, Bangladesh, 1999.

[44] Parvin M. The Rate of Decline and Trend Line Analysis of Groundwater Underneath Dhaka and Gazipur City. *Journal of Water Resource and Protection*, 11(3): 348-356, 2019. https://doi.org/10.4236/jwarp.2019.113020.

[45] Patel K, Mehta M, Patel T. A Series Solution of Moisture Content in Vertical Ground Water Flow through Unsaturated Heterogeneous Porous Media, *International Journal of Mathematics and Engineering*, 159(1): 1467-1477, 2012.

[46] Perko L. Differential Equations and Dynamical Systems. Third Edition, Springer-Verlag, New York, USA, 2001.

[47] Rahman M, Haque MM, Tareq SM. Appraisal of groundwater vulnerability in south-central part of Bangladesh using DRASTIC model: An approach towards groundwater protection and health safety. *Environmental Challenges* 5(100391): 11-20, 2021. https://doi.org/10.1016/j.envc.2021.100391.

[48] Safi MA, Garba SM. (2012) Global Stability Analysis of SEIR Model with Holling Type II Incidence Function. *Computational and Mathematical Methods in Medicine*, e826052. https://doi.org/10.1155/2012/826052.

[49] Scibek J, Allen DM. Modeled Impacts of Predicted Climate Change on Recharge and Groundwater Levels. *Water Resources Research*, 42: 11405-11423, 2006. https://doi.org/10.1029/2005wr004742

[50] Shijina V, Adithya U and Sunil J J. Similarity Measure of Multiple Sets and its Application to Pattern Recognition. *Informatica*, 44: 335–347, 2020. https://doi.org/10.31449/inf.v44i3.2872.

[51] Surface Water and Groundwater. https://www.britannica.com/technology/water-supplysystem/Surface-water and-Groundwater. [Accessed on 2 August 2020].

[52] Tesfaye TW, Dhanya CT, Gosain AK. Modeling the impact of climate change on the environmental flow indicators over Omo-Gibe basin, Ethiopia. *Model Earth Syst. Environ*, 6: 2063–2089, 2020. https://doi.org/10.1007/s40808-020-00813-x.

[53] Ross SL. Differential Equations. Third Edition, John Wiley and Sons Inc., New York, USA, 2004.

[54] Vargas C. Global Stability of Infectious Disease Models with Contact Rate as A Function of Prevalence Index. *Mathematical Biosciences and Engineering*, 14: 1019-1033, 2017. https://doi.org/10.3934/mbe.2017053.

[55] Verma V. Optimal control analysis of a mathematical model on smoking. *Model Earth Syst Environ*, 6: 2535–2542, 2020. https://doi.org/10.1007/s40808-020-00847-1.

[56] Weather and Climate Forecast Bangladesh. https:// www.weather atlas.com/en/Bangladesh/Dhaka-climate [Accessed on 2 August 2020].

[57] Wilson AB, Jenkins EW, Wang J, Husson SM. Numerical simulation of chemical separations using multimodal adsorption isotherms. *Results in Applied Mathematics*, 7(100122): 1-10, 2020. https://doi.org/10.1016/j.rinam.2020.100122.

Predicting the Usefulness of E-Commerce Products' Reviews Using Machine Learning Techniques

Dimple Chehal, Parul Gupta, Payal Gulati
Department of Computer Engineering, J.C. Bose University of Science and Technology, YMCA, Faridabad, India
E-mail: dimplechehal@gmail.com, parulgupta_gem@yahoo.com, gulatipayal@yahoo.co.in

Keywords: classification, e-commerce, machine learning, recommender system, usefulness, user reviews

Received: May 5, 2022

User-generated reviews are an essential component of e-commerce platforms. The presence of a large number of these reviews creates an information overload problem, making it difficult for other users to establish their purchase decision. A review voting mechanism, in which users can vote for or against a review, addresses this issue (as helpful or not). The helpful votes on a review reflect its usefulness to other users. As voting on usefulness is optional, not all reviews receive this vote. Furthermore, reviews posted recently by users are not associated with any vote (s). The aim of this paper is to predict the usefulness of user reviews through machine learning techniques. Using the Amazon product review dataset of cell phones, classification models are built on eight features and compared on seven performance measures. As per results, all the classification models performed well, except Linear Discriminant Analysis. The classification performance of Logistic Regression, Decision Tree, Random Forest, AdaBoost, and Gradient Boost was unaffected by feature selection or outlier removal. The performance of Linear Discriminant Analysis improved after feature selection but decreased after outlier removal, whereas ET and KNN classifiers improved in both cases.

Povzetek: Uporaba tehnik strojnega učenja za napovedovanje uporabnih ocen izdelkov e-trgovine.

1 Introduction

Online consumer reviews have evolved for e-commerce users and its stakeholders as an electronic word of mouth (eWoM) [32],[30]. Product reviews comprise of detailed experience of the customer(s) with the product(s). They help the consumers in their purchase decision, indicate any improvement required in the products' quality, thereby helping the business organizations in improving the products' sales. Mining customer reviews through sentiment analysis or topic modeling techniques reveal the customer's inclination towards a product. This helps in building the customer profile and understanding his/her preference for unseen products. Many platforms such as Amazon, Yelp, TripAdvisor, IMDB and Netflix are hosting a large number of user reviews [35]. However, the ever-growing rise in the number of products, customers and product reviews on the e-commerce platform, has led to the information overload problem and has made it infeasible for the customers to browse all the product reviews. To overcome this problem, voting a review as helpful by other customers had been introduced. While the rating of a product depicts a user's experience with a product, the votes gained by a review indicate its usefulness. The solution can be browsing user reviews according to their helpfulness or usefulness. But, due to factors such as humongous volume of electronic word of mouth, voluntary helpfulness voting mechanism, level of visibility and their recentness, all reviews do not receive this vote [5],[27]. Hence, the objective of this study is to categorize the product review according to its usefulness. This will not only help the customers to identify the products as useful or useless even if the review has not gained any votes but can also be fed as input to the recommender system for generating useful recommendations to the users. Also, through this study, the following questions have been answered-

- Which is the most efficient machine learning algorithm for the forecasting the usefulness of a product review?
- Which features should help in determining the usefulness of product review?

The results to the above questions have been obtained with the help of cell phone and accessories dataset taken from Amazon [3]. Eight different machine learning models, namely, Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), AdaBoost (ADA), Gradient Boost (GB), Extra Trees (ET), K Nearest Neighbors (KNN) and Linear Discriminant Analysis (LDA), have been trained and tested on existing and derived features and have been evaluated on seven evaluation metrics such as Area under the Curve (AUC), Accuracy (ACC), F1-score (F1) , Precision (P), Recall (R), Mathew's Correlation Coefficient (MCC) and Kappa score [19]. The best model has been fine tuned for prediction of usefulness of review. This study's contributions are stated as follows:

1. Through this research, features such as overall rating, user review, review summary, review

votes, word count of review, character count of review, review's sentiment score, average word count of review have been used to predict a review's usefulness.

2. Along with the already existing features in the Amazon dataset such as overall rating, user review, review summary and review votes, other features have been derived from user review and used in combination as input to the prediction model.
3. This study enables customers to identify useful reviews and e-commerce managers, merchants, retailers to improve the listing of product reviews based on the review usefulness.

The rest of the study is structured as follows: Section 2 consists of the related work, Section 3 details the research methodology followed. While section 4 discusses the result of different experiments conducted on the dataset. Lastly, section 5 concludes by discussing the limitations and future work.

2 Related work

Online user testimonials have gained much-needed prominence in the literature as they instill trust in other potential consumers in the online community [9], [17]. Product reviews can be viewed as a type of passive recommendation process or visibility of user sentiment for their past purchases [12]. Critical management choice for policy-makers is to manage user review to improve customer review efficacy. The academic evidence on review usefulness is largely driven and aided by review hosting platforms, which offer users' opinions on reviews' helpfulness explicitly. For instance on Amazon, customers not only access the rating and text content of each user review, but they also view the number of votes the review obtains from the fellow users and the number of helpful votes [35], [25]. Consumers benefit from informative reviews while making buying decisions. Some customers believe that favorable and unfavorable reviews are useful because such deeply divided records help to validate or invalidate purchase options. Others, on the contrary, find mixed reviews useful because they illustrate both the positives and negatives of the product under consideration. The perceived importance of a review to the end-user is also conveyed through the review's usefulness [28], [18]. This functionality, in particular, makes use of crowd-sourcing to assess the usefulness of reviews [6]. Every review includes the question, "Was this review helpful to you?" Customers who read the reviews may up vote or down vote the review [9], [12].

The research on review usefulness is roughly classified into two categories, prediction-based techniques to ascertain the review's usefulness and understanding of review usefulness. Machine learning classifiers, regression and deep learning approaches have been used to predict review helpfulness in the past [10], [8], [14], [16]. The review length, review timestamp, reviewer's expertise, and manner of writing reviews all have been used previously to predict helpful reviews [5]. Early indicators used to identify review usefulness through review length and review star rating also [24]. Deviation from the mean review length of a product, review's polarity and rating from the same user or on the same item to estimate review helpfulness helped in filtering high-quality ratings thereby improving the collaborative item recommendation process [27]. The moderate-length reviews outperform brief and lengthy ones as review length has inverted-U-shaped impact on usefulness [15]. Further, the more the review matches the language style of the target user, the more it is said to be readable. As a result, it is classified as a domain-specific indicator [22]. The semantic analysis of reviews comprises a wide range of methodologies that make use of structural characteristics like the count of product features cited in a review and its length [34]. The most useful reviews are said to be medium in length, have a lower score, and are negative or neutral in polarity [13]. Both critical evaluations containing data on service failures and favorable reviews highlighting essential product functionalities, technical elements, and aesthetics are seen as beneficial for usefulness prediction [1]. Besides the semantic aspect, neutral polarity reviews are regarded to be also useful [31]. The inclusion of adjectives, status and action verbs, as well as grammatical structure, are vital predictors of helpfulness, particularly when paired with factors such as review age, rating, readability, and subjectivity [21]. Highly readable reviews have proven to be the most beneficial. Based on previously performed emotion-based analysis, it has been concluded that male readers were more inclined to reviews with positive emotions, whereas female readers were more attracted to reviews with negative emotions. Previous findings also indicated that several features such as the review title's polarity, the review's sentiment and polarity, and the cosine similarity between the product review and the product title are contributing factors to determine the usefulness of user reviews [24]. As per the literature review, previous studies are deficient in terms of the combination of natural language processing tools and machine learning techniques for estimation of review usefulness. This study considering the above employs user voting as the target label to build the helpfulness or usefulness prediction system.

This depiction of helpfulness votes differs across platforms. Some platforms show the most helpful votes for a review, whereas others represent the usefulness as the "X of Y" concept. However, in prior methods, a ratio of 0.6 was considered as a helpfulness threshold in the "X of Y" approach of the consumer voting mechanism. Review usefulness, in particular, is critical in product rankings and recommendations [12]. Prediction of review usefulness enables users to compose meaningful reviews that shall assist retailers in intelligent website management by guiding it s users in purchase decisions [24]. The incorporation of a usefulness estimation model can aid in increasing the effectiveness of a collaborative filtering-based recommender system through optimization of data selection for user ratings estimation. This is a great resource for identifying relevant user

Table 1: Comparison of existing studies on identification of useful reviews

S. No	Paper	Model	Dataset	Input features	Performance Metric	Key points	Classification / Regression problem
1.	[20]	MLP, CNN with Trans E	Amazon dataset: CDs, Electronics, Video Games, Books	Product, Review, Reviewer features	Accuracy, F1-score	Dependence solely on hand crafted features leads to poor accuracy. Along with CNN another technique is required for mapping between the reviewer, product and reviews	Regression
2.	[7]	R:LNR, C:Log Reg, Both: DT, RF, GBT, NN	Yelp Shopping reviews	Product, Review, Reviewer features	RMSE, MSE, RAE, RSE, RRSE, MAE,R2 and CC (R), Accuracy, AUC, Precision, Recall, and F1 score (C)	Authors examine the impact of friends on review usefulness by introducing social network features. For classification, reviews receiving more than 3 votes are marked as helpful, 0 votes as unhelpful and discarded otherwise	Classification, Regression
3.	[11]	MLP, CNN	SiteJabber.com, ConsumerAffairs.com (DomainsDating, Wedding Dresses, Marketplace, Car Insurance, Travel Agencies, Mortgages)	Review features	Accuracy	Adjacent or neighbor reviews impact a user's helpfulness perception of a review. For classification, reviews receiving more than 2 helpful votes labeled as helpful and unhelpful otherwise.	Classification
4.	[27]	Linear Support Vector Regression, RF Regression	Yelp hotel stores reviews, Yelp food stores reviews	Review features	Pearson and Spearman correlation values	Deviations in star ratings, review's length and review's polarity with respect to user and item impact usefulness. Authors do not consider reviewer features. Random Forest was a better helpfulness predictor. Integration of such an estimation model improves the CF system performance.	Regression
5.	[24]	Multivariate adaptive regression, 'C' and 'R' tree, RF, NN, deep NN	Amazon multidomain sentiment analysis dataset	Review, Reviewer, Product features	MSE, RMSE,RRSE	Review type characteristics standout as effective characteristics to determine review's helpfulness as compared to reviewer and product characteristics. Combining all three characteristics yield best performance.	Regression
6.	[2]	DT, RF	Amazon Product dataset (Books, Office Products)	Review, Reviewer features	Accuracy, F-measure	Helpfulness threshold ratio set to value of 0.6. Features such as text, reviewer, and readability perform better than summary features. RF performed better than decision trees	Classification
7.	[26]	MLP, CART	Contributed dataset of	Review, Reviewer,	MSE, RAE, RMSE,	More the comments, polarity and sentiments in a review, more are the	Regression

		, Multivariate adaptive regression, Generalized Linear model, Ensemble model	34 product categories from Amazon.com	Product features	RRSE, MAE	helpful votes. Reviews with at least 10 votes are selected. Best results were obtained using hybrid features with ensemble model performing the best	
8.	[34]	RF	Dataset from JD.com	Review features, informativeness, length	Accuracy, AUC	Classification threshold for search and experience products to be different. Threshold of 4 for search products such as electronics and 2 for experience products such as skin gave the best model performance	Classification

reviews for decision-making [27]. Table 1 lists the key takeaway points from the existing literature.

