Fk means를 이용한 동적객체그룹관리기반 지능형 멀티 에이전트 분산플랫폼☆

Intelligent Multi-Agent Distributed Platform based on Dynamic Object Group Management using Fk-means

이 재 완* 나 혜 영** 마테오 로미오***

Jae-wan Lee Hye-young Na Romeo Mark A. Mateo

요 약

효율적인 자원공유 및 동적인 시스템구성을 위한 지능형 분산 접근방식에서 주로 멀티에이전트 시스템을 사용한다. 또한 객체중복은 고장허용시스템을 구축하여 시스템에 예기치 않은 결함의 문제를 해결하기 위해 흔히 사용된다. 본 논문은 동적인 객체그룹관리에 기반한 지능형 멀티에이전트 분산플랫폼을 제시하고, 제안한 filtered *k*-means (F*k*-means)를 기반으로 하여 객체검색기법을 제시한다. 객체 결함의 경우에, 대체 객체를 검색하여 클라이언트에게 적절한 객체를 투명하게 재 연결 시켜주기 위해 F*k*-means를 사용한다. 검색방법을 효율적으로 수행하고, 그룹 내의 적절한 객체를 포함시키기 위해 F*k*-means의 여과 범위를 설정한다. 시뮬레이션 결과 제안한 기법이 분산객체그룹에 대해 빠르고 정확한 검색을 나타내었다.

Abstract

Multi-agent systems are mostly used to integrate the intelligent and distributed approaches to various systems for effective sharing of resources and dynamic system reconfigurations. Object replication is usually used to implement fault tolerance and solve the problem of unexpected failures to the system. This paper presents the intelligent multi-agent distributed platform based on the dynamic object group management and proposes an object search technique based on the proposed filtered *k*-means (F*k*-means). We propose F*k*-means for the search mechanism to find alternative objects in the event of object failures and transparently reconnect client to the object. The filtering range of F*k*-means value is set only to include relevant objects within the group to perform the search method efficiently. The simulation result shows that the proposed mechanism provides fast and accurate search for the distributed object groups.

☞ keyword : Multi-agent system, distributed object group, clustering, object replication

1. Introduction

Recently, most researchers integrate multi-agent systems (MAS) for their proposed system to implement the distributed task oriented and intelligent approaches. Moreover, the abstraction of tasks using MAS is easily understood by the designers. In the distributed object systems, most of the design uses componentized approach where each component has specialized function and includes scalable design that allows additional components to integrate easily. However, these components are not provided with intelligence to

* 종신회원 : 군산대학교 교수
jwlee@kunsan.ac.kr
** 준 회 원 : 군산대학교 석사-정보통신전파
jyunbba@hanmail.net
*** 정 회 원 : Kunsan National University, Ph.D. student
rmmateo@kunsan.ac.kr

[2008/07/10 투고 - 2008/7/15 심사 - 2008/11/10 심사완료]
☆ This work is supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2006-521- D00372)

adapt on the changes based on current system requirements. The intelligent distributed framework is proposed [5] to address the integration of MAS to the distributed object group system. In implementing the reliable distributed object system, failures and other unexpected breakdowns are managed by the system transparently. The basic mechanisms of fault management are replication scheme and transparent forwarding of request where the access of client to object is seamlessly forwarded to another source whenever a server that hosts the object or the object itself fails. The fault tolerance is needed to be designed correctly so that the interaction of components to access object replica is not ambiguous and efficiently selects the appropriate object for the service. Object group management is usually used to manage these groups of object replicas [1]. Each group member uses group interactions in order to provide consistent updates among the replicated objects. The FT-CORBA was introduced by OMG (Object Management Group) for the specification of CORBA to implement the fault tolerance that includes object cooperation from object groups [2,3]. Implementing the IORG (Interoperable Object Group Reference) in the object group to inform the changes of object properties to all other group is studied [4]. However, as the system expands, the demand for new objects also increases. Base on the needs of accurate search of resources and system expansion, we use a knowledge based clustering algorithm to significantly improve the accuracy of object search.

This paper presents an efficient fault management for distributed objects using the proposed object group management of the intelligent distributed platform [5] and proposes an object search technique based on filtered *k*-means (F*k*-means) for accurate searching of objects within the objects groups. The F*k*-means enables client to manage IOGR to support accurate and fast object search. The proposed system manages the creation of groups, implements fault tolerance and provides fast search of objects. Moreover, in case of object disconnection, the system redirects the request to a replica object. The F*k*-means algorithm is used to classify objects and the system selects an appropriate object that will be used as an alternative object. In able to provide the right object, the object reference group manager updates each IOR concurrently to other members of object group in real time. A simulation is done to compare the performance of the proposed algorithm to classical method based on throughput time and classification accuracy.

2. Related Works

Independency, transparency and scalability are significant features required from the fault tolerance schemes for modern cluster of computers. The RADIC (Redundant Array of Distributed Independent Checkpoints) [6] is proposed for fault tolerance to cope with the requirements for grid of computers. RADIC implements the fault tolerance activities, transparently to the user application, using a message log rollback recovery protocol. Fault management based on mobile agents is a distributed approach and popular to implement the fault tolerance. Chameleon [7] is an agent based infrastructure focused on fault tolerant system where services are fully distributed so that every node can be used as alternative service in the event of failures. Each of the components is active for a certain period. If a component fails during its active phase, there is a provision for recovery, either by switching to another node or by regenerating the component. However, the procedure is lack of accuracy in choosing the appropriate service or object.

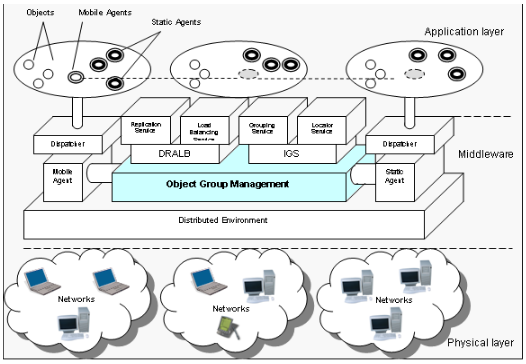
CORBA is a robust environment for middleware development based on the Object Request Broker (ORB) architecture that complies with the standards of the Object Management Group (OMG) [8]. The ORB supports a flexible system that allows applications to run on heterogeneous networks. CORBA supports a highly distributed system consisting of many remote objects, but has limited functionality to deal with load balancing or fault tolerance. The FT-CORBA which extends the CORBA specifications provides an additional service: failure detection and failure recovery – ways to improve an object group using fault tolerance [2]. This paper improves the specification of FT-CORBA by proposing a clustering technique to search the appropriate object in case of disconnection based on the preference of the client. Cluster analysis is used for data analysis in solving classification problems [9]. Mostly, it is used to implement the intelligent information retrieval [10]. The goal of cluster analysis is categorization