International Journal of Smart Home
Vol. 9, No. 5 (2015), pp. 185-194
http://dx.doi.org/10.14257/ijsh.2015.9.5.18

Autonomous Network-Based Integration Architecture for Multi-Agent Systems under Dynamic and Heterogeneous Environment

Deng Xian-Rui[1], Liu Yu-Bin[2] and Feng Yu-Fen[3]

[1]*Department of Computer Science, Tangshan Normal University, Tangshan 063000, China*
[2]*Department of Computer Science, Tangshan Normal University, Tangshan 063000, China*
[3]*Department of Mathematics and Information Science, Tangshan Normal University, Tangshan 063000, China*
190233349@qq.com

Abstract

Multi-agent systems fit nicely into domains that are naturally distributed and require artificial intelligence technology. The autonomous network-based information services integration architecture has been designed to satisfy the multiservice utilization in rapidly changing environments. However, the increase in the total number of user requests and changes lead to the unbalancing load in the system and the overload in the locality. This paper proposes a new strategy to solve such problem. Autonomous load distribution can be achieved through the integrated access method, which reduces the total load of the system for the number of Pull-MAs sent to the system decrease. In addition, the information structure of integrated service area is effective to improve the ratio of the satisfaction of Pull-MAs with joint request on one node. As a result, the homogeneous distribution of the separated services requests and correlated services requests is guaranteed autonomously. The simulation witnesses the success of the proposed mechanism.

Keywords: *multi-agent system; artificial intelligence; autonomous network-based information service; load distribution.*

1. Introduction

With the advances in communication technologies and the decreasing costs of computers, Global Information Systems, such as Internet has become an attractive alternative for satisfying the information service needs of global users. Currently, most information service systems enable one-to-one interactions, which take place between one user and one provider. But sometimes the user requests multi-service, for example, a complete travel package made up of several flights and hotel information. In this case, the adaptability and timeliness have to be assured by the system. Conventional information service systems based on client/server model cannot meet users' heterogeneous and dynamically changing requirements.

The demand for information services is increasing at an explosive rate. Due to large geographical region of business operations and external competitive pressures, information services are expected to collect, store, retrieve, integrate and distribute timely at remote and dispersed nodes. In order to reduce network traffic and improve users' access time, we are developing a demand-oriented architecture called Faded Information Field (FIF), sustained by push/Pull mobile agents [1]. The characteristic of the FIF is to balance the cost of information allocation performed by push mobile agents and the cost of the access to the information performed by the pull mobile agents [2]. The information structure

ISSN: 1975-4094 IJSH

consequently permits to preserve the same access time to all unspecified users, whatever their current demand trend and volume.

Several standards have been offered to solve the problem of enterprise application and information integration such as CORBA [11] and UDDI [12]. The CORBA standard allows distributed objects to communicate with each other using a commonly defined interface language. However, it is not efficient in large decentralized information systems. Web services standard UDDI is a specification for distributed Web based information registries of Web services. It focuses more on information registry and discovery issues than integration issue.

2. Autonomous Correlated Services Access

Through the autonomous integration and dynamic online allocation described in the paper [3], majority of users' multiservice requests can be satisfied on one node at the same time. But still have some Pull-MAs with joint requests must get the correlated CHs from the separated node in each service area. In this case, the access time of the Pull-MA with joint request is composed of the time to move inside integrated service area to the corresponding nodes which store the requested correlated CHs, the time of processing at each node, and the time of returning to the user.

In a study on the performance of mobile agent systems, it is validated that the transmission time and the processing time strongly depends on the amount of information volume retrieved by the Pull-MA. For multi-service access, in order to obtain customized information, the users have to specify their personal conditions about the correlated services and the correlated CHs of each service.

As an example shown in Fig. 1, the Pull-MA is emitted with joint request about CH(11, 12, 14) of CC1 and CH(21, 22, 24) of CC2. But the most correlated CHs in the integrated area only consist of CH(11, 12, 21, 22). The Pull-MA cannot be satisfied on one node for multi-service request. As a result, it must move to the upper node to get the required CHs of CC1 at a certain node. And then through the integrated service area to move to a node to get the correlated CHs of CC2. If the Pull-MA gets all matching CHs from node C and carries these CHs to move to the node G, the transmission time between node C and G will be high. As CH(11, 12, 21, 22) allocated in the integrated area are common correlated CHs for all nodes, the Pull-MA can also get them at node G. It is enough that the Pull-MA just gets the lacking CHs of CC1 from node C. For instance, if the Pull-MA only takes CH14 moving from node C to node G and gets other correlated CHs from node G, the response time can be improved while transmission time decrease.