In Table 1, AUC stands for Area Under the Curve, 'C'-Classification, CNN-Convolutional Neural Network, CC-Correlation Coefficient, DT-Decision Tree, GBT-Gradient Boosted Tree, Log Reg-Logistic Regression R-Regression, RAE-Relative Absolute Error, RF-Random Forest, RMSE-Root Mean Square Error, R2 -R Squared, RSE-Relative Squared Error, RRSE-Root Relative Squared Error, MAE-Mean Absolute Error, MLP-Multi Linear Perceptron, MSE-Mean Squared Error and NN-Neural Network.

3 Research methodology

The review-based recommendation methods in the studied literature utilize review contents and do not incorporate the associated helpfulness or usefulness scores. Incorporating this information of reviews helps in better exploitation of the user reviews. Since, several reviews don't have helpfulness scores, it is essential to predict the usefulness of these reviews [16]. The steps undertaken as part of prediction of usefulness of user reviews are shown in Figure 1 and are as follows:

3.1 Data collection

Data collection and its processing are the initial steps in all the machine learning methods [36]. Amazon cell phone and accessories dataset has been considered for this task [29] [3]. As shown in Table 2, the dataset with `(1048572, 12)` size has the following columns:

{
"reviewerID": "`A62MUEQU8I52E`*",*
"asin": "`B007SJZUSI`*",*
"reviewerName": " H. Moyer *",*
"vote": 3,
"style": { 'Color:': ' Gold' },
"reviewText": "Not a huge capacity power bank but a very good capacity for its very compact size. Exactly what I need, to have with me all of the time, just in case. One micro USB power input port for charging it, and one standard USB port for charging another device, either one using the same most standard cable in the industry. For most of us, power banks are for emergency only, so multiple output ports just add size unnecessarily. One state of charge gauge with 4 LED indicator lights, and one pushbutton. Very simple.",
"overall": 5.0,
"summary": " SIMPLE, COMPACT, AND POWERFUL FOR ITS SIZE ",
"unixReviewTime": `1490659200.0`*,*
"reviewTime": " 03 28, 2017 ",
"image": nan,
"verified": True
}

Table 2: Dataset description

Column name	Column description
reviewerID	Specifies the reviewer's unique identifier e.g. A284QS51P9P9V1
asin	Specifies the product's unique identifier e.g. B00UVSNVHA
reviewerName	Represents the name of the user/reviewer
vote	Represents the count of helpful votes received by a review
style	Represents a dictionary of the product metadata
reviewText	Implies the text contained in the review
overall	Represents the star rating given to a product
summary	Represents the textual summary of a product review
unixReviewTime	Represents the time at which review was generated (unix time)
reviewTime	Represents the time at which review was generated (raw)
image	Represents the product images that users post when they review the product

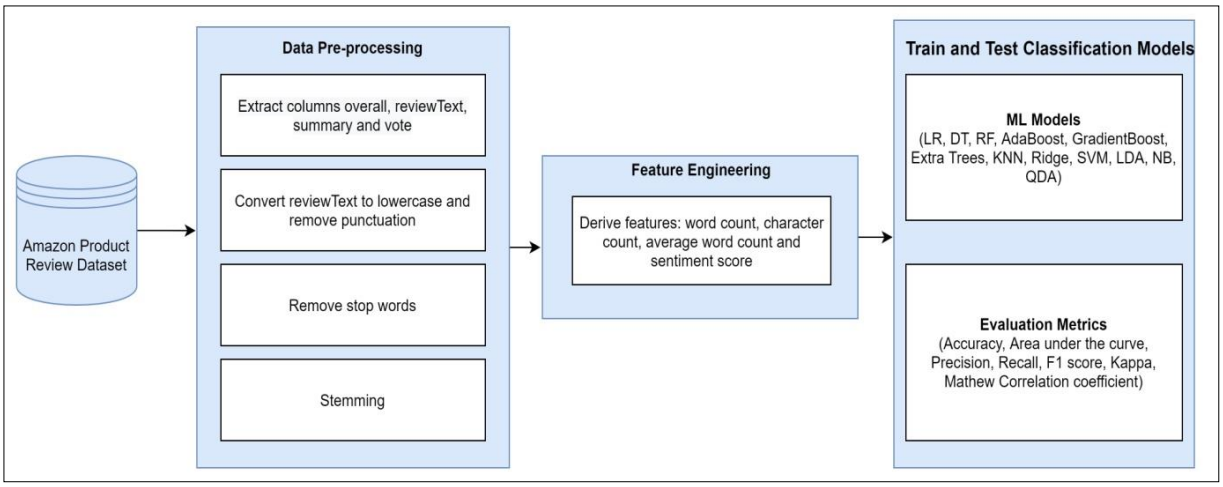


Figure 1: Research method

3.2 Pre-processing

The dataset consists of 12 columns and 1048572 rows. In order to categorize the reviews as useful or useless the following pre-processing steps have been undertaken:

1. Out of the 12 columns available, only columns- *overall*, *reviewText*, *summary* and *vote* have been utilized.
2. Next, *reviewText* column has been converted to lower case and punctuation has been removed.
3. After performing the below mentioned feature engineering steps, stop words using Python's *nltk* library have been removed.
4. Step 3 has been followed by stemming process in which porter stemmer has been used to apply stemming on *reviewText* column.

3.3 Feature engineering

Apart from the columns considered during the pre-processing phase, below mentioned columns have been derived:

1. *Word count*: This column represents number of words in a review
2. *Char count*: This column indicates number of characters in a review
3. *Avg word count*: This column stands for average word length of a review
4. *Sentiment score*: This column represents polarity of a review ranging from minus one (indicating extremely negative) to plus one (indicating extremely positive) which has been determined with the help of Python's *vaderSentiment* library

3.4 Preliminary analysis

1. The top ten most frequently occurring words, as shown in Table 3, after removal of stop words from the dataset are given below:

Table 3 Top ten frequently occurring words

Word	Frequency
Phone	165691
case	117779
one	62104
screen	57831
like	51122
use	43841
great	39611
battery	39595
would	38616
good	37078

As seen in Table 3, as the dataset is related to cell phones, the top ten frequently occurring words are related to this domain. The users have provided reviews mostly related to phone, case, screen and battery. To obtain these words, the frequency of words in the user reviews was obtained and then the top ten words were extracted.

2. The ten least frequently occurring words in the dataset, with only single occurrence are- Performancebattery, gummybearlike, amazonsunvalleytek, knive , terd, hh, nomy, 4siphone, Loosey, caseseems
3. The percentage of overall rating provided by users is provided in Table 4. The review dataset contains the majority of user reviews with the highest rating of the product, that is, 5, followed by user rating 4. The dataset contains more one-

star ratings than compared to three-star and two-star ratings.

Table 4 Distribution of user rating in the dataset

Rating	Count	Percentage (%)
5	49894	55.02
4	15243	16.81
3	8094	8.93
2	5509	6.08
1	11942	13.17

Supervised learning algorithms require input and output examples for training the model. In order to predict the review usefulness, the target column has been contributed which identifies each review as useful or not. To help the classification models learn if a review is useful or useless all the reviews with more than 10 votes have been marked as useful else useless.

3.5 Model selection

Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), AdaBoost (ADA), Gradient Boost (GB), Extra Trees (ET), k Nearest Neighbor (KNN) and Linear Discriminant Analysis (LDA) were used to categorize the usefulness of user reviews [23], [4]. All the models have been implemented in Python using *pycaret* library.

3.6 Data setup

Classification estimators were used in this study to predict the user review's usefulness. The target type is binary, with two possible values as useful or useless. The data has been partitioned into 70:30 partitions to obtain the training and testing sets. To allow row shuffling during the train-test split, the data split shuffle was set to true. The predictive models' performance was evaluated using stratified ten-fold cross-validation.

4 Result and discussion

Usefulness is treated as dependent variable and overall, reviewText, summary, vote, word count, character count, average word length and sentiment score are treated as independent variables. The model's performance can be assessed using a variety of evaluators, some of which are more appropriate than others. The models have been assessed in terms of accuracy (calculated using (1)), area under the curve, precision (calculated using (2)), recall (calculated using (3)), f1-score (calculated using (4)), kappa score (calculated using (5)) and MCC (calculated using (6)) as shown in Table 5. The Table also displays the time taken (TT) in seconds for the models to be trained.

- *Accuracy:* It is the most widely used performance metric and is calculated as the number of correct predictions over all predictions [33].

$$Accuracy = \left(\frac{TP + TN}{TP+TN+FP+FN}\right) * 100 \quad (1)$$

Where, TP stands for true positive, TN stands for true negative, FP stands for false positive and FN stands for false negative.

- *Area under the Curve:* The plot of sensitivity versus (1-specificity) is given by Receiver Operating Characteristic curve. AUC converts the curve to a numeric value. The ranges of the curve and their corresponding interpretations are grouped as excellent for range varying from 1 to 0.90; good from 0.90 to 0.80; fair from 0.80 to 0.70; poor from 0.70 to 0.60 and fail from 0.60 to 0.50.

- *Precision:*

$$Precision = TP/TP + FP \quad (2)$$

- *Sensitivity:* Sensitivity is the ratio of actually true classes that are identified correctly. Another name for sensitivity is true positive rate or recall. To reframe, it measures how often true predictions are correct.

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

- *F1 Score:* It's an accuracy metric that considers the trade-off between precision and recall.

$$F1\ Score = 2 * \frac{(Precision*Recall)}{Precision+Recall} \quad (4)$$

- *Kappa:* The Kappa score handles multi-class as well as imbalanced class problems.

$$Kappa = p_o - p_e/1 - p_e \quad (5)$$

Where, p_o and p_e denote the observed and expected agreement, respectively. In general, it reflects how a classifier performs as compared to another classifier that simply guesses at random based on each class's frequency. Cohen's kappa is never greater than 1. When the value of kappa is zero, the classifier is useless.

- *Matthews Correlation Coefficient (MCC):* The Matthews correlation coefficient assesses the quality of a binary classification problem; it is a balanced measure for unbalanced dataset as well. It outputs a value between minus one and plus one where, plus one indicates complete agreement between predicted and observed value, minus one indicates total disagreement, and zero value indicates random predicted values [33].

$$MCC = \frac{TP*TN-FP*FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (6)$$

Table 5: Performance of ML models

Model	Accuracy	AUC	Recall	Precision	F1-Score	Kappa	MCC	TT (sec)
LR	0.99	0.99	0.99	0.99	0.99	0.99	0.99	11.5
DT	0.99	0.99	0.99	0.99	0.99	0.99	0.99	**0.19**
RF	0.99	0.99	0.99	0.99	0.99	0.99	0.99	2.47
ADA	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.23
GB	0.99	0.99	0.99	0.99	0.99	0.99	0.99	6.48
ET	0.9657	0.9995	0.9997	0.9612	0.98	0.8596	0.8683	9.14
KNN	0.9263	0.9471	0.9912	0.9263	0.9576	0.6762	0.7004	1.92
LDA	0.5788	0.517	0.6078	0.6427	0.6233	0.3348	0.3439	27.73

As shown in Table 5 and Figure 2, most of the classification models are performing decently when contrasted according to the evaluation parameters. In order to test the model's robustness, ten-fold cross-validation has been employed. Due to lack of sufficient system RAM, the model has been fed a random sample of 5000 rows, leading to the above performance. Also, the methods' black-box state diminishes the results' interpretability. In comparison to others, LDA is unable to provide a reasonable prediction. The models have been trained again after performing feature selection and outlier removal to check the improvement in their performance. The near perfect performance of these models can be attributed to the size of data being fed to these models. Decision Tree model takes the least amount of time i.e. 0.19 seconds for generating the above results.

