

A Study of Fairness of Information Distribution and Utilization in a Mobile Agent-Based Adaptive Information Service System

Iftikhar Ahmed and Mohammad Jafar Sadiq,
 Department of Electrical & Computer Engineering
 Faculty of Engineering, International Islamic University Malaysia
 E-mail: a.iftikhar@iiu.edu.my

Keywords: mobile agents, information systems, intelligent systems

Received: July 19, 2005

The information handling capacity of wide area networks is ever increasing and becoming dynamic, thus posing a challenge for their optimum utilization. In such a dynamic environment it is very difficult to maintain a coherent picture of the information services. The current information services on the internet therefore need to be optimized. A novel technique proposed recently employs autonomous mobile agents in faded information field architecture (FIF) to facilitate information provision and servicing on wide area networks. The architecture is decentralized in nature whereby both the service provider and information seeker communicate through mobile agents. However, the degree of fairness of information distribution and access is still a problem to be addressed. This paper investigates the issue of fairness among information users for access to information in a networked environment employing mobile agent technology. A FIF was simulated employing a number of algorithms characterizing its behaviour with respect to fairness of information distribution by various service providers in the network. Simulations results indicate that in a given information system, both the service providers and the information seekers may be given equal access to the distribution and utilization of information provided a number of suitable algorithms are employed depending on the nature of information service available on the wide area network.

Povzetek: Narejena je analiza porazdelitve v agentnem informacijskem storitvenem sistemu.

1 Introduction

Information in today's business environment has become the prime ingredient of a successful business enterprise. The rapid advances in Information and Communication Technologies (ICT) have given new impetus to the way business is conducted on the internet with information being the key element of a successful business enterprise. The traditional paradigms of business, learning and other essential services are undergoing an evolution. The internet is a huge data repository that is expanding without a central authority. The number of worldwide internet users is predicted to exceed one billion by the end of 2005 [1]. Moreover 300 terabytes of information is published online every year [2]. As the performance demands of the internet and its usage increase exponentially, the emphasis is shifting from platform centric computing to network centric computing. Information systems have to be designed that are distributed, dynamic and have high assurance and fault tolerance to meet the heterogeneous demands of users. As ICT advances, the dynamics of e-commerce tend to be more data intensive and complex. The traditional business to customer paradigm is encompassing the business to business transactions as well. Companies have to comprehend the trends and demands of the users quickly in order to survive in the competitive

environment of today. The expectations of users and customers have soared high and they seek a flawless any time any where service syndrome.

The traditional information services based on the client-server model therefore cannot cope with the rising demand placed by the complexities of data intensive heterogeneous computing. It is therefore imperative to employ new techniques for optimum and accurate information retrieval on the web. Various techniques being investigated for this purpose [3-5] mostly tend to improve upon the current method of information retrieval, i.e., search engines Information fading based on a demand-oriented service architecture is one technique that may bring about a conceptual change in the information search method in vogue [6], in addition to Semantic web [7]. The architecture balances the cost of information provision and utilization on the network by employing push and pull mobile agents to service user requests and conducts the business of information provision and utilization on the network. In the faded information field (FIF) the service provider distributes the most demanded and popular information contents closer to its vicinity on various nodes in the network. The volume of distributed information on a node is inversely proportional to the distance of information storing node

from the service provider. Thus the FIF permits the fairness of information distribution to unspecified users with equal access time.

A number of attributes of FIF architecture have been investigated recently [8-11]. This paper will analyse the fairness of information access to various users in a FIF information system. The paper is organized as follows; Information service system performance requirements are discussed in section 2, followed by a description of FIF architecture. Mobile agents will be surveyed in section 4. The attribute of fairness of information distribution and utilization will be discussed in section 4. The paper will be concluded in section 5.