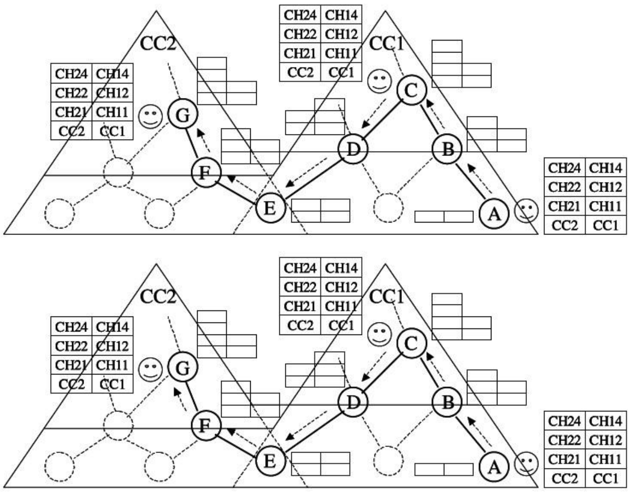


Figure 1. Autonomous Moving with Correlated CHs

The information structure of integrated service area permits to efficiently distributed Pull-MAs with joint requests between the levels. However, under rapidly changing environment, local overload in certain nodes might arise depending upon changing arrival rate. As the gradient of the information structure, any request bounded for a lower level can be processed in a higher level of information if some free resources are available. Similarly, the processes of higher-level requests for multi-service access can be executed at the lower nodes with the lacking CHs. But extra access overhead are induced and impair the overall response time in the case that access processes are executed in more different levels. To distribute the multi-service access in separated process is benefit only if all upper nodes able to satisfy the request are currently congested. In other conditions, to satisfy the multi-service request at appropriate nodes is more efficient.

To analyze the load distribution of Pull-MA with joint request, the simple information structure is considered, as shown in Fig. 2, and the M/M/k queuing model is applied. There are two kinds of Pull-MA's request in this model:

- **Single service request**: the Pull-MA only requests CHs of CC2 information service.
- **Multiple services request**: the Pull-MA requests correlated CHs of CC1 and CC2 at the same time.

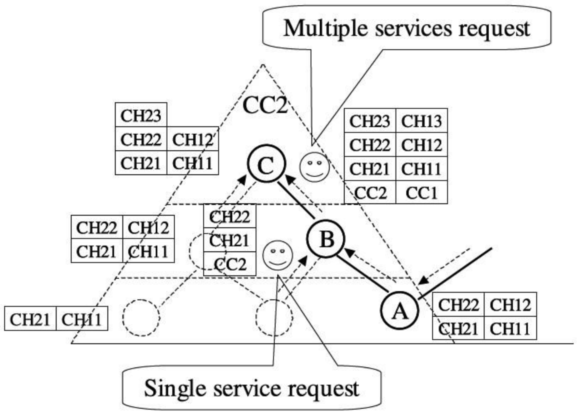


Figure 2. The Model of the Load Distribution Analysis

Table 1. Illustration of Parameters

λ01	arrival rate of Pull-MAs to request CH(21, 22)
λ02	arrival rate of Pull-MAs to request CH(21, 22, 23)
λ11	arrival rate of Pull-MAs with joint request for CH(11, 12) and CH(21, 22)
λ12	arrival rate of Pull-MAs with joint request for CH(11,12,13) and CH(21,22, 23)
μ0	service rate of Pull-MAs with single request
μ1	service rate of Pull-MAs with joint request
μ2	service rate of Pull-MAs for just getting the lacking CH

If only considering the average processing time of each node, two process methods exist for Pull-MAs with joint requests that cannot be satisfied at one node.

- **Forward to process**: the Pull-MAs with joint requests are processed in the appropriate level, i.e. the node B always forwards the unsatisfied Pull-MAs to the upper level.
- **Get and process**: the unsatisfied Pull-MAs with joint requests just get the lacking CH from the upper level and return back to process at the lower node.