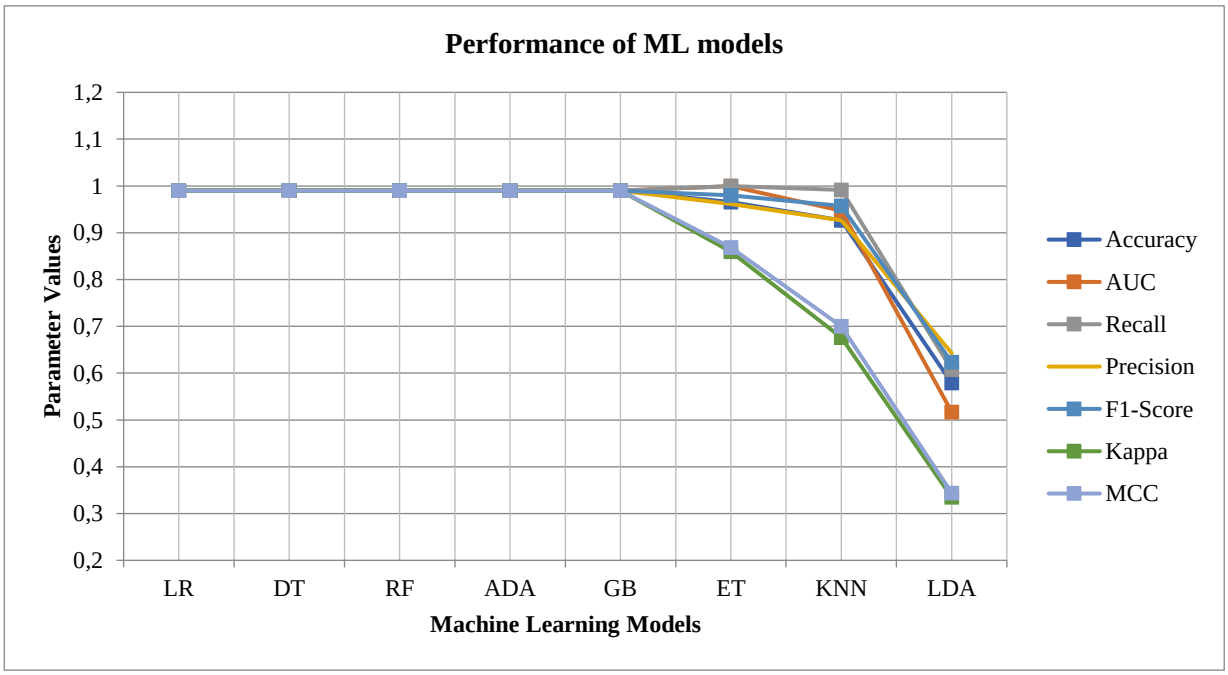


Figure 2: Performance of ML models

In order to improve the performance and reduce the training time of the above models, feature selection has been performed. Upon performing feature selection, the accuracy of LDA model jumps to 0.8411, AUC increases to 0.732, recall, precision, f1-score, kappa and MCC turn out to be 0.892, 0.842, 0.866, 0.638 and 0.658 respectively. The threshold value used for feature selection is set to 0.8 and the classic method of permutation feature importance techniques is used. Even after performing feature selection, the performance of LR, DT, RF, ADA, and GB classifiers remains unaffected as shown in Table 6.

As seen from Table 5 and Table 6, the training time of all the models reduced. Training time of model- LR reduced to 6.32 from 11.5 (without feature selection), DT remained the same as 0.19, ADA classifier remained the same as 0.23, ET reduced to 9.05 from 9.14, KNN reduced to 1.90 from 1.92 and LDA reduced to 26.75 from 27.73 seconds. Only two models RF and GB had their training time increased to 2.62 from 2.47 and 6.51 from 6.48 respectively. This increase

Table 6: Performance of ML models after feature selection

Model	Accuracy	AUC	Recall	Precision	F1-Score	Kappa	MCC	TT (sec)
LR	0.99	0.99	0.99	0.99	0.99	0.99	0.99	6.32
DT	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.19
RF	0.99	0.99	0.99	0.99	0.99	0.99	0.99	2.62
ADA	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.23
GB	0.99	0.99	0.99	0.99	0.99	0.99	0.99	6.51
ET	0.9634	0.99	0.99	0.959	0.979	0.848	0.858	9.05
KNN	0.9729	0.991	0.996	0.973	0.984	0.892	0.895	1.90
LDA	0.8411	0.732	0.892	0.842	0.866	0.638	0.658	26.75

Table 7 represents performance of classifiers after removal of outliers. Outliers from the training data have been reduced using the Singular Value Decomposition and the outlier threshold has been set to 0.05, that is, five percent of the outliers have been removed from the training dataset. Again, the performance of LR, DT, RF, ADA, and GB classifiers remained unaffected. While the accuracy of ET and KNN classifiers increased, that of LDA decreased significantly. This implies that ET, KNN and LDA classifiers are affected due to removal of outliers whereas the rest of the classifiers are not affected with this processing step.

Table 7: Performance of ML models after outlier removal

Model	Accuracy	AUC	Recall	Precision	F1-Score	Kappa	MCC	TT (sec)
LR	0.99	0.99	0.99	0.99	0.99	0.99	0.99	6.25
DT	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.18
RF	0.99	0.99	0.99	0.99	0.99	0.99	0.99	2.37
ADA	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.22
GB	0.99	0.99	0.99	0.99	0.99	0.99	0.99	6.42
ET	0.9759	0.9999	0.99	0.9726	0.9861	0.8969	0.9019	6.43
KNN	0.9801	0.9938	0.9979	0.9793	0.9885	0.9168	0.9192	1.83
LDA	0.7523	0.7173	0.7673	0.8541	0.806	0.4644	0.4877	23.48

As shown in Table 5, Table 6 and Table 7, LR, DT, RF, ADA and GB are performing perfectly for the sample dataset provided to the models with and without feature selection and outlier removal process. LDA model showed performance improvement after feature selection process, but degradation after outlier removal and accuracy of ET and KNN models improved after removal of outliers.

5 Limitations and future work

In this study, review usefulness prediction models were built and compared using collected features from the publicly available Amazon's cell phone and accessories dataset such as overall, reviewtext, summary, and vote, as well as derived features such as word count, character count, average word count, and sentiment score. Seven different performance measures namely, accuracy, area under the curve, precision, recall, f1-score, Kappa score and MCC were used to compare eight machine learning models- Logistic Regression, Decision Tree, Random Forest, AdaBoost, Gradient Boost, Extra Trees, K nearest Neighbor and Linear Discriminant Analysis. All the classification models performed well except LDA. Feature selection and outlier removal techniques had no effect on the classification performance of Logistic Regression, Decision Tree, Random Forest, AdaBoost, and Gradient Boost. The performance of LDA improved after feature selection but decreased after outlier removal, whereas ET and KNN depicted improvement in both cases. The results of this research can assist e-commerce platforms in gaining more clarity of the usefulness of online reviews. They can automatically analyze the usefulness of product reviews by utilizing prediction models as stated above. A system that uses ML models to predict useful reviews will benefit all stakeholders, including end users, product owners, and e-commerce platform regulators. In cases where the review has received no votes from people, such a system would be beneficial. In that instance, stakeholders might utilize the models' predictions to find reviews of interest or usefulness. This would ultimately save a significant amount of time spent reviewing the enormous number of available user reviews. This study was limited due to lack of sufficient system RAM; the models were fed a random sample of 5000 rows. Also, the methods' black-box state diminishes the results' interpretability. The study can be strengthened by improving the prediction models by removing fake reviews, incorporating emoticons for online review helpfulness prediction, employing unsupervised learning techniques instead of supervised learning, and developing deep learning models.

Availability of data

The dataset is available through URL: https://jmcauley.ucsd.edu/data/amazon/

References

[1] Ahmad, S.N. and Laroche, M. 2017. Analyzing electronic word of mouth: A social commerce

construct. *International Journal of Information Management*. 37, 3 (Jun. 2017), 202–213. DOI:https://doi.org/10.1016/J.IJINFOMGT.2016.08.004.

[2] Akbarabadi, M. and Hosseini, M. 2020. Predicting the helpfulness of online customer reviews: The role of title features. *International Journal of Market Research*. 62, 3 (2020), 272–287. DOI:https://doi.org/10.1177/1470785318819979.

[3] Amazon Review Data: 2018. *https://jmcauley.ucsd.edu/data/amazon/*. Accessed: 2021-05-14.

[4] Ampomah, E.K. et al. 2021. Stock market decision support modeling with tree-based AdaBoost ensemble machine learning models. *Informatica*. 44, 4 (Mar. 2021), 477–489. DOI:https://doi.org/10.31449/inf.v44i4.3159.

[5] Arif, M. et al. 2019. A Survey of Customer Review Helpfulness Prediction Techniques. *Advances in Intelligent Systems and Computing*. Springer International Publishing. 215–226.

[6] Bilal, M. et al. 2019. Profiling and predicting the cumulative helpfulness (Quality) of crowd-sourced reviews. *Information (Switzerland)*.

[7] Bilal, M. et al. 2021. Profiling reviewers' social network strength and predicting the "Helpfulness" of online customer reviews. *Electronic Commerce Research and Applications*. 45, (Jan. 2021), 101026. DOI:https://doi.org/10.1016/j.elerap.2020.101026.

[8] Chen, C. et al. 2019. Multi-domain gated CNn for review helpfulness prediction. *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*. (2019), 2630–2636. DOI:https://doi.org/10.1145/3308558.3313587.

[9] Chua, A.Y.K. and Banerjee, S. 2016. Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality. *Computers in Human Behavior*. 54, (2016), 547–554. DOI:https://doi.org/10.1016/j.chb.2015.08.057.

[10] Dey, D. and Kumar, P. 2019. A novel approach to identify the determinants of online review helpfulness and predict the helpfulness score across product categories. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer. 365–388.

[11] Du, J. et al. 2021. Neighbor-aware review helpfulness prediction. *Decision Support Systems*. April (2021), 113581. DOI:https://doi.org/10.1016/j.dss.2021.113581.

[12] Enamul Haque, M. et al. 2018. Helpfulness prediction of online product reviews. *Proceedings of the ACM Symposium on Document Engineering 2018, DocEng 2018*. (2018). DOI:https://doi.org/10.1145/3209280.3229105.

[13] Eslami, S.P. et al. 2018. Which online reviews do consumers find most helpful? A multi-method investigation. *Decision Support Systems*. 113, (Sep. 2018), 32–42. DOI:https://doi.org/10.1016/J.DSS.2018.06.012.

[14] Fan, M. et al. 2019. Product-aware helpfulness prediction of online reviews. *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*. 2, Ccl (2019), 2715–2721. DOI:https://doi.org/10.1145/3308558.3313523.

[15] Fink, L. et al. 2018. Longer online reviews are not necessarily better. *International Journal of Information Management*. 39, (Apr. 2018), 30–37. DOI:https://doi.org/10.1016/J.IJINFOMGT.2017.11.002.

[16] Ge, S. et al. 2019. Helpfulness-aware review based neural recommendation. *CCF Transactions on Pervasive Computing and Interaction*. 1, 4 (Dec. 2019), 285–295. DOI:https://doi.org/10.1007/s42486-019-00023-0.

[17] Hamad et al. 2018. Review helpfulness as a function of Linguistic Indicators. *IJCSNS International Journal of Computer Science and Network Security*. 18, 1 (2018), 234–240.

[18] Hong, H. et al. 2017. Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems*. 102, (2017), 1–11. DOI:https://doi.org/10.1016/j.dss.2017.06.007.

[19] Kaddoura, S. et al. 2022. A systematic review on machine learning models for online learning and examination systems. *PeerJ Computer Science*. 8, (May 2022), e986. DOI:https://doi.org/10.7717/PEERJ-CS.986.

[20] Kong, L. et al. 2022. Predicting Product Review Helpfulness - A Hybrid Method. *IEEE Transactions on Services Computing*. 15, 4 (2022), 2213–2225. DOI:https://doi.org/10.1109/TSC.2020.3041095.

[21] Krishnamoorthy, S. 2015. Linguistic features for review helpfulness prediction. *Expert Systems with Applications*. 42, 7 (May 2015), 3751–3759. DOI:https://doi.org/10.1016/J.ESWA.2014.12.044.

[22] Liu, A.X. et al. 2019. It's Not Just What You Say, But How You Say It: The Effect of Language Style Matching on Perceived Quality of Consumer Reviews. *Journal of Interactive Marketing*. 46, (May 2019), 70–86. DOI:https://doi.org/10.1016/J.INTMAR.2018.11.001.

[23] Luo, Y. and Xu, X. 2019. Predicting the helpfulness of online restaurant reviews using different machine learning algorithms: A case study of yelp. *Sustainability (Switzerland)*. 11, 19 (2019). DOI:https://doi.org/10.3390/su11195254.

[24] Malik, M.S.I. 2020. Predicting users' review helpfulness: the role of significant review and

reviewer characteristics. *Soft Computing*. 24, 18 (Sep. 2020), 13913–13928. DOI:https://doi.org/10.1007/s00500-020-04767-1.

[25] Malik, M.S.I. and Hussain, A. 2018. An analysis of review content and reviewer variables that contribute to review helpfulness. *Information Processing and Management*. 54, 1 (2018), 88–104. DOI:https://doi.org/10.1016/j.ipm.2017.09.004.