2 Information service system performance requirements

The wide area network services are gradually creeping into our daily lives with the number of users of these services constantly on the rise. The information systems now involve a constantly changing environment where stringent performance demands are placed on the network. The businesses are taking advantage of the wide area network facility to offer bargains and put various items on sales and special promotions resulting in the users having much more flexibility and choice available making the process of information provision and utilization a complex matter to deal with on the network. It is thus imperative for an information system to respond to both user needs and service provider (SP) requirements to respond in time to these needs in a rapidly changing environment. The main desirable attributes of such a system therefore should be flexibility, reliability and quick reaction time. Only a system with these properties can successfully satisfy both users and SPs in a dynamic network environment.

The traditional data retrieval methodologies focused on the optimization of digging in huge data base to satisfy a search request. However, the FIF employs the concept of autonomous provision and utilization of information based on mobile agents. The amount of information to be faded away from SP is a function of network conditions like congestion, popularity of information contents and any other criteria programmed into mobile agents [6].

3 The FIF architecture

The information in a FIF is distributed on various nodes in the system. The information pattern in the field is analogous to the electromagnetic radiation from an antenna, high in intensity near the antenna and low in intensity away from the antenna. The SP fades away the information contents to the nodes in the field with the richness in information contents inversely proportional to the distance of the node from the SP through Push mobile agents (MAs).

The user on the other hand sends out Pull MAs seeking the desired information [6, 12]. The general schematic of FIF architecture is depicted in figure 1.

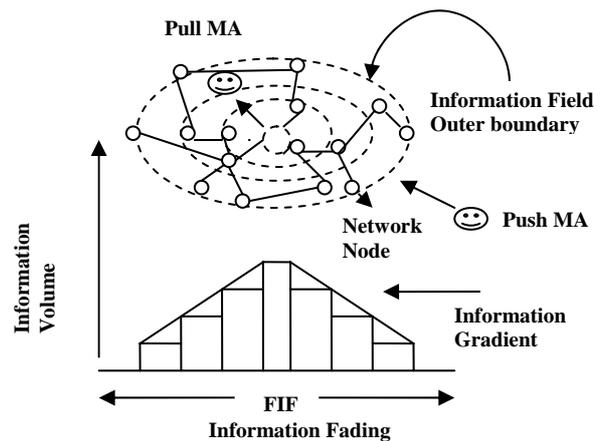


Figure 1: The layout of FIF architecture.

3.1 FIF System Components

The system essentially consists of logically connected nodes through which users and service providers correspond. Mobile agents are used by both parties to acquire and provide information respectively, under evolving/changing situation. The mobile agents (MA) generated by service providers are termed as push mobile agents (Push MAs). Push MAs carry out the function of autonomous coordination and negotiation with other nodes for information fading according to the network traffic status and the level of importance attached to the particular information contents.

The level of importance of particular information content is based on its popularity, determined from a high hit rate of query by the information seeker on the network. The pull mobile agents (Pull MAs) are generated by users and they autonomously navigate in search of the required information on the network nodes in a step-by-step fashion. Once the required information is located, these agents report back to the information-seeking source. The push and pull MAs have no direct correspondence with each other.

The third important subsystem of a FIF is the node itself. It is a platform for both storage of information and program execution. It monitors the local information-based system conditions and autonomously makes decisions for allocation requests by the SP. It determines the upper directions to the information service in addition to response times on its upper directions. The upper directions or stages of information eventually point to the SP node whereas the lower directions to the information service lead to the outer nodes away from the SP. Each subsystem is autonomous in terms of control to execute its operations and coordination with other nodes under evolving network conditions.

3.2 Communication Format in the FIF

The conventional communications techniques cannot cope with the evolving conditions in a heterogeneous network environment where the state of nodes, the status of the SPs and the stability of connections are highly unpredictable as the user demand to access the

information is dynamic in nature. FIF therefore employs a different technique referred to as content Code communication technique [6]. The information contents in a FIF are uniquely defined by their content codes (CCs). These are further elaborately specified by their characteristic codes (CHs). Push MAs are sent out in the FIF by SPs specifying Content Codes (CCs) of information to the nodes using the message format as depicted in the FIF communication format example in Figure 3. The information about a university is structured according to the degree of importance in this example. For instance, SP1 specifies the university name followed by CH₁ indicating the location of university with respect to a particular country and city. CH₂ depicts the programs offered by the university and so on.