These two methods are compared in this analysis. Therefore, the average response time of the node B and node C in two models can be expressed as follows, and Table 1 gives the illustration of parameters used in the formulas.

Forward to process:

$$\begin{cases} T_B = \dfrac{1}{\mu_0 - \lambda_{01}} + \dfrac{1}{\mu_0 - \lambda_{11}} \\ T_C = \dfrac{1}{\mu_0 - \lambda_{02}} + \dfrac{1}{\mu_0 - \lambda_{12}} \end{cases}$$

Get and process:

$$\begin{cases} T_B^{'} = \dfrac{1}{\mu_0 - \lambda_{01}} + \dfrac{1}{\mu_1 - (\lambda_{11} + \lambda_{12})} + \dfrac{1}{\mu_2 - \lambda_{12}} \\ T_C^{'} = \dfrac{1}{\mu_0 - \lambda_{02}} + \dfrac{1}{\mu_2 - \lambda_{12}} \end{cases}$$

From above analysis, we can see that the variation of the average response time of upper and lower nodes for the Pull-MAs relates to the changes in the users' preferences. If the sum of $\lambda 11 + \lambda 12$ is a constant and $\mu 1 < \mu 2$, the average processing time of *TC* increases rapidly in the case of overload with the $\lambda 12$ increase. But in such case, the lower node becomes underload. It is efficient to improve the average response time of the upper node by using *Get and process* method. Both waiting time and processing time at the upper node are reduced. On the other hand, this approach has the disadvantage to burden the Pull-MAs that can be satisfied at the lower node with the increasing waiting time.

A. Autonomous Processing

If the lower nodes only forward unsatisfied Pull-MAs with joint request to upper nodes to process, the local overload in certain nodes might arise depending upon changing users' preferences. The access load distribution must provide both efficiency in response time and fairness among the Pull-MAs with heterogeneous requests. Therefore, the autonomous access distribution mechanism is proposed to fairly reduce congestion under dynamically changing users' preference conditions. According to the current condition, Pull-MAs with joint request execute two process methods adaptively.

- If the current load is more than that of the upper node, forwards Pull-MAs directly to upper level and executes processes on upper node.
- If the heavy load in upper level is detected, gets the lacking CH from the upper node and executes processes on the lower node to reduce the processing overhead of upper level.

In the system, each node autonomously determine its process behavior for Pull-MAs to dynamically balance the load in the locality and offload some congested nodes by adaptive executing forward to process or get and process. The *forward to process* is reactive to a current overload situation making a reduction of the waiting time of Pull-MAs. The *get and process* proactive to reduce the servicing time on the upper node with heave load. Both processes assure the dynamic adaptation to unbalance of workload in the locality. If T_{up} and T_{low} are the current processing time of the upper and lower nodes, a new process behavior leads to new processing times as $T_{up}^{'}$ and $T_{low}^{'}$. The fairness for the process behavior of Pull-MAs with joint request permits to the total decrease in the response time of the latter is equivalent or improved compared to the increase in the response time of the former.

$$\left(T_{up} - T_{up}^{'}\right) + \left(T_{low} - T_{low}^{'}\right) \geq 0$$
$$(T_{up} > T_{low} \; or \; T_{up} < T_{low})$$

B. Autonomous Monitoring

Under rapidly changing conditions, neither the demand correlated CHs nor the required conditions of each user can be predicted in the system. The workload unbalance hence cannot be solved through the centralized approach without incurring large time overheads. It is necessary that the decision of the process behaviors for Pull-MAs with joint request at each node be only based on the node's local information. The autonomous process behavior of each node assures to keep the fairness of the changes of the processing times in current and upper nodes. The local processing time of each node is mainly due to the node utilization, which is defined as the ration between the arrival rate of satisfied Pull-MAs over the service rate. To estimate the load in the locality, each node monitors the node utilization of current and upper nodes. As shown in Fig. 3, the node utilization of the current node B and the upper node C can be given by:

$$\rho_{cu} = \frac{1}{2}\left(\frac{\lambda_{01}}{\mu_0} + \frac{\lambda_p}{\mu_1}\right)$$

$$\rho_{up} = \frac{1}{3}\left(\frac{\lambda_{02}}{\mu_0} + \frac{\lambda_f}{\mu_1} + \frac{\lambda_p}{\mu_2}\right)$$

• λp: process rate of Pull-MAs with joint request that cannot be satisfied at the current node, i.e. number of Pull-MAs undertaking the *get and process* behavior.
• λf : forward rate of unsatisfied Pull-MAs with joint request to the upper node, i.e. number of Pull-MAs undertaking the *forward to process* behavior.