[26] Malik, M.S.I. and Hussain, A. 2020. Exploring the influential reviewer, review and product determinants for review helpfulness. *Artificial Intelligence Review*. 53, 1 (2020), 407–427. DOI:https://doi.org/10.1007/s10462-018-9662-y.

[27] Mauro, N. et al. 2021. User and item-aware estimation of review helpfulness. *Information Processing and Management*.

[28] Mitra, S. and Jenamani, M. 2021. Helpfulness of online consumer reviews: A multi-perspective approach. *Information Processing and Management*. 58, 3 (2021), 102538. DOI:https://doi.org/10.1016/j.ipm.2021.102538.

[29] Ni, J. et al. 2020. Justifying recommendations using distantly-labeled reviews and fine-grained aspects. *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*. (2020), 188–197. DOI:https://doi.org/10.18653/v1/d19-1018.

[30] Orimaye, S.O. et al. 2016. Learning Sentiment Dependent Bayesian Network Classifier for Online Product Reviews. *Informatica (Slovenia)*. 40, 2 (2016), 225–235.

[31] Salehan, M. and Kim, D.J. 2016. Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems*. 81, (Jan. 2016), 30–40. DOI:https://doi.org/10.1016/J.DSS.2015.10.006.

[32] Saumya, S. et al. 2018. Ranking online consumer reviews. *Electronic Commerce Research and Applications*. 29, (2018), 78–89. DOI:https://doi.org/10.1016/j.elerap.2018.03.008.

[33] Sidhu, R.K. et al. 2020. Machine learning based crop water demand forecasting using minimum climatological data. *Multimedia Tools and Applications*. 79, 19–20 (2020), 13109–13124. DOI:https://doi.org/10.1007/s11042-019-08533-w.

[34] Sun, X. et al. 2019. Helpfulness of online reviews: Examining review informativeness and classification thresholds by search products and experience products. *Decision Support Systems*. 124, (Sep. 2019), 113099. DOI:https://doi.org/10.1016/J.DSS.2019.113099.

[35] Wu, J. 2017. Review popularity and review helpfulness: A model for user review effectiveness. *Decision Support Systems*. 97, (2017), 92–103. DOI:https://doi.org/10.1016/j.dss.2017.03.008.

[36] Yenkikar, A. et al. 2022. Semantic relational machine learning model for sentiment analysis using cascade feature selection and heterogeneous classifier ensemble. *PeerJ Computer Science*. 8, (Sep. 2022), e1100. DOI:https://doi.org/10.7717/PEERJ-CS.1100.

Design and Development of Mobile Terminal Application Based on Android

Wen Yang*, Ming Zhan, Zhijun Huang, Wei Shao
Jiangxi Medical College, Shangrao Jiangxi, 334000, China
E-mail: wenyan317@126.com*, mingzhan61@163.com, zhijunhuang@126.com, weishao291@163.com
*Corresponding author

Keywords: android, mobile terminal, design and development, OCR application

Received: February 15, 2022

This article addresses the design and development of mobile terminal application based on Android. This paper proposes the design and development of character recognition system application based on Android platform. The implementation process of image acquisition module, image clipping module, image preprocessing module, character recognition, recognition history display module and recognition result post-processing module is introduced in this article. This article presents the method of character training by Tesseract. By using the training tool based on LSTM (long short - term memory) neural network to train the sample set, a specific character training set is obtained, and OCR (optical character recognition) can be used in specific occasions. The function and performance of the system are tested, and the experimental results are analyzed. After the whole OCR system is built and deployed, the operation and completion of the whole OCR system can be understood through the test and analysis of system functions. It can be seen from the test results that the average response time of text images in pure English is the fastest. The response time of text images in pure Chinese is the second, and the slowest is the mixed arrangement of Chinese and English. For pure English text images, the character recognition accuracy is about 90%, and for pure Chinese text images, the recognition accuracy is close to 90%. However, for the mixed arrangement of Chinese and English, the accuracy of character recognition is lower than that of pure Chinese and English. The accuracy of most commercial character recognition software is about more than 90%. Except for the mixed arrangement of Chinese and English, other products can basically achieve the accuracy of general commercial character recognition software. The product has certain practicability and can be applied to recognize text images taken in natural scenes in daily life.

Povzetek: Članek predstavi novo aplikacijo za prepoznavanje znakov na platformi Android. Z uporabo LSTM nevronske mreže in OCR je izboljšana natančnost prepoznave angleških in kitajskih znakov.

1 Introduction

With the growth of wireless network coverage year by year and the continuous increase of mobile terminals, the number of mobile Internet users has increased rapidly. Mobile phones have not only met simple communication functions for people, but also developed into a tool with internet access, music, video, games and other functions [1]. At the same time, with the rapid development of mobile Internet, the amount of data used by mobile clients is gradually approaching the PC end, and the main thrust of this phenomenon should be attributed to the rapid development of mobile app applications [2]. Moreover, the coverage of WiFi has been continuously expanded, which also provides a larger platform for the development of APP applications on the mobile terminal, which has always been a limitation that the PC terminal is difficult to surpass [3]. OCR (optical character recognition), which means optical character recognition, is a process of obtaining text and layout information through a series of decomposition processing of text image files [4]. Specific to the scene, it is to automatically identify and input your business card, ID card, driver's license and bank card into terminals such as computer or mobile phone through scanning, eliminating the process of manual input. Due to the shift of the consumer market to mobile terminals, more than 60% of the data in the future will come from smart mobile terminals such as mobile phones and tablets. Smart camera has become the main entrance of data acquisition. More unstructured data needs to be transformed into retrievable data at the front and back. This transformation process requires OCR technology to show its skills. However, the fundamental way of the current mobile phone OCR system is to use the mobile phone to capture images, compress the images, transmit them to the background server through the network, recognize characters and return the corresponding results. However, with the improvement of mobile phone performance, the captured images will be more and more fresh, and the image size will gradually increase [5]. If the original image is transmitted directly without compressing the image, it will lead to large traffic consumption and prolonged network transmission; If the image is compressed, some

information will be lost, and the compression algorithm itself has a certain time consumption. At the same time, this method depends on network connection, not strictly mobile phone-based OCR system. The overall process of OCR system is shown in Figure 1. With the rapid improvement of the computing power of mobile intelligent terminal, the feasibility of using the computing power of mobile terminal to complete picture preprocessing, feature extraction and character recognition at the terminal is gradually improved [6]. In this way, the image compression and transmission process can be avoided, there is no network transmission delay, and it can be independent of the mobile network.

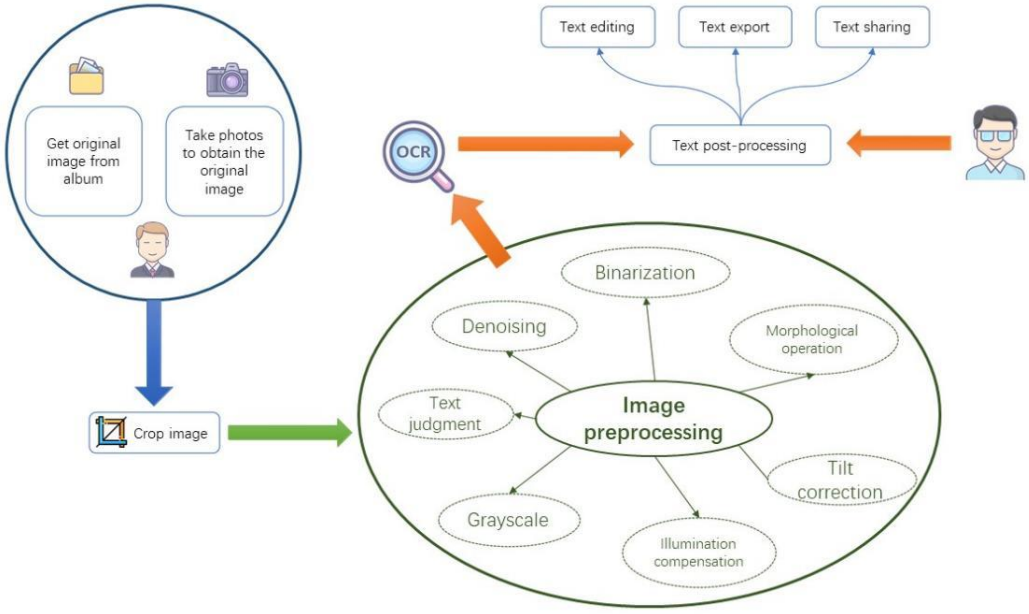


Figure 1: Overall process of OCR system.

At the same time, feature extraction and character recognition are carried out on the original image, resulting in less information loss and higher OCR accuracy [7].

This article addresses the design and development of mobile terminal application based on Android platform by proposing a design and development of character recognition system application. The implementation process of image acquisition module, image clipping module, image preprocessing module, character recognition, recognition history display module and recognition result post-processing module is introduced. The article presents the method of character training by Tesseract. By using the training tool based on LSTM (long short - term memory) neural network to train the sample set, a specific character training set is obtained, and OCR (optical character recognition) can be used in specific occasions. The function and performance of the system are tested, and the experimental results are analyzed. After the whole OCR system is built and deployed, the operation and completion of the whole OCR system can be understood through the test and analysis of system functions. It can be seen from the test results that the average response time of text images in pure English is the fastest. Except for the mixed arrangement of Chinese and English, other products can basically achieve the accuracy of general commercial character recognition software. The product has certain practicability and can be applied to recognize text images taken in natural scenes in daily life.

The rest of this article is arranged as: Literature review is presented in section 2 and the methods discussing the traditional character recognition algorithm, tesseract recognition algorithm, image acquisition, clipping and preprocessing, recognition and text post-processing is presented in section 3. Results and analysis are presented in section 4 followed by conclusion in section 5.

2 Literature review

To solve this problem, Xu *et al.* studied the design and development of intelligent webcam application based on Android platform [8]. Wang *et al.* found that Android is an open software system containing many source codes. Its system architecture is divided into four levels [9]. The software named wavesecure launched by Feng, is a mobile phone security protection software, which is specially used to backup and restore data on a variety of devices, so that the important data in the user's mobile phone can be fully protected [10]. Cai *et al.* studied a mobile device management solution, which mainly manages and configures enterprise mobile devices and employees' own mobile devices based on roles, and ensures the security of enterprise content stored in the devices [11]. Ameigeiras *et al.* studied and analyzed that text secure is an application software for SMS encryption of mobile terminals, which provides SMS encryption,

SMS session encryption and other functions [12]. Lin *et al.* found that Twitter has opened the source code of the project on git hub. The application uses a combination of symmetric encryption algorithm and asymmetric encryption algorithm to protect the transmitted data, which has good security [13].

Cao *et al.* found that Twitter uses digital signature technology to ensure the non-repudiation of data. Although the application software enhances the security of short message communication between individual users and enterprise users, it can only encrypt short messages and short message sessions, and cannot provide more comprehensive protection for data and devices in mobile terminals. Therefore, it also has limitations that cannot be ignored [14]. Xin *et al.* proposed a file remote data backup system on Android platform based on the research of remote-control technology [15]. Zhang proposed a remote security management model of Android mobile terminal based on RFB (remote frame buffer) protocol, which can support remote opening of camera, recorder, Bluetooth, positioning device, erasing data and other functions. However, this protocol is a thin client protocol, so that the service platform cannot monitor the client status [16].

Based on the current research, this paper proposes the design and development of character recognition system application based on Android platform. The implementation process of image acquisition module, image clipping module, image preprocessing module, character recognition, recognition history display module and recognition result post-processing module is introduced in detail, and the implementation algorithm is analyzed. The method of character training by Tesseract is introduced. The training tool is used based on LSTM neural network to train the sample set, a specific character training set is obtained, and OCR (optical character recognition) can be used in specific occasions. The experimental results show that the system can quickly carry out image preprocessing and character recognition by relying on gradually powerful mobile computing resources, and efficiently complete the character recognition task at the mobile end.

3 Research methodology

This section includes the discussion of adopted methodology of character recognition algorithm and image acquisition and recognition module.

3.1 Traditional character recognition algorithm

Pattern recognition can recognize the target object through two machine learning methods: supervised learning and unsupervised learning [17]. Supervised pattern recognition classifiers need to judge the results based on a priori known conditions, that is, through known labels; The similarity classifier, which does not use a priori known conditions to train the character set, but expresses the results according to a set of feature vectors, is an unsupervised pattern recognition classifier. Template matching is one of the simplest pattern recognition algorithms. It is often used to find sub image regions that are the same or highly similar to the template image from an unknown image according to the predefined template image in image processing. Therefore, template matching requires two inputs, one template image and the other is the target image to be detected [18]. Template matching uses the calculation method based on image pixel similarity, which is easy to be affected by illumination intensity and object geometric distortion, and has high accuracy only when the brightness and resolution are constant and there is no geometric distortion. The template matching methods based on pixel calculation include the following 6 kinds.