The message format components thus lead to the breakdown of the principal source of information available on the SP into its uniquely defined characteristic codes (CHs). Push MAs are sent out in the FIF by SPs specifying Content Codes (CCs) of information to the nodes using the message format as shown in Figure 3. The selection of information storage/allocation is autonomously carried out by the nodes based on CHs. Similarly the Pull MAs sent out by the users search for the required information based on CHs.

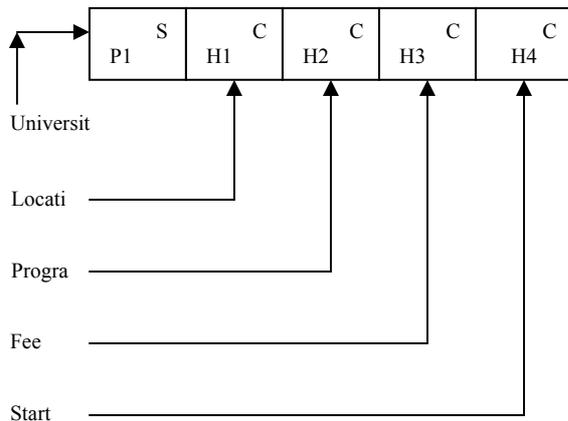


Figure 2: The message format in FIF architecture.

4 A survey of mobile agents

A distributed environment is most suitable for the employment of mobile agents. The details of such applications in a distributed environment can be found elsewhere [13-15]. Only, a brief introduction of mobile agents and their role in a network based information system will be discussed in this paper. The term mobile agent is often context-dependent and has two separate and distinct concepts: mobility and agency. The term agency implies having the same characteristics as that of an agent. These are self-contained and identifiable computer programs that can move within the network from node to node and act on behalf of a user or other entity. These can halt execution from a host without human interruption [16].

The current distributed network environment is based on the traditional client-server paradigm. In the case of mobile agents employed in a network, the service provision and utilization can be distributed in nature and is dynamically configured according to the changing network performance metrics like congestion and user demand for service provision [7]. The two network environments are depicted in Figure 1 both for client-server and mobile agents.

Mobile agents are typically suited to applications requiring structuring and coordinating wide area network and distributed services involving a large number of remote real time interactions [17].

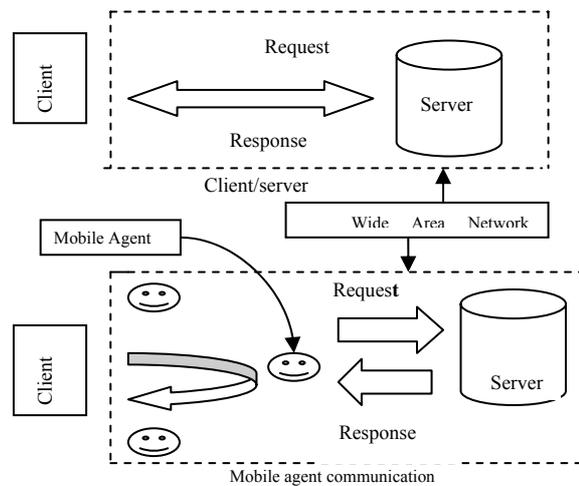


Figure 3: A mobile agent is a network aware program that can conserve bandwidth in a complex adaptive information system environment.

5 Fairness of information distribution and utilization simulation

The purpose of Information fading in a network is twofold:

- *Congestion Reduction* – The server hosting a popular site is highly likely to get congested. The information provider could employ additional machines (child nodes) to which the original server would distribute or fade its information contents. A pull MA entering the network of the parent server could then access any of the child nodes, which would service the agent’s request for the desired information and therefore reduce the service request traffic load on the parent server (PS) in the network.
- *Fairness of Access* – Service providers (whether commercial or non-commercial) would like their information contents to be easily accessible by the users on the network. The established and renowned service providers stand out compared to new competitors as search engines already index their search details. Not only the new

service providers suffer a long delay in getting their nomenclature indexed by a search engine, but they are the tail enders on the search engine index. This particular scheme of indexing by search engine creates disparity among Service providers on the network. Faded information field architecture therefore, can aid in creating fairness among various contenders of information provision and utilization on the network.