C. Autonomous Adjusting

Any change of the access behavior has the effect on the processing times in current and upper nodes. The average response time of each node is mainly due to the length of the waiting queue. Based on the monitored value of the node utilization, each node estimates the average length of queue in current and upper nodes. Meanwhile, the forward ratio $r\ f$ which is defined as the ratio of the forward rate to the total number of unsatisfied Pull-MAs with joint request is dynamically adjusted according to the changes of node utilization. The more the processing time of the upper node increases, the less the forward ratio is adjusted. Consequently, we can modify the forward ration with a new value, by comparing the average length of queue in current and upper nodes.

$$r_f' = \begin{cases} r_f \square \dfrac{\rho_{cu}}{\rho_{up}}, & \dfrac{\rho_{cu}}{1-\rho_{cu}} - \dfrac{\rho_{up}}{1-\rho_{up}} > \lambda_f \\[2ex] r_f \square \dfrac{\rho_{cu}}{\rho_{up}}, & \dfrac{\rho_{up}}{1-\rho_{up}} - \dfrac{\rho_{cu}}{1-\rho_{cu}} > \lambda_f \end{cases}$$

Through the autonomous adjusting mechanism, the access distribution of Pull-MAs with joint request is gradually refined to improve the response time and assure the fairness of multi-service access. This step-by-step approach permits to achieve timeliness under rapidly changing environments with high autonomy of information services.

3. Performance Evaluation

In this section, two simulations are illustrated to show the efficiency of the autonomous correlated services access technology based on the autonomous integrated information structure. We have developed Autonomous Information Service System Simulator (AIS3) [10] for performance evaluation.

- **Load reduction**: to show that the autonomous integrated information structure and integrated access techniques are not only to improve the response time of multi-service access, but also to reduce the total load of the system.
- **Access distribution**: to show that the proposed autonomous process adjusting technique is efficient to reduce the local overload in rapidly changing user demand.

A. Load Reduction Evaluation

a) Simulation Model

To evaluate the performance of load distributionn, we consider two SPs to provide CC1 and CC2 services and heterogeneous user distributions. Figure 3 shows the fixed information structure of each service that has been constructed to map a certain volume of user requests. The CC1 service area is composed of 30 nodes dispatched on 5 levels and from SP to the edge level the information volume is 10CHs, 8CHs, 6CHs, 4CHs and 2CHs respectively. The CC2 service area is set up by 20 nodes with 4 levels and 10CHs, 8CHs, 6CHs and 3CHs are allocated from SP to the edge level.

The distributions of user demands for separated service and integrated service is shown in Table 2. Total number of CHs is 10 for each SP and the data size of each CH is 1 kBytes. Table 3 gives the initial parameters to perform the simulation.

b) Results

For comparison, three experiments are executed on the simulation model.

- The separated access method is applied for Pull-MAs with joint request on the separated service areas. The average response time of users who request only one service is observed.
- The integrated access method is applied for Pull-MAs with joint request on the separated service areas. The average response time of multi-service accesses is observed.

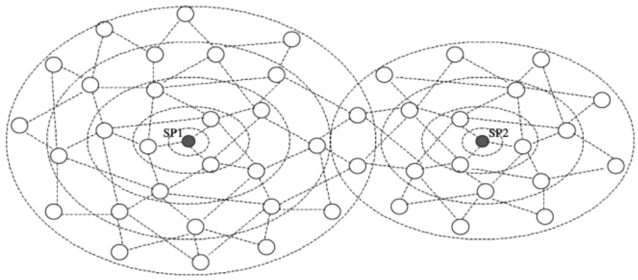


Figure 3. The Simulation Model of Load Distribution

Table 2. The User Distributions

CC1	Info. ratio	0.2	0.4	0.6	0.8	1.0
	users%	40	30	17	10	3
CC2	Info. ratio	0.3	0.6 0.8			1.0
	users%	50	30 15			5
CC1&CC2	Info. ratio	0.2	0.4	0.6	0.8	1.0
	users%	45	25	15	10	5

Table 3. Parameters of Simulation

Parameter	Value
The processing time of each Pull-MA	10 ms
Data size of the Pull-MA	2 kBytes
Transmission time of the Pull-MA	2ms/hop

- The integrated access method is applied for Pull-MAs with joint request on the integrated service area. The average response times of not only separated requests but also joint requests are recorded.