TM_SQDIFF square difference:

$$R(x,y)=\sum_{x',y'}\left(T(x',y')-I(x+x',y+y')^2\right) \tag{1}$$

TM_SQDIFF_NORMEND normalized squared difference:

$$R(x,y)=\frac{\sum_{x',y'}\left(T(x',y')-I(x+x',y+y')^2\right)}{\sqrt{\sum_{x^1,y^1}T(x',y')^2\cdot\sum_{x^1,y^1}I(x+x',y+y')^2}} \tag{2}$$

TM_CCORR relevance:

$$R(x,y)=\sum_{x',y'}\left(T(x',y')\right)\cdot I(x+x',y+y') \tag{3}$$

TM_CCORR_NORMEND normalized correlation:

$$R(x,y)=\frac{\sum_{x',y'}\left(T(x',y')\cdot I(x+x',y+y')\right)^2}{\sqrt{\sum_{x',y'}T(x',y')^2\cdot\sum_{x',y'}I(x+x',y,y')^2}} \tag{4}$$

TM_CCOEFF correlation factor:

$$R(x,y)=\sum_{x',y'}\left(T'(x',y')\right)\cdot I'(x+x',y+y') \tag{5}$$

TM_CCORR_NORMEND normalized correlation factor:

$$R(x,y)=\frac{\sum_{x',y'}\left(T'(x',y')\cdot I(x+x',y+y')\right)^2}{\sqrt{\sum_{x',y'}T'(x',y')^2\cdot\sum_{x',y'}I'(x+x',y+y')^2}} \tag{6}$$

3.2 Tesseract recognition algorithm

Recurrent neural network RNN is a kind of neural network used to process sequence $X^{(1)},\ldots,X^{(T)}$, which can be extended to longer sequences. However, the gradient of RNN tends to disappear after many stages of propagation. Even if we assume that the cyclic network

parameters are stable, the difficulty of long-term dependence comes from the weight smaller than the index of short-term interaction. In theory, RNN can deal with long-term dependence problems, and people can carefully select parameters to solve the most primary form of such problems, but in practice, RNN can certainly not successfully learn this knowledge. RNN cannot process sequences with a long distance because the gradient is likely to disappear during training, that is, exponential reduction is likely to occur during training through formula (7), so that RNN loses its ability to perceive distant times [19].

$$\frac{\partial E}{\partial W}=\sum{}^{t}\frac{\partial E_t}{\partial W}=\sum\nolimits_{k=0}^{t}\frac{\partial E_t}{\partial net_t}\frac{\partial net_t}{\partial s_t}\left(\Pi_{j=k+1}^{t}\frac{\partial s_t}{\partial_{sK}}\right)\frac{\partial_{sK}}{\partial W} \quad (7)$$

Because RNN has the problem of gradient disappearance, Tesseract introduces LSTM algorithm as the neural network structure for training and recognition. LSTM (long short-term memory) is a long-term and short-term memory network. It is a time recursive neural network, which is suitable for processing and predicting important events with relatively long interval and delay in time series. LSTM is a special RNN, which can learn long-term dependent information. LSTM has achieved great success in many applications, such as unconstrained handwriting recognition, speech recognition, handwriting generation, machine translation, image title generation and parsing. LSTM is different from a single neural network layer.

The structure of the repetitive network module of LSTM is much more complex. It realizes three gate calculations, namely forgetting gate, input gate and output gate. Each gate is responsible for different tasks, and the forgetting gate is responsible for determining how many cell states from the previous time to the current time [20]. The input gate is responsible for determining how much current time input is reserved to the unit state of the current time; The output gate is responsible for determining how many outputs the unit state has at the current time. In addition to the external RNN cycle, LSTM also has an internal "LSTM cell" cycle. Therefore, LSTM is not a gating unit system with simple affine transformation of input and loop units and element by element nonlinearity. Similar to ordinary loop networks, each unit has the same input and output, but also has more parameters and control information flow. The most important component is the state unit $S_i^{(t)}$. the self loop weight is controlled by the forgetting Gate $f_i^{(t)}$, and the sigmoid unit sets the weight to a value between 0 and 1. The forgetting gate formula is shown in equation (8).

$$f_i^{(t)}=\sigma\left(b_i^{f}+\sum_j U_{i,j}^{f}x_j^{(t)}+\sum_j W_{i,j}^{f}h_j^{(t-1)}\right) \quad (8)$$

Where $X^{(t)}$ is the current input vector, h^t is the current hidden layer vector, and h^t contains the outputs of all LSTM cells. B^f, U^f, W^f are offset, input weight and cycle weight of forgetting gate respectively. Therefore, the internal state of LSTM cells is updated in the way of formula (9), in which there is a conditional self-ring weight $f_i^{(t)}$.

$$S_i^{(t)}=f_i^{(t)}s_i^{(t-1)}+g_i^{(t)}\sigma\left(b_i+\sum_j U_{i,j}x_j^{(t)}+\sum_j W_{i,j}h_j^{(t-1)}\right) \quad (9)$$

Where, b, U and W are the bias, input weight and cycle weight of forgetting gate in LSTM cells respectively. The external input gate unit $g_i^{(t)}$ is updated in a manner similar to the forgetting gate, but has its own parameters, as shown in equation (10).

$$g_i^{(t)}=\sigma\left(b_i^{g}+\sum_j U_{i,j}^{g}x_j^{(t)}+\sum_j W_{i,j}^{g}h_j^{(t-1)}\right) \quad (10)$$

The output h_i^t of LSTM cells (see equation (11) can be closed by the output gate q_i^t (see equation (12).

$$h_i^{(t)}=\tanh\left(s_i^{(t)}\right)q_i^{(t)} \quad (11)$$

$$q_i^{(t)}=\sigma\left(b_i^{0}+\sum_j U_{i,j}^{0}x_j^{(t)}+\sum_j W_{i,j}^{0}h_j^{(t-1)}\right) \quad (12)$$

Where b^0, U^0, W^0 are the cyclic weight of offset, input weight and forgetting gate respectively. LSTM network is easier to learn and rely on for a long time than simple cyclic architecture, which is very suitable for optical character recognition of long sequences.

3.3 Image acquisition module

The image acquisition module is used to obtain the text image to be recognized. There are two ways to obtain the text image. It mainly completes the task of text image through the two sub modules of obtaining the original image through photo album and photographing. The function of photo album to obtain the original image is realized by the photo album calling interface of Android system. Taking photos to obtain the original image is to obtain the image by taking photos with the camera provided by Android. The image acquisition module is the basis for the operation of the whole software.

After entering the system, the user first needs to select the image acquisition method. The original image obtained from the album of the system is the default interface to enter the system. The text image in the album is loaded in the interface. By selecting the photographing mode in the column at the bottom of the system, the source image can also be obtained by photographing.

Only through the image acquisition module can the next step of image preprocessing and image recognition [21].

3.4 Image clipping module

The original image selected by the user may contain non text areas. The main function of the system is to recognize characters. Therefore, the non-text areas need to be cropped and the text area ROI of interest to the user is selected, so as to reduce unnecessary calculation during later image processing and recognition, so as to improve the accuracy of character positioning and recognition. The implementation part of the system is that the user cuts out the text area manually.

The clipping module mainly includes clipping, image rotation, clipping region scaling and clipping region moving functions. Clipping is used for the user to manually select the text area of interest through block diagram selection [22]. Rotation is used to deal with large angle text tilt. It mainly aims at the image rotation greater than 90 degrees caused by the photographing angle, and realizes the function of 90 degrees rotation of the image after cutting. For the inclined image less than 90 degrees, the image needs to be corrected by tilt correction. The image zoom function is to zoom in and out the image of the clipping area; Picture movement is to move the position of the background original image of the clipping area, and fine tune the user's clipping area through the movement function.

3.5 Image preprocessing module

The input image of the image preprocessing module is the Bitmap image after image clipping by the image clipping module, and the output end is the bitmap format image after graying, illumination compensation, filter denoising, binarization, tilt correction and morphological operation. The function code of the picture preprocessing module is implemented on the recognition page. After the user clicks the character recognition button, the image preprocessing module is called for pre-processing. Because image preprocessing is a time-consuming operation, if it is operated directly on the main identification interface, the system will get stuck and even the program will not respond and crash. In order to solve this problem, the system uses the asynchronous call mode of Handler+Loo per+Message Queue in Android to process the time-consuming operation of the image processing module. Asynchronous processing sub threads mainly do some time-consuming operations. After the sub thread completes the operation, it sends a message to the main thread. The system communicates between the sub thread and the handler in the main line through the message circulation mechanism. The specific work is that the sub thread sends a message after completing the time-consuming operation, and the handler refreshes the UI on the main thread after receiving the message, so as to avoid the user interface getting stuck and unresponsive due to the time-consuming preprocessing process in the background. In this paper, events related to the user interface in Android asynchronous tasks will be handled by the main thread. In the Android system, the UI interface of the system needs to be updated through the main thread. In order to avoid blocking the main thread of the operation interface, this paper processes time-consuming tasks such as image preprocessing and character recognition in the background by opening a new thread [23].

3.6 Image recognition module

The input end of the image recognition module is the output image preprocessed by the background image, and the output end is the structured text information after character recognition. If the identification module is called in the main thread, the system will crash if the waiting time of the main thread is too long. Therefore, the background of the system also needs to adopt the asynchronous call mode of Handler+Loo per+Message Queue to deal with the time-consuming operation of the character recognition module. The interface of the character recognition module mainly uses the material design control Floating Action Button, which can easily realize the effect of floating buttons. For floating buttons, you can set the transparency and color of buttons. For the selection of recognition text, this paper uses the spinner drop-down box button to realize the recognition text. The recognition text can choose to recognize Chinese and English.

3.7 Text post-processing module

This module is used for post-processing the recognized text, mainly for correcting the misrecognized text in the source text image. The text post-processing module includes the functions of text original image comparison and correction, text editing, text copying, exporting text, sharing and so on. The proofreading function is to load the processed gray-scale image. In the same activity, the user can edit and change the original image. Exporting the text can help the user export the recognized text into PDF format. The text can be saved on the user's computer and shared with friends to help friends with text recognition. The text post-processing module is developed based on the Android native API. The whole implementation interface of the text post-processing module is simple and easy to use.

The system implementation flow of character recognition module is shown in Figure 2. For the function realization of the core module of character recognition, this paper uses Tesseract recognition engine to realize the function of character recognition, and calls Tesseract through JNI to realize the complete character recognition function of the image in the form of bitmap at the input and string characters at the output.

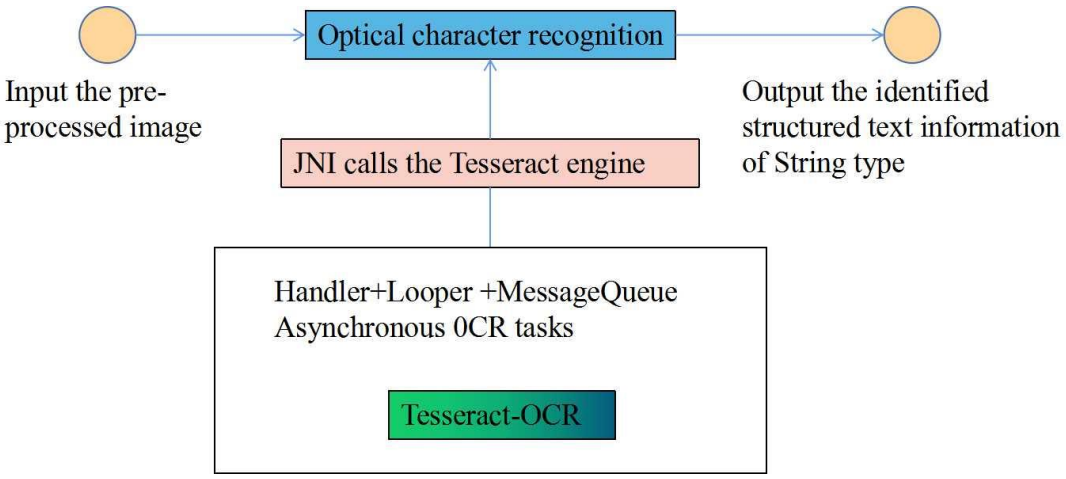


Figure 2: OCR implementation process.

4 Results and analysis

The function and performance of the system are tested, and the experimental results are analyzed. After the whole OCR system is built and deployed, the operation and completion of the whole OCR system can be understood through the test and analysis of system functions.

4.1 System test environment

The OCR character recognition application based on Android needs the support of the camera, so it cannot continue to use the virtual Android machine in the Android Studio development environment for testing. The application needs to support two ways: photo album and photographing to obtain the text source image to be recognized. The application is aimed at ordinary Internet users. The software running environment is determined through analysis, including hardware environment and software environment. The specific test running environment of the system is shown in Table 1.