For congestion reduction, the service provider must use its own resources to create the field size since the congestion reduction benefits only the service provider. However, in a typical business application like advertising for example, each information provider cannot be expected to have the resources to create a field for itself alone. It is therefore possible that each server participates in the faded information field as being part of the field in addition to being a source of information. Thus, no one is the owner of the faded information field resources, and there is neither prohibition nor delay in becoming a member of the field.

For the sake of *localization* (prioritising a nearby service provider over a distant one), information should be faded in a geographic area as close as possible around the server. The faded information distance away from the SP was approximated using propagation delays in this simulation.

For the sake of *fairness*, the field size of all SPs should be equal. This is taken to be the number of nodes that constitute the field. Since each SP is autonomous in creating the field, there has to be an algorithm to ensure that each SP creates fields of the same size. The following are possible algorithms for an SP to autonomously determine its faded information field given that all SPs are aware of the nodes available for them to create FIFs. (*In essence, these algorithms compute the FIF size with the given variables as described next*):

- *Sorting*. Each SP orders the list of available FIF nodes into a list sorted by delay (as a substitute for physical distance). It can thus determine how “far” it needs to send its push-MAs to create the standard field size, and imposes a travel restriction beyond the envisaged perimeter of its field. The push-MAs will be required to keep updating the field until they reach the travel restriction set by the SP (which can be in the form of travel time). Alternatively the push-MAs can be multicast to the required nodes in the field. The sorting technique is inherently resource intensive and it must be re-employed to account for cost update in the WAN as and when it occurs. However, it is expected that all SPs will have the same number of nodes in the field, since each one selects the same number of nodes from the sorted list.
- *Step Size*. The SP determines the closest and furthest distances from the SP to the nodes, and

then divides that distance by the number of nodes to get the average ‘hop’ distance that a push-MA might take to travel from one node to the next. The inter-nodal step distance D is computed by using the following equation:

$$D = (\max(d) - \min(d)) / (N - 1) * (R / 100 * N) + \min(d)$$

Where D = step distance

d = distance from SP to node

N = number of nodes

R = required field size as a percent of total available nodes. The approximate distance that the SP envisages its push-MAs to cover in its FIF and reach a particular node is determined by the product of step distance and the n_{th} node ($D \times N$).

The SP will now send push-MAs that traverse the required distance. Alternatively, the server would determine which nodes fall in the required distance and multicast to those nodes. The step size must be recomputed every time there is an update in the system, but the major computation is finding the maximum and minimum distances, which is much faster than sorting.

- *Averaged Step Size*. In order to account for skewing of SPs’ distances, an average distance is computed for each server that is used to determine the step size instead. In this method, the determination of maximum and minimum distances is replaced by finding the total distance to all the nodes. The algorithm was simulated using the following equation:

$$D = (\sum d / (N - 1) - \min(d)) / (N - 1) * 2 * (R / 100.0 * N) + \min(d)$$

Where D represents the inter-nodal distance, N is the total number of nodes in the FIF and d is the distance of the node currently being visited by the MA from SP.

- *Modified Averaged Step Size*. The averaged step size algorithm still gets affected by the skewing of nodes. The algorithm was modified by computing the average step size of all the costs excluding the minimum and maximum costs of system information updates. In this method, the total has to be computed and the maximum and minimum has to be determined. The maximum and minimum are subtracted from the total. The average step size for push MA visits is computed using the following equation given below. The variables were incremented as the simulations iterations were performed to compute the step distance D of push MA visit to determine the FIF size.

$$D = ((\sum d - N_{min} * \min(d) - N_{max} * \max(d)) / (N - 1 - N_{min} - N_{max}) - \min_2(d)) / (N - 1 - N_{min} - N_{max}) * 2 * (R/100.0 * N - N_{min}) + \min_2(d)$$

Where

N_{min} = number of nodes at the minimum distance

N_{max} = number of nodes at the maximum distance and $min_2(d)$ = distance of the second closest node

The performance of each algorithm was tested by simulation using a wide area network with randomly distributed servers connected to randomly distributed gateways. The simulation involved 50 servers and 200 routers/gateways. Dijkstra’s algorithm [18] was used to determine the server-to-server costs. Under a standard policy each server was allowed to maintain a field size of 10 nodes. Figure 4 demonstrates the field size distribution for this wide area network.