The users' preferences realized in the experiments are same, and among the number of users 80% for separated service and 20% for integrated service. We computed the average response time of each case for number of users from 200 to 2000 per second. The average response time obtained in the condition without integration is compared with the proposed technique in Fig. 4. The improvement for users' separated requests of each service and joint requests of integrated service is about 20% and 40% in average. The main reason is that the integrated access method reduces the total arrival rate of Pull-MAs to the system in one unit time. In addition, the information structure of integrated service area is effective to improve the ratio of the satisfaction of Pull-MAs with joint request on one node and the service rate of the system is therefore improved. As a result, the total load of the system is reduced autonomously thanks to the proposed integrated information service structure and access method.

B. Access Distribution Evaluation

a) Simulation Model

The performance of the autonomous access distribution is evaluated on the model that contains two SPs, as shown in Fig. 5. Each service area is composed of 10 nodes dispatched on 3 levels. From SP to the edge level the information volume is 3CHs, 2CHs and 1CH respectively. The integrated service area has been constructed to map a certain multi-service requests. The nodes in the system are assumed to be uniform in processing. Each node has three parallel queues for the different requests of Pull-MAs. The processing time depends on the the information volume requested by the Pull-MA.

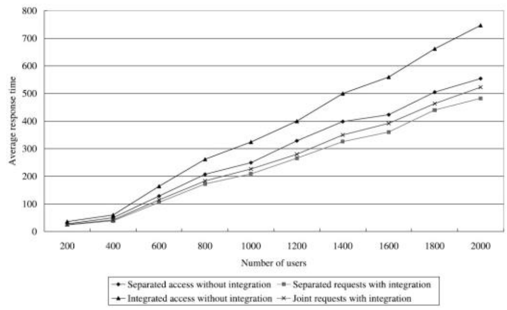


Figure 4. Comparison of the Average Response Time

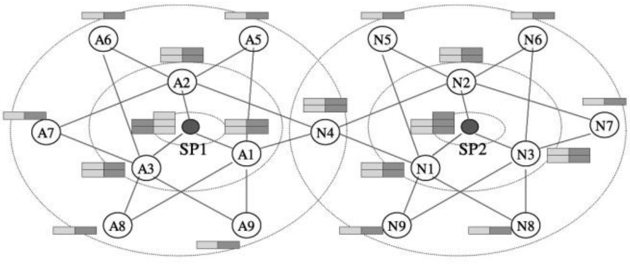


Figure 5. The Simulation Model of Access Distribution.

Table 4. Initial Setting

Parameter	Means	Value
λ0	arrival rate of Pull-MAs to request one service	110
λ11	arrival rate of Pull-MAs with joint request for CH(11, 21)	66
λ12	arrival rate of Pull-MAs with joint request for CH(11, 12, 21, 22)	Dynamic
λ13	arrival rate of Pull-MAs with joint request for all CHs	Dynamic
μ0	service rate of Pull-MAs with single request	40MA/s
μ1	service rate of Pull-MAs with joint request	20MA/s
μ2	service rate of Pull-MAs for just getting the lacking CH	60MA/s

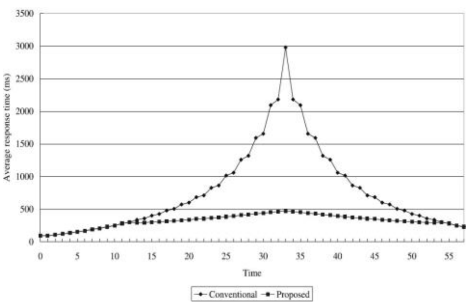


Figure 6. The Result of Comparison (a)