Software / Hardware	Version / Configuration
Operating system	Android 7.0
Database	Sq1Lite
Image processing	Open CV 3. 2
Recognition engine	Tesseract 4.0
CPU	Qualcomm Snapdragon 835
RAM	6G
ROM	128G
Camera	21 00W
Resolving power	1080*1920

Table 1: Test environment

4.2 Function test

The system function test is mainly to verify whether the system can run the application normally and stably, whether each functional module can run normally, whether the data obtained by the system is accurate, and whether the speed of system image processing and the accuracy of character recognition meet the design requirements. After the development of each functional module of the system is completed, we have deployed the character recognition system in the above operating environment. First, install the character recognition software on the Android system, enter the first interface of the system, run the installed and deployed OCR recognition application, and start testing the main functions of the recognition system. According to the test cases of the system, the functional unit module and system test are carried out for the image acquisition module, cutting module, character recognition module, character recognition history display module and text post-processing module of the system. The test results show that all functional modules of the system operate normally, indicating that the functional modules of the system meet the product design requirements. The UI interface of each functional module of the system is convenient and practical, and the overall layout is reasonable, which is in line with the aesthetics of public users. It shows that the interface effect based on material design style is good and meets the interface requirements of the system.

4.3 Performance test

The main purpose of system performance test is to test the operation effect of the system under heavy load, and then find some possible defects of the system. In this system, the performance test object is mainly to test the system response time in different languages and different text mixing. For the character recognition system, the two key factors affecting the system response time are image preprocessing and character recognition. Image preprocessing and character recognition are functional modules with large amount of calculation. For character recognition, the two main indicators affecting the response time are the working efficiency of the classifier and the size of the character library. The larger the character library, the more character sets the system needs to match, and the more time it needs to consume. System response time is an important index to measure software performance. Through software testing, the

time-consuming of image preprocessing and classification recognition in image recognition is counted. Through detailed analysis of the time-consuming of each small step, the operation bottleneck in the system is found, so as to provide scientific data basis for subsequent optimization of software performance. The test text images used in this paper are 150, 50 pure Chinese, 50 pure English and 50 Mixed Chinese and English, with about 50 characters each. After many tests, the statistical test results are shown in Figures 3, 4 and 5.

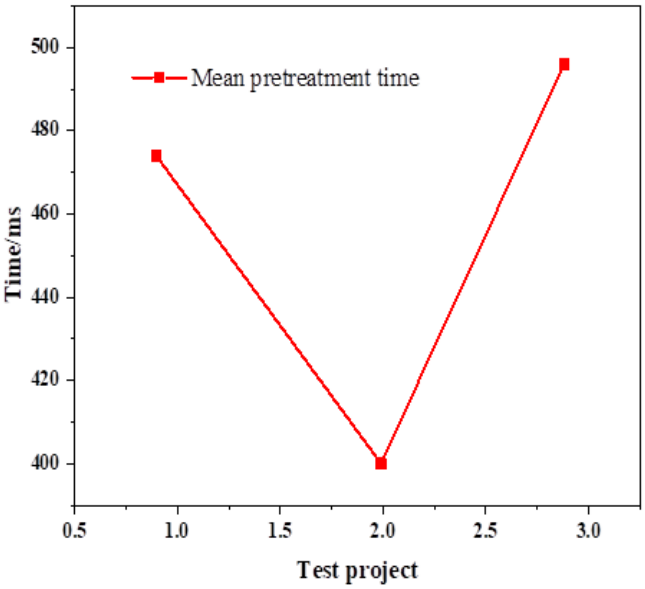


Figure 3: Test results of average response preprocessing time.

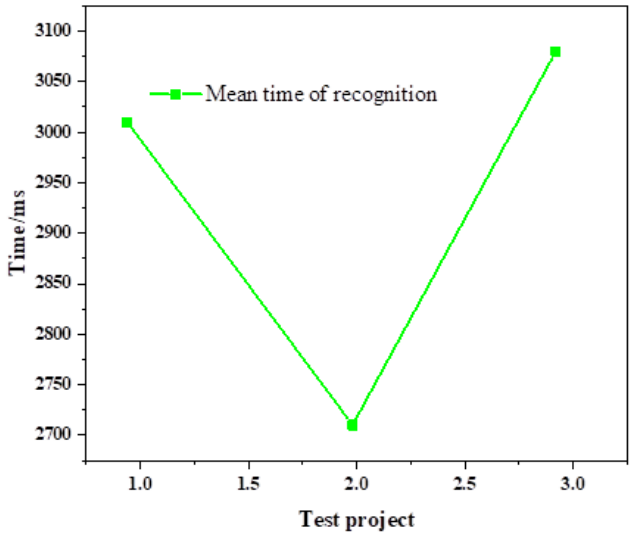


Figure 4: Average response identification time test results.

It can be seen from the test results that the average response time of text images in pure English is the fastest, the response time of text images in pure Chinese is the second, and the slowest is the mixed arrangement of Chinese and English. The analysis results show that because there are more Chinese features than English features and the matching time is longer, for the mixed arrangement of Chinese and English, multiple recognition is required to determine the text result, and the complex layout will also increase the character recognition time. Therefore, the text type and layout of the original image should not be too complex during character recognition, which can improve the accuracy of character recognition. The total average time of observation and recognition is about 3s, most of the mainstream commercial character recognition software is within 3s, and most of the software using cloud recognition is within 2s.

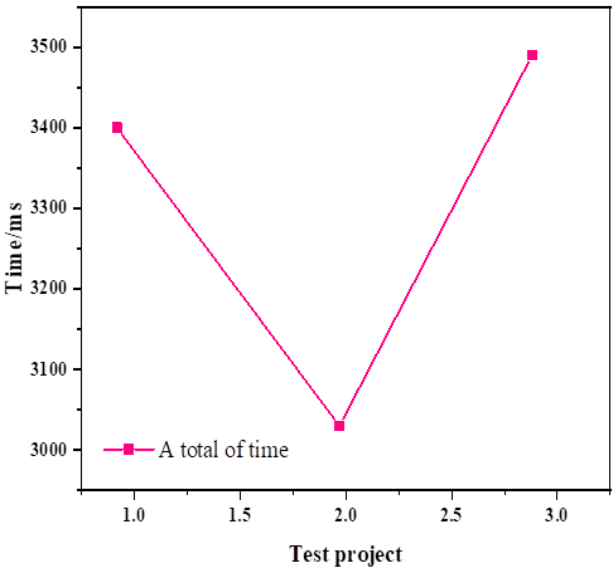


Figure 5: Total response time test results.

The recognition of the system is about 3s in many cases. Therefore, the system can basically achieve the recognition response time performance of the mainstream commercial software.

4.4 Identification accuracy test

After the function test and performance test, the recognition accuracy of OCR system needs to be tested. The recognition accuracy is an important index to evaluate the system. The recognition accuracy directly affects the user experience. The system mainly describes the recognition accuracy of the system through three parameters: correct recognition rate (A), error recognition rate (S) and recognition accuracy (P). Among them, recognition accuracy (P) is an important index to evaluate the recognition effect of a character recognition system. The higher the recognition accuracy, the better the recognition effect of the system. The identification accuracy is expressed by formula, see formula (13).

$$P = \frac{A}{A+S} * 100\% \tag{13}$$

This section also uses 150 pictures used in the performance test, 50 in Chinese, 50 in English and 50 in Chinese and 50 in English. Each picture contains about 50 characters. After many tests and statistical data

analysis, the test results are shown in Table 2 and it is graphically presented in Figure 6.

Table 2: Statistical results of recognition accuracy

	Total number of characters	Number of false identifications	Correct recognition rate A	Error recognition rate S	Recognition accuracy P
Chinese characters (English/Mathematics)	2467	343	86.09%	13.91%	86.09%
Symbols (English)	2345	262	88.82%	11.18%	88.82%
Mixed arrangement of Chinese and English	2656	423	84.07%	15.93%	84.07%

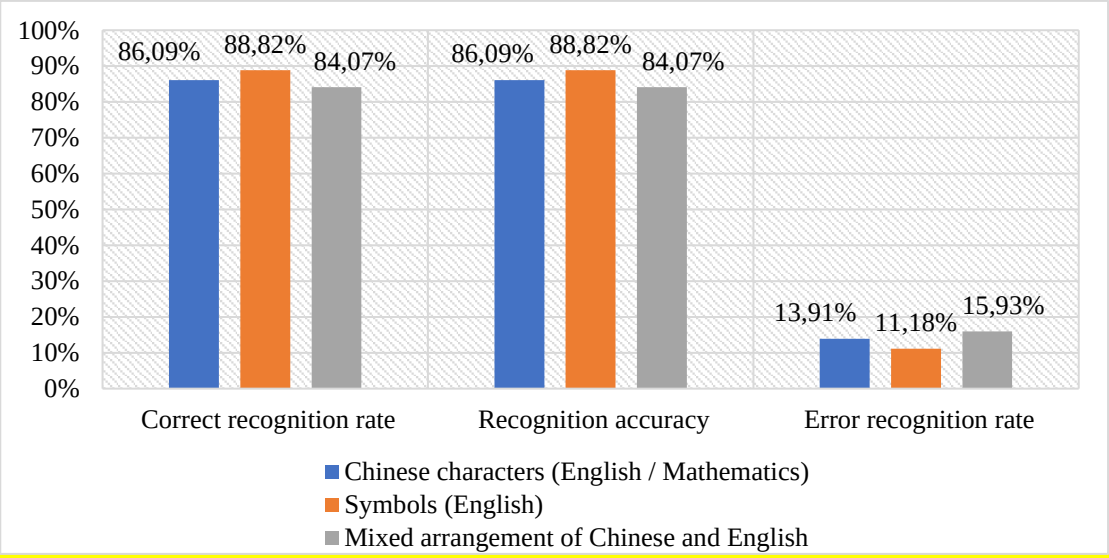


Figure 6: Graphical presentation of statistical results of recognition accuracy.

The experimental results show that the accuracy of character recognition is about 90% for pure English text images and close to 90% for pure Chinese text images. However, for the mixed arrangement of Chinese and English, the accuracy of character recognition is lower than that of pure Chinese and English. Nowadays, the accuracy of most commercial character recognition software is about more than 90%. Except for the mixed arrangement of Chinese and English, other products can basically achieve the accuracy of general commercial character recognition software. The product has certain practicability and can be applied to recognize text images taken in natural scenes in daily life.

4.5 System comprehensive evaluation

The optical character recognition system based on Android determines the system architecture and functional modules through user demand research and business analysis. The interface design part adopts the design style of material design. After the system test and user experience feedback on the UI effect of OCR system, the evaluation is mainly based on the scores of system function, response performance, accuracy and interface. The system application evaluation survey is shown in Table 3. According to the application evaluation table, the system has complete functional modules, high availability, neat and beautiful system interface and good operability. However, the recognition response performance and recognition accuracy score of the system are not very good. Therefore, in the later stage, we should deeply study the detailed technology of character recognition, break through the key technical points and improve the comprehensive ability of the software.

Table 3: System application evaluation

Evaluation content	Average score
Functional module	92
Response performance	84
Accuracy	87
Interface layout, aesthetic degree	93

5 Conclusion

This paper proposes the design and development of mobile terminal application based on Android, which mainly explains the implementation process from the implementation process of functional modules along with UI implementation effect and some implementation codes. For the UI implementation part, the design of this system is based on Google material design. This design style is a new set of interface design language invented by Google design engineers based on traditional excellent design principles and combined with rich creativity and science and technology, including visual, sports, interaction and other characteristics. The experimental tests are done on the character recognition system, and the experimental results are analyzed. The test shows that under ideal conditions, both Chinese and English recognition have good recognition accuracy. The average response time of the system is about 3s, which can basically meet the response time performance of mainstream commercial OCR software within 3S. However, compared with some cloud-based OCR software whose recognition time is generally less than 2s, there is still a certain gap between this software and cloud-based OCR software. With the gradual upgrading of Android mobile terminal hardware in the future and the further improvement of CPU and GPU processing. The system can quickly carry out image preprocessing and character recognition by relying on gradually powerful mobile terminal computing resources, and efficiently complete the character recognition task at the mobile terminal. However, the recognition response performance and recognition accuracy score of the system are not very good. Therefore, the future research directions should focus on detailed technology of character recognition, break through the key technical points and improve the comprehensive ability of the software.