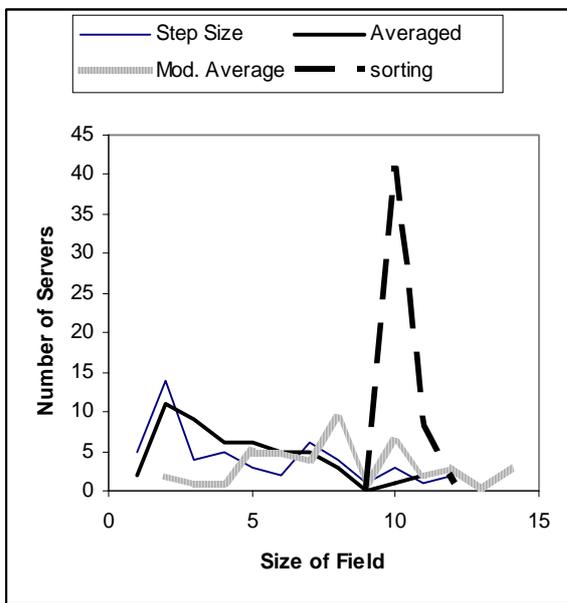


Figure 4: Field size distribution versus the number of nodes in the network.

It was observed that the sorting algorithm gave the closest approximation to the required distribution. However, the step size method failed to give the required distribution when it was required to provide a field size of 10 nodes. The averaged step size method fared a little better. It gave better field distribution since the distances are skewed to the higher end thus diminishing the field size below the required size. However, the modified average method gave a very good average of 10.8 nodes per faded information field but the servers did not all have equal numbers of nodes in their fields in this case. In order to account for skewing, a multiplier was introduced into the algorithms for step size and averaged step size. The step size algorithm was set to produce a field size of 15 nodes and the averaged step size algorithm a field size of 16 nodes. The results are shown in figure 5.

It can be seen that the field sizes are better distributed with no processing overhead. The averaged step size results finally show a significant improvement over the

normal step size results, as expected. In the second simulation run, the sorting algorithm had an average field size of 10.2 nodes per information provider, the step size algorithm had 10.3, the averaged step size had 10.5 and the modified average had 10.8 nodes respectively. Also, the averaged step size algorithm gave a fair distribution of field size, with most of the servers within a certain range.

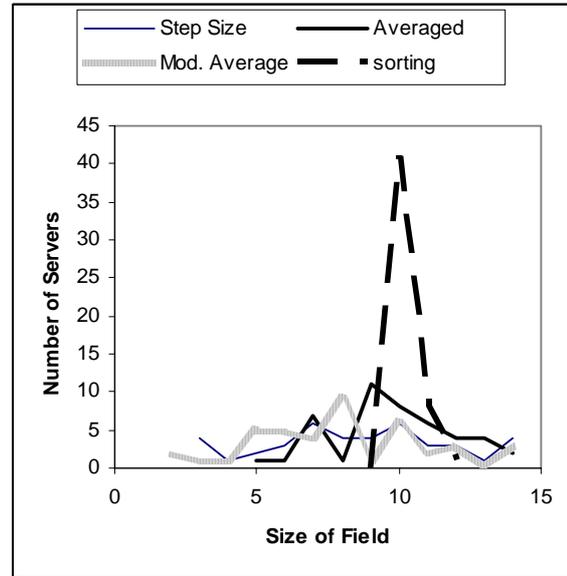


Figure 5: Compensation for skewing.