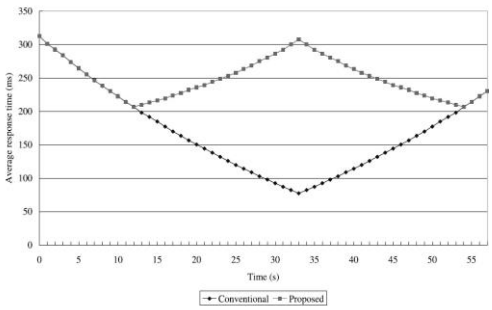


Figure 7. The Result of Comparison (b)

The arrival rate of Pull-MAs for each edge node is assumed to be constant and equal to 20MA/s. Two kinds of Pull-MAs are considered in this simulation.

- Among of them, 10 Pull-MAs request only one service and percentages of 1, 2 and 3 CHs requests are 60%, 30% and 10%, respectively.
- Other 10 Pull-MAs are joint requests and the percentage of CH(11, 21) requests is 60%. The proportion of users who want to get correlated CHs of CH(11, 12, 21, 22) and CH(11, 12, 13, 21, 22, 23) changes dynamically.

The proposed autonomous access distribution mechanism is measured and compared with the conventional process that Pull-MAs with joint request are forwarded to the corresponding level according to their required correlated CHs. The initial setting to perform the simulation is given in Table 4.

b) Results

In this experiment, each Pull-MA is dynamically guided towards the relevant service going from node to node. And the changing arrival rates of users' preferences are realized in the experiment. Figure 6 shows the average response time of users who want to get all correlated CHs. The autonomous processing and adjusting mechanisms clearly permit to maintain the access time of the Pull-MAs in an acceptable range. The proposed technique makes a 200% improvement in average compared to the *forward to process* case. On the other hand, the average response time of lower level is shown in Fig. 7. In the case of *get and process*, the average response time of lower level is increased. But take into account the improvement of the access time, the time cost in lower level is acceptable. As a result, with the changing arrival rates, the autonomous access distribution technique permits much improvement of the access time of joint requests when compared to the conventional process case.

4. Conclusion

The autonomous network-based information services integration architecture has been designed to satisfy the multiservice utilization in rapidly changing environments [3]. However, the increase in the total number of user requests and changes in the users' preferences cause the unbalancing load in the system and the overload in the locality. In this paper, a new approach of autonomous correlated services access is proposed to reduce the load of the system and achieve the timeliness of correlated services utilization.

Autonomous load distribution can be achieved through the integrated access method, which reduces the total load of the system for the number of Pull-MAs sent to the system decrease. In addition, the information structure of integrated service area is effective to improve the ratio of the satisfaction of Pull-MAs with joint request on one node, i.e. the time of Pull-MAs staying in the system decreases. In other words, the service rate of the system is improved. As a result, the homogeneous distribution of the separated services requests and correlated services requests is guaranteed autonomously.

The results of simulation show that the system can improve the average response time not only for joint requests of correlated services, but also for separate requests of each service under the static distribution of users' preferences. However, in many conditions, the users' preferences change dynamically. Then the problem is: how to avoid the local overload in the system with rapidly changing users' demands. The proposed autonomous access distribution is to solve this problem. The autonomous processing and adjusting mechanisms permit to dynamically balance the load in the locality and offload some congested nodes by adaptive executing forward to process or get and process to achieve timeliness and assure the fairness under rapidly changing environments.

Acknowledgement

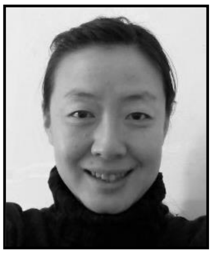
This work is supported by the Hebei province science and technology plan project, No. 13220319D and Doctor Foundation of Tangshan Normal University No. 07A01.

References

[1] K. Mori, "Autonomous fading and navigation for information allocation and search under evolving service system", Proc. IEEE Conf. on APSITT, **(1999)**.
[2] H. F. Ahmad and K. Mori, "Autonomous information provision to achieve reliability for users and providers", Proc. IEEE Conf. on ISADS, **(2001)**.
[3] X. D. Lu, K. Moriyama, I. Luque, M. Kanda, Y. Jiang, R. Takanuki and K. Mori, "Timeliness and reliability oriented autonomous network-based information services integration in multi-agent systems", IEICE Trans. Inf. & Syst., vol. E88-D, no. 9, **(2005)**, pp.2089–2097.
[4] O. Etzioni, K. Golden and D. Weld, "Sound and efficient closeworld reasoning for planning", Journal of Artificial Intelligence, vol. 89, no.1-2, **(1997)**, pp.113–148.
[5] O. M. Duschka and A. Y. Levy, "A recursive plans for information gathering", Proc. 15th Internatioal Joint Conference on Artificial Intelligence, **(1997)**.
[6] O. M. Duschka, "Query optimization using local completeness", Proc. 14th Natioal Conference on Artificial Intelligence, **(1997)**.
[7] X. D. Lu and K. Mori, "Autonomous integration and optimal allocation of heterogeneous information services for high-assurance in distributed information service system", IEICE Trans. Inf. & Syst., vol. E86-D, no. 10, **(2003)**, pp. 2087–2094.
[8] SUN, "Jini network technology", http://www.sun.com/software/jini/.
[9] K. Mori, "Autonomous decentralized systems: Concepts, data field architecture and future trends", Proc. IEEE Conf. on ISADS, **(1993)**.

[10] Y. Zhou, H. F. Ahmad, H. Arfaoui and K. Mori, "Autonomous information allocation through mobile agents to achieve load blanceing in distributed information service system", Proc. IEEE Conf. on IWADS, **(2002)**.
[11] OMG, "The common object request broker: Architecture and specification,"http://www.omg.org/corba/.
[12] UDDI.ORG, "UDDI technical white paper," http://www.uddi.org/.
[13] J. He, Y. Geng and K. Pahlavan, "Modeling Indoor TOA Ranging Error for Body Mounted Sensors", IEEE 23nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), **(2012)**; Sydney, Australia.
[14] S. Li, Y. Geng, J. He and K. Pahlavan, "Analysis of Three-dimensional Maximum Likelihood Algorithm for Capsule Endoscopy Localization", 5th International Conference on Biomedical Engineering and Informatics (BMEI), **(2012)**; Chongqing, China.
[15] M. Huhns, N. Jacobs, T. Ksiezyk, W. M. Shen, M. Singh and P. Canata, "Enterprise information modeling and model integration in carnot", Proc. Enterprise Integration Modeling, MIT Press, **(1992)**.
[16] R. J. Bayardo, Jr., W. Bohrer, R. Brice, A. Cichocki, J. Fowler, A. Helal, V. Kashyap, T. Ksiezyk, G. Martin, M. Nodine, M. Rashid, M. Rusinkiewicz, R. Shea, C. Unnikrishnan, A. Unruh and D. Woelk, "InfoSleuth: Agent-based semantic integration of information in open and dynamic environments", Proc. ACM SIGMOD, **(1997)**.
[17] Y. Arens, C. A. Knoblock and W. Shen, "Query reformulation for dynamic information integration", Journal of Intelligent Information Systems, vol. 6, no. 2-3, **(1996)**, pp.99–130.
[18] Y. Geng, J. He, H. Deng and K. Pahlavan, "Modeling the Effect of Human Body on TOA Ranging for Indoor Human Tracking with Wrist Mounted Sensor", 16th International Symposium on Wireless Personal Multimedia Communications (WPMC), **(2013)**; Atlantic City, NJ.

Authors



Deng Xian-Rui, she received her PhD in Control Theory and Control Engineering in 2007 from the Institute of Automation, Chinese Academy of Sciences, Beijing, China. She is currently associate professor in the Computer Science Department of Tangshan Normal College, Tangshan, Hebei, China. Her research interests include Internet of Things and computer simulation.

Liu Yu-Bin, he received his B.E degree in application of electronics from Hebei Normal University in 2003, the Master degree of engineering in software engineering from Tongji University. Now he is lecturer at Department of computer science, Tangshan Normal University. His research interests mainly include Internet of Things, graphics and image processing.

Feng Yu-Fen, she received her B.Sc degree in computer science from Lanzhou University in 1992, the Master degree of engineering in computer technology from University of Science and Technology Beijing. Now she is an associate professor at Department of Mathematics and Information Science, Tangshan Normal University. She is a member of China Computer Federation. Her research interests mainly include virtual reality, graphics and image processing.