References

[1] Kaur, S., & Dhindsa, K. S. (2020). Design and development of android based mobile application for specially abled people. Wireless Personal Communications, 111(4), 2353-2367. *https://doi.org/10.1007/s11277-019-06990-y*

[2] Zhang, W., Li, X., Xiong, N., & Vasilakos, A. V. (2016). Android platform-based individual privacy information protection system. *Personal and Ubiquitous Computing*, *20*(6), 875-884. *https://doi.org/10.1007/s00779-016-0966-0*

[3] Sun, Y., & Jiang, X. (2017). The design and application of english pronunciation training software based on android intelligent mobile phone platform. Revista de la Facultad de Ingenieria, 32(5), 756-765. *https://doi.org/10.1155/2021/5537101*

[4] Conti, R., Allotta, B., Meli, E., & Ridolfi, A. (2015). Development, design and validation of an assistive device for hand disabilities based on an innovative mechanism. Robotica, 35(4), 892-906. *https://doi.org/10.1017/S0263574715000879*

[5] Wen, C., & Zhang, J. (2014). Design of a microlecture mobile learning system based on smartphone and web platforms. *IEEE Transactions on Education*, *58*(3), 203-207. *10.1109/TE.2014.2363627*

[6] Hongyan, L. (2015). Design and realization of smart home terminal applications based on iot technology. International Journal of Smart Home, 9(8), 123-132. *http://dx.doi.org/10.14257/ijsh.2015.9.8.14*

[7] Tong, J. (2016). Design and implementation of music teaching platform in college based on android mobile technology. International Journal of Emerging Technologies in Learning (iJET), 11(5), 4. *https://www.learntechlib.org/p/173410/*

[8] Xu, H., Zhao, J., Li, Y. L., & Xu, C. Q. (2015). The development of shs-swths designing software based on windows and android mobile device platforms. Renewable Energy, 84(DEC.), 138-144. *https://doi.org/10.1016/j.renene.2015.06.058*

[9] Wang, Y., Zhou, J. T., & Song, X. (2018). A raas model based on emotion analysis and double labeling applied to mobile terminal. IEEE Access, PP(99), 1-1. *10.1109/ACCESS.2018.2880738*

[10] Feng, Y. (2020). Mobile terminal video image fuzzy feature extraction simulation based on surf virtual reality technology. IEEE Access, PP(99), 1-1. *10.1109/ACCESS.2020.3019070*

[11] Cai, Y., Wang, Z., He, P., Lu, Z., & Tang, J. (2017). The industrial network system design based on profinet1. Acta Technica CSAV (Ceskoslovensk Akademie Ved), 62(1), 21-31.

[12] Ameigeiras, P., Ramos-Munoz, J. J., Schumacher, L., Prados-Garzon, J., Navarro-Ortiz, J., & Lopez-Soler, J. M. (2015). Link-level access cloud architecture design based on sdn for 5g networks. IEEE Network, 29(2), 24-31. *10.1109/MNET.2015.7064899*

[13] Lin, T. C., Jun-Zhe, Y., Yu, C. S., & Liu, C. W. (2019). Development of a transmission network fault location platform based on cloud computing and synchrophasors. IEEE Transactions on Power Delivery, PP(99), 1-1. *10.1109/TPWRD.2019.2929290*

[14] Cao, C., Yang, Q., Li, J., Liu, X., & Yu, L. (2019). Design of water surface evaporation on-line detection device of pulse type evaporator. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering, 35(1), 106-113. *10.11975/j.issn.1002-6819.2019.01.013*

[15] Xin, C., Qiao, D., Hongjie, S., Chunhe, L., & Haikuan, Z. (2018, June). Design and implementation of debris search and rescue robot system based on internet of things. In *2018 International Conference on Smart Grid and Electrical Automation (ICSGEA)* (pp. 303-307). IEEE. *10.1109/ICSGEA.2018.00082*

[16] Zhang, Y. (2021, October). Application of

Computer Aided Design in the Major of Exhibition Design. In *2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture* (pp. 1117-1120). *https://doi.org/10.1145/3495018.3495347*

[17] Jiang, K., Li, J., Zhang, X., & Liu, S. (2015). The development and application of one thermal–hydraulic program based on ansys for design of ceramic breeder blanket of cfetr. Journal of Fusion Energy, 34(5), 1088-1093. *https://doi.org/10.1007/s10894-015-9923-6*

[18] Luo, M. (2014, September). The design of English listening learning system based on Android platform. In *2014 IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA)* (pp. 644-647). IEEE. *10.1109/WARTIA.2014.6976344*

[19] Al-Obaidi, A. S. M., Al-Qassar, A., Nasser, A. R., Alkhayyat, A., Humaidi, A. J., & Ibraheem, I. K. (2021). Embedded design and implementation of mobile robot for surveillance applications. *Indonesian Journal of Science and Technology*, *6*(2), 427-440. *https://doi.org/10.17509/ijost.v6i2.36275*

[20] Nie, J. (2015, December). Research on mobile learning platform construction in higher vocational colleges based on cloud computing. In *2015 11th International Conference on Computational Intelligence and Security (CIS)* (pp. 91-94). IEEE. *10.1109/CIS.2015.30*

[21] Sun, Z., Anbarasan, M., & Praveen Kumar, D. J. C. I. (2021). Design of online intelligent English teaching platform based on artificial intelligence techniques. *Computational Intelligence*, *37*(3), 1166-1180. *https://doi.org/10.1111/coin.12351*

[22] Yu, J., Du, S., & Wu, Y. (2020). Design and application of soil moisture content monitoring system based on cloud-native technology. Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering, 36(13), 165-172. *10.11975/j.issn.1002-6819.2020.13.020*

[23] Tahiroğlu, K., Tanaka, A., Parkinson, A., & Gibson, S. (2012, June). Designing musical interactions for mobile systems. In *Proceedings of the Designing Interactive Systems Conference* (pp. 807-808). *10.1145/2317956.2318084*

Detecting Temporal and Spatial Anomalies in Users' Activities for Security Provisioning in Computer Networks

Aleks Huč
University of Ljubljana, Faculty of Computer and Information Science, Ljubljana, Slovenia
aleks.huc@fri.uni-lj.si

Thesis Summary

Keywords: anomaly detection, incremental learning, unsupervised learning, clustering, adaptive windowing, profiling, network security, network flows

Received: June 1, 2023

The paper summarizes a Doctoral Thesis that focuses on two new approaches for detecting anomalies in computer networks based on network flows. The approaches use incremental hierarchical clustering algorithms and monitor changes in the data structures to detect anomalies. Both approaches achieved prediction performance comparable to the state-of-the-art supervised approaches (F1 score over 0.90), even when taking into account that our approaches see every data point only once and then discard it and they operate without the prerequisite learning phase with labeled data.

Povzetek: Članek povzema vsebino doktorske disertacije, v kateri se osredotočimo na dva nova pristopa za detekcijo anomalij v računalniških omrežjih. Pristopa temeljita na omrežnih tokovih, inkrementalnem hierarhičnem gručenju in spremljanju sprememb v podatkovnih strukturah z namenom detekcije anomalij. Oba pristopa dosežeta primerljivo stopnjo detekcije (mera F1 preko 0.90) v primerjavi z najnovejšimi nadzorovanimi metodami, tudi ko upoštevamo, da naša pristopa vidita vsak podatek le enkrat in ga nato pozabita ter delujeta brez predhodne faze učenja z označenimi podatki.

1 Introduction

The goal of computer network security is to provide a secure environment for a computer network, its resources, data in storage and transit and all its users [1]. Network security starts with intrusion detection, which is defined as a deliberate unauthorized attempt (successful or not) by an intruder to gain access to, manipulate or misuse a computer system or network [1]. Examples include Trojans, viruses, malware and denial of service, brute force and probe attacks.

Over the years of active development, two main categories of intrusion detection approaches have emerged: signature-based and anomaly-based [2]. Signature-based approaches detect intrusions on the basis of signature databases while anomaly-based approaches detect intrusions on the basis of deviations from normal activity models. Various anomaly detection approaches have already been proposed but have problems with today's dynamic computer networks with large volume and high velocity, variety and variability due to their use of supervised and batch learning. Newer methods have switched to unsupervised, incremental and adaptable methods to improve upon and augment traditional approaches and provide overall better anomaly detection.

This paper summarizes a Doctoral Thesis [3] that provides two new approaches for improving the current state-of-the-art anomaly detection using unsupervised, incremental, adaptable and hierarchical clustering.

2 PHICAD

PHICAD (Profile- and Hierarchical Incremental Clustering-based Anomaly Detection) is a single-layer, unsupervised and incremental algorithm that detects network activity anomalies in real-time. The input is a stream of chronologically ordered flows. The algorithm receives a new flow and sends it to the appropriate two profiles based on source and destination IP addresses. A profile models the incoming and outgoing activity of an individual network entity. The algorithm then extracts, transforms, and normalizes the features from the flow into a real-valued vector. The vector is then clustered inside the appropriate profile hierarchical clustering tree structure.

The anomalies are determined in the leaf nodes where, if the new vector is merged with the existing leaf, we track the distance between the new vector and the leaf, the leaf centroid changes and the leaf size; or if the new vector becomes a new leaf, we track the distance between the new leaf and the centroid of neighboring leaves. The predictions from all detection mechanisms are put into a short-term model that discards mechanisms that trigger too often and reports final predictions.

3 PHI2CAD

PHI2CAD (Profile- and Hierarchical Incremental Two-layer Clustering-based Anomaly Detection) builds upon our single-layer PHICAD with an additional second layer unsupervised and incremental clustering algorithm which detects anomalies in profiles and groups of profiles.

The input into our approach is again a stream of chronologically ordered flows. First, the flow is sent to the first layer where the PHICAD algorithm creates and updates profiles of network entities and detects network anomalies in each individual profile separately. For each flow, the PHICAD produces two updated profiles with predictions for possible anomalies, one for the source and one for the destination IP address, which are then sent to the PHI2CAD algorithm on the second layer.

PHI2CAD first checks for each updated profile if it has already been clustered into its tree data structure and if it has been, it checks if the updated profile is still inside the leaf or not. If it is still inside, we check for possible anomalies caused by the updated profile and produce possible anomaly predictions, by tracking the distance between the updated profile and the leaf, the leaf centroid changes, and the leaf size.

Otherwise, if an updated profile has not been clustered yet or it falls outside the leaf it has previously been clustered to, we cluster the updated profile into PHI2CAD tree data structure, while its previous version, if it exists, is removed from the tree. Finally, possible anomaly predictions are determined in the leaf to which the updated profile has been clustered. If the updated profile is merged with the existing leaf, we track the distance between the updated profile and the leaf, the leaf centroid changes and the leaf size; or if the updated profile becomes a new leaf, we track the distance between the new leaf and the centroid of the neighboring leaves. The predictions from all detection mechanisms are input into a short-term model that discards mechanisms which trigger too often and reports the final predictions.

4 Conclusion

The goal of this dissertation was to research if we can devise an anomaly detection approach with the following operational constraints: incremental execution, unsupervised learning, real-time response, ability to analyze large data sets, lightweight design and ability to adapt to changes over time, while still providing comparable performance to classic approaches and/or providing us with additional new insights.

We have evaluated our two approaches using a state-of-the-art data set CICIDS2017 [4] that comprises the most common network anomalies. To measure the predictive performance we used standard machine learning metrics such as precision, recall and F1 score and also the execution time against the supervised approaches. To further explain the achieved prediction performance we analyzed the influence of individual features on the predictions and performed sensitivity analysis of the main parameters.

Our approaches can successfully detect Denial of Service, Distributed Denial of Service, Port Scan and Web attacks when analyzing each anomaly separately and are also able to detect anomalies even when they analyze entire data sets with multiple types of anomalies. Performance is good where anomalous patterns clearly differ from the normal activity (F1 score over 0.90), however, they have problems detecting attacks that are presented with flows that are similar to normal flows or that are executed on higher layers of the network stack or are a part of packet payloads. But we have to be mindful of the diminishing importance of packet-payload analysis, due to the increasing use of packet-payload encryption. The results were also published in a peer-reviewed journal paper [5].

References

[1] Kizza, J. M. (2020) *Guide to computer network security*, Springer.

[2] Thakkar, A. and Lohiya, R. (2021) A survey on intrusion detection system: feature selection, model, performance measures, application perspective, challenges, and future research directions, *Artificial Intelligence Review*, Springer, pp. 1–111.

[3] Huč, A. (2022) *Detecting temporal and spatial anomalies in users' activities for security provisioning in computer networks*, doktorska disertacija, Ljubljana, https://repozitorij.uni-lj.si/IzpisGradiva.php?id=137562.

[4] Sharafaldin, I. and Lashkari, A. H. and Ghorbani, A. A. (2018) Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterizationy, *4th International Conference on Information Systems Security and Privacy (ICISSP)*, pp. 108–116.

[5] Huč, A. and Trček, D. (2021) Anomaly detection in IoT networks: From architectures to machine learning transparency, *IEEE Access*, IEEE, pp. 60607–60616.

Tradeoffs In Using Blockchain Technology For Security, Privacy, And Decentralization:theoretical And Empirical Perspectives

Aleksandar Tošić
Faculty of Mathematics, Natural Sciences and Information Technologies
University of Primorska
E-mail: aleksandar.tosic@upr.si

Thesis Summary

Keywords: Edge Computing, Blockchain, Container orchestration, Consensus mechanism

Received: June, 15, 2023

This paper is and extended abstract of the doctoral thesis [1]. It identifies four selected topics in which blockchain technology can have a positive or transformative effect on existing solutions. We propose new protocols, which change the current standards to add functionality, improve performance or overcome limitations of existing blockchain networks. Specifically, we focus on container orchestration on the edge using a unique blockchain protocol for security, verifiability, and trust.

Povzetek: Pričujoče delo je razširjen povzetek doktorske disertacije [1]. Delo predstavlja raziskavo, ki se osredotoča na uporabo tehnologije blockchain za izboljšanje obstoječih rešitev na štirih izbranih področjih. Poseben poudarek je na razvoju in implementaciji edinstvenega blockchain protokola za orkestracijo kontejnerjev na robu omrežja, ki zagotavlja varnost, preverljivost in zaupanje.