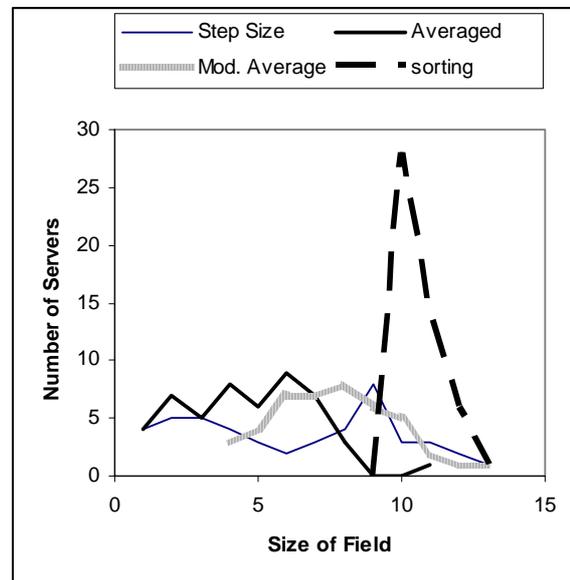


Figure 6: Field size without a multiplier.

The simulation was repeated using 150 routers instead of 200 to determine the effect of network size on the field size. The number of servers was kept the same. Figure 6 shows the field size without any multiplier and figure 7 depicts the field size with a multiplier. Again, the sorting and modified average algorithms both produced the required field size, whereas the other two algorithms did not. The step size algorithm this time needed only a

multiplier of 1.2 to give an acceptable average, but the average step size algorithm still needed a multiplier of 1.6. The latter produced a decent distribution of field size. Further simulation is required to determine the effects of network size on the above algorithms, as well as the effect of required field size. The preliminary simulations demonstrated that the sorting algorithm gives the best distribution of field size. Without a multiplier, the modified average algorithm gives the best performance, whereas, with a multiplier, the averaged step size algorithm gives the best distribution among the three non-sorting algorithms.

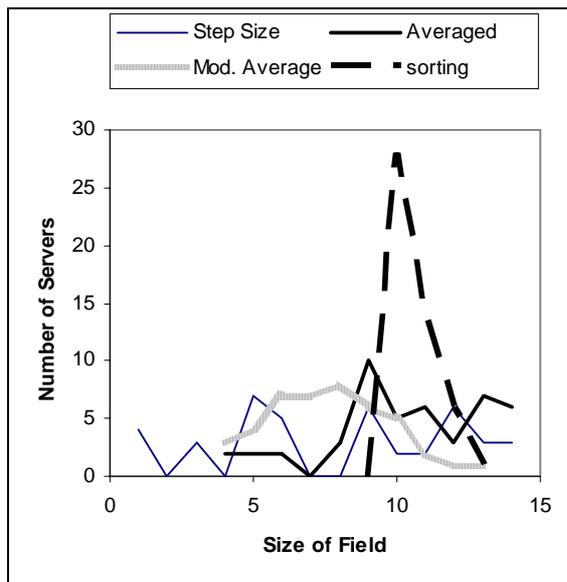


Figure 7: Field size with a multiplier.

5 Conclusion

Faded information field architecture has been reviewed in addition to a survey of mobile agents in this paper. A faded information field was simulated employing various algorithms to characterize its behaviour with respect to fairness of information distribution by various service providers in the network. The simulation provided a comparative analysis of algorithms to achieve this goal in addition to discovering a need to introduce a multiplicative factor to fine tune the performance of the algorithms evaluated for the purpose. It was discovered that without a multiplier, the modified average algorithm gave the best performance. But with a multiplier, the averaged step size algorithm turned out to be the best in distributing the information fairly among the three non-sorting algorithms. These results indicate that the faded information technique can be used to provide equal access rights to the information seekers and equal information distribution rights to the service providers by using the appropriate algorithm suited to the system requirement. Fairness of information provision and utilization can then be ensured on the network in addition to high reliability of information contents being updated regularly by the service provider.

Acknowledgement

This work was supported by the Research Centre International Islamic University Malaysia.

References

- [1] L. G. Roberts, "Beyond Moore's Law: Internet Growth Trends," *IEEE Computer*, Vol. 33, No. 1, p.117, 2000.
- [2] P. Lyman and H. R. Varian, "How much Information" *Journal of Electronics Publishing*, Vol. 6, Issue 2, December 2000.
- [3] Pavel Pereira Calado, Berthier Ribeiro-Neto, "An information Retrieval Approach for Approximate Queries," *IEEE Transactions on Knowledge and Data Engineering*, Jan-Feb 2003, Vol. 15, Issue. 1, p 236-239.
- [4] Kalervo Jarvelin, Juana Kekalainen, "Cumulated Gain-based Evaluation for IR Techniques," *ACM Transactions on Information Systems*, Oct 2002, Vol. 20, Issue.4, p. 422-447.
- [5] Caroline M. Eastman; Bernard J. Jansen, "Coverage, Relevance, and Ranking: the impact of query operators on Web search engine results," *ACM Transactions on Information Systems*, Oct 2003, Vol, 21, Issue 4, 383-412.
- [6] H. F. Ahmad, H. Arfaoui and K. Mori, "Autonomous Information Fading by Mobile Agents for improving User's Access Time and Fault Tolerance," *Proceedings, 5th International Symposium on Autonomous Decentralized Systems*, 2001. P.279-283.
- [7] Patrick Thibodeau, "The Web's Next Leap," *Computer world*, 21 April, 2003.
- [8] I. Ahmed, J Sadiq, "Characterization of Information Distribution Fairness in Mobile Agent-Based Adaptive Information System" *Proceedings, Sixth International Conference on Web Integration and Services*, September 27-29 2004, Jakarta, Indonesia, p. 761-770.
- [9] I Ahmed, J. Sadiq, "Nodal Congestion Analysis of an Autonomous Mobile Agent-Based Distributed Information Service System for Efficient Utilization of Information Contents," *International Journal of Information Technology*, 2004, Vol.10, No.1, p.64-70.
- [10] I. Ahmed, J Sadiq, "A Novel Autonomous Architecture for the Integration of Information and control in an Automated Manufacturing System," *Proceedings, International Conference on Intelligent Agents, Web Technologies and Internet Commerce*, 12-14 July 2004, Gold Coast Australia, p. 260-267.
- [11] I Ahmed, "Balancing the user demand and information provision in a Web-based linguistic learning system employing Faded Information Field Architecture," *Proceedings, ASIA CALL International Conference on Information Technology and Linguistic Learning*, Bangkok, Thailand, 3-5 December 2003, p. 8-13.

- [12] H. Arafoui and K. Mori, "Autonomous Navigation Architecture for Load Balancing User demands in Distributed Information Systems," *IEICE Transactions on Communications*, 2001, E84-B (10), p.1085-1093.
- [13] A. Karmouch and V. A. Pham, "Mobile Software Agents: An Overview," *IEEE Communications Magazine*, July 1998, Vol. 36, No. 7.
- [14] M. K. Perdikeas *et al*, "Mobile Agents Standards and Available Platforms, *Comp.Net*" 1999, Vol. 31, No. 19, p. 1999-2016.
- [15] D. B. Lange and M. Oshima, "Seven Good Reasons for Mobile Agents," *Communications of the ACM*, 1999, Vol. 42, No. p. 88-89.
- [16] J. Cao, G. H. Chan, W. Jia, and T. Dillon, "Checkpointing and Rollback of wide-area Distributed Applications using Mobile Agents," *Proceedings International Parallel and Distributed Processing Symposium. San Francisco: IEEE Computer Society Press*, April 2001.
- [17] S. Funfrocken, "Integrating Java-based Mobile Agents into Web Servers under Security Concerns," *Proceedings 31st Hawaii International Conference on System Sciences*, 1998, p.34-43.
- [18] Marta Steenstrup, "Routing in Communications Network," Prentice Hall New Jersey USA, 1995, p.143-144.