1 Introduction and problem statement

In recent years, cloud computing became a commonly used architecture for most applications. The shift of the geography of computation was incentivized by many factors ranging from ease of software maintenance [2], reliable quality of service(QoS), hardware flexibility, and cost (CapEx to OpEx) [3], etc However, with the expected growth of data generation and consumption and storage and service provisioning in cloud computing environments, the architecture is pushing network bandwidth requirements to the limit [4]. Edge computing in it's simplest form can be defined as an architecture in which computation is moved to the edge of the network in order to make use of the geographic proximity to decrease latency and improve bandwidth. This recent paradigm shift attempts to address the overly geographically-centralized cloud architecture. However, distributing services to the edge introduces new challenges such as resource allocation, service and application migration, trust, etc.. Blockchain technology may be used to address some of the issues. It can serve as a layer of trust between the system, and the end user by providing a verifiable and transparent ledger of the state of the system. To achieve this, a new protocol is required that would overcome the latency constraint, decentralized resource allocation, and real-time container migrations [5].

2 Methodology

We design, and develop a new blockchain protocol aimed at autonomous decentralized container orchestration suitable for edge devices. The proposed protocol uses verifiable delay functions(VDFs)[6] as the entropy source for secure randomness. Nodes participating in consensus compute a function $p = vdf(bh, bd)$ where p is the proof, bh is the SHA256 hash of the current block, and bd is the difficulty of the current block. We show that p is a sufficiently secure source of entropy for generating randomness. Moreover, given delay imposed on the nodes computing the VDF prevents malicious nodes to peek into the future. Using the shared seed, nodes are able to self-elect into consensus roles for each slot without communication overhead as shown in Figure 1.

3 Evaluation methodology and results

We performed extensive testing of our reference implementation simulating networks as large as 1000 nodes. The telemetry obtained from logging the state of all nodes shows that the protocol is scalable, and can efficiently converge towards average resource utilization of the entire network by performing real-time container migrations between nodes using Checkpoint/Restore In Userspace (CRIU). Moreover, our results show that using CRIU significantly improves the performance making our protocol viable in practice.

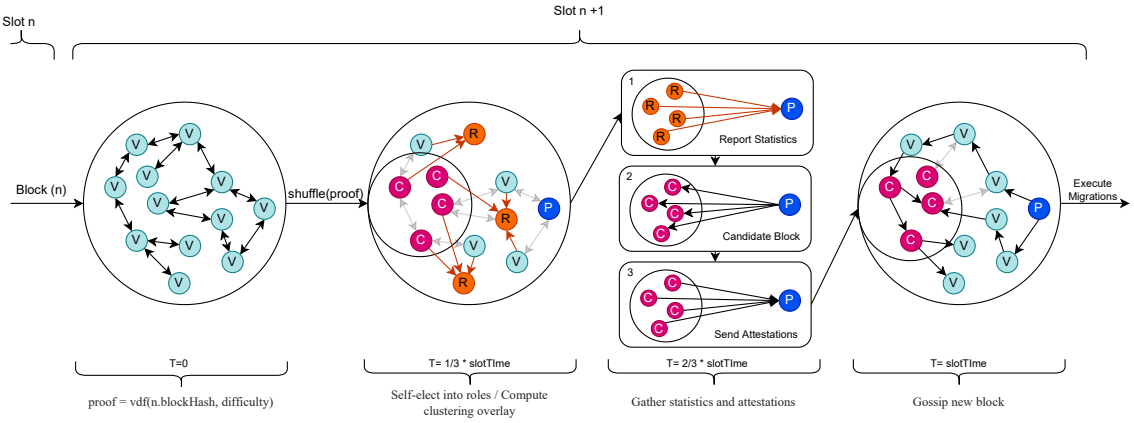


Figure 1: Role based consensus mechanism using VDFs as an entropy source

4 Discussion and further work

Our results showcase the feasibility of the proposed protocol for large networks of edge devices with limited compute resources. However, decentralized networks must address Byzantine behaviour of nodes. To secure the protocol against malicious actors the protocol must secure containerized application and guarantee the execution. Existing solutions such as Intel SGX are not generic and inherently impose hardware restrictions on the protocol. Research should focus overcoming specific hardware implementations of trusted computation.

References

[1] A. Tošić, "Empirična študija uporabe tehnologije veriženja blokov v obstoječih sistemih in arhitekturah: doktorska disertacija," Ph.D. dissertation, Univerza na Primorskem, Fakulteta za matematiko, naravoslovje in …, 2022.

[2] E. Bayrak, J. Conley, and S. Wilkie, "The economics of cloud computing," 2011.

[3] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica *et al.*, "A view of cloud computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.

[4] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, 2016.

[5] A. Tošić, J. Vičič, M. Burnard, and M. Mrissa, "A blockchain protocol for real-time application migration on the edge," *Sensors*, vol. 23, no. 9, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/9/4448

[6] D. Boneh, J. Bonneau, B. Bünz, and B. Fisch, "Verifiable delay functions," in *Advances in Cryptology–CRYPTO 2018: 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19–23, 2018, Proceedings, Part I*. Springer, 2018, pp. 757–788.

JOŽEF STEFAN INSTITUTE

Jožef Stefan (1835-1893) was one of the most prominent physicists of the 19th century. Born to Slovene parents, he obtained his Ph.D. at Vienna University, where he was later Director of the Physics Institute, Vice-President of the Vienna Academy of Sciences and a member of several sci- entific institutions in Europe. Stefan explored many areas in hydrodynamics, optics, acoustics, electricity, magnetism and the kinetic theory of gases. Among other things, he originated the law that the total radiation from a black body is proportional to the 4th power of its absolute tem- perature, known as the Stefan–Boltzmann law.

The Jožef Stefan Institute (JSI) is the leading independent scientific research institution in Slovenia, covering a broad spectrum of fundamental and applied research in the fields of physics, chemistry and biochemistry, electronics and information science, nuclear science technology, en- ergy research and environmental science.

The Jožef Stefan Institute (JSI) is a research organisation for pure and applied research in the natural sciences and technology. Both are closely interconnected in research de- partments composed of different task teams. Emphasis in basic research is given to the development and education of young scientists, while applied research and development serve for the transfer of advanced knowledge, contributing to the development of the national economy and society in general.

At present the Institute, with a total of about 900 staff, has 700 researchers, about 250 of whom are postgraduates, around 500 of whom have doctorates (Ph.D.), and around 200 of whom have permanent professorships or temporary teaching assignments at the Universities.

In view of its activities and status, the JSI plays the role of a national institute, complementing the role of the uni- versities and bridging the gap between basic science and applications.

Research at the JSI includes the following major fields: physics; chemistry; electronics, informatics and computer sciences; biochemistry; ecology; reactor technology; ap- plied mathematics. Most of the activities are more or less closely connected to information sciences, in particu- lar computer sciences, artificial intelligence, language and speech technologies, computer-aided design, computer architectures, biocybernetics and robotics, computer automa- tion and control, professional electronics, digital communi- cations and networks, and applied mathematics.

The Institute is located in Ljubljana, the capital of the in dependent state of S**love**nia (or S♡nia). The capital today isconsidered a crossroad bet between East, West and Mediter-ranean Europe, offering excellent productive capabilities and solid business opportunities, with strong international connections. Ljubljana is connected to important centers such as Prague, Budapest, Vienna, Zagreb, Milan, Rome, Monaco, Nice, Bern and Munich, all within a radius of 600 km.

From the Jožef Stefan Institute, the Technology park "Ljubljana" has been proposed as part of the national strat- egy for technological development to foster synergies be- tween research and industry, to promote joint ventures be- tween university bodies, research institutes and innovative industry, to act as an incubator for high-tech initiatives and to accelerate the development cycle of innovative products.

Part of the Institute was reorganized into several high-tech units supported by and connected within the Technology park at the Jožef Stefan Institute, established as the beginning of a regional Technology park "Ljubljana". The project was developed at a particularly historical moment, characterized by the process of state reorganisation, privatisation and private initiative. The national Technology Park is a shareholding company hosting an independent venture-capital institution.

The promoters and operational entities of the project are the Republic of Slovenia, Ministry of Higher Education, Science and Technology and the Jožef Stefan Institute. The framework of the operation also includes the University of Ljubljana, the National Institute of Chemistry, the Institute for Electronics and Vacuum Technology and the Institute for Materials and Construction Research among others. In addition, the project is supported by the Ministry of the Economy, the National Chamber of Economy and the City of Ljubljana.

Jožef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
Tel.:+386 1 4773 900, Fax.:+386 1 251 93 85
WWW: http://www.ijs.si
E-mail: matjaz.gams@ijs.si
Public relations: Polona Strnad

INFORMATICA

AN INTERNATIONAL JOURNAL OF COMPUTING AND INFORMATICS

INVITATION, COOPERATION

Submissions and Refereeing

Please register as an author and submit a manuscript at: http://www.informatica.si. At least two referees outside the author's country will examine it, and they are invited to make as many remarks as possible from typing errors to global philosoph- ical disagreements. The chosen editor will send the author the obtained reviews. If the paper is accepted, the editor will also send an email to the managing editor. The executive board will inform the author that the paper has been accepted, and the author will send the paper to the managing editor. The paper will be pub- lished within one year of receipt of email with the text in Infor- matica MS Word format or Informatica LaTeX format and figures in .eps format. Style and examples of papers can be obtained from http://www.informatica.si. Opinions, news, calls for conferences, calls for papers, etc. should be sent directly to the managing edi- tor.

SUBSCRIPTION

Please, complete the order form and send it to Dr. Drago Torkar, Informatica, Institut Jožef Stefan, Jamova 39, 1000 Ljubljana, Slovenia. E-mail: drago.torkar@ijs.si

Since 1977, Informatica has been a major Slovenian scientific journal of computing and informatics, including telecommunications, automation and other related areas. In its 16th year (more than twentyeight years ago) it became truly international, although it still remains connected to Central Europe. The basic aim of Informatica is to impose intellectual values (science, engineering) in a distributed organisation.

Informatica is a journal primarily covering intelligent systems in the European computer science, informatics and cognitive community; scientific and educational as well as technical, commercial and industrial. Its basic aim is to enhance communications between different European structures on the basis of equal rights and international refereeing. It publishes scientific papers ac- cepted by at least two referees outside the author's country. In ad- dition, it contains information about conferences, opinions, criti- cal examinations of existing publications and news. Finally, major practical achievements and innovations in the computer and infor- mation industry are presented through commercial publications as well as through independent evaluations.

Editing and refereeing are distributed. Each editor can conduct the refereeing process by appointing two new referees or referees from the Board of Referees or Editorial Board. Referees should not be from the author's country. If new referees are appointed, their names will appear in the Refereeing Board.

Informatica web edition is free of charge and accessible at http://www.informatica.si.
Informatica print edition is free of charge for major scientific, ed- ucational and governmental institutions. Others should subscribe.

Informatica

An International Journal of Computing and Informatics

Web edition of Informatica may be accessed at: http://www.informatica.si.

Subscription Information Informatica (ISSN 0350-5596) is published four times a year in Spring, Summer, Autumn, and Winter (4 issues per year) by the Slovene Society Informatika, Litostrojska cesta 54, 1000 Ljubljana, Slovenia.
The subscription rate for 2022 (Volume 46) is
– 60 EUR for institutions,
– 30 EUR for individuals, and
– 15 EUR for students
Claims for missing issues will be honored free of charge within six months after the publication date of the issue.

Typesetting: Blaž Mahnič, Gašper Slapničar; gasper.slapnicar@ijs.si
Printing: ABO grafika d.o.o., Ob železnici 16, 1000 Ljubljana.

Orders may be placed by email (drago.torkar@ijs.si), telephone (+386 1 477 3900) or fax (+386 1 251 93 85). The payment should be made to our bank account no.: 02083-0013014662 at NLB d.d., 1520 Ljubljana, Trg republike 2, Slovenija, IBAN no.: SI56020830013014662, SWIFT Code: LJBASI2X.

Informatica is published by Slovene Society Informatika (president Niko Schlamberger) in cooperation with the following societies (and contact persons):
Slovene Society for Pattern Recognition (Vitomir Štruc)
Slovenian Artificial Intelligence Society (Sašo Džeroski)
Cognitive Science Society (Olga Markič)
Slovenian Society of Mathematicians, Physicists and Astronomers (Dragan Mihailović)
Automatic Control Society of Slovenia (Giovanni Godena)
Slovenian Association of Technical and Natural Sciences / Engineering Academy of Slovenia (Mark Pleško)
ACM Slovenia (Ljupčo Todorovski)

Informatica is financially supported by the Slovenian research agency from the Call for co-financing of scientific periodical publications.

Informatica is surveyed by: ACM Digital Library, Citeseer, COBISS, Compendex, Computer & Information Systems Abstracts, Computer Database, Computer Science Index, Current Mathematical Publications, DBLP Computer Science Bibliography, Directory of Open Access Journals, InfoTrac OneFile, Inspec, Linguistic and Language Behaviour Abstracts, Mathematical Reviews, MatSciNet, MatSci on SilverPlatter, Scopus, Zentralblatt Math

Volume 47 Number 2 June 2023 **ISSN 0350-5596**

Informatica

An International Journal of Computing and Informatics

Informatica **47** (2023) Number 2, pp. 143–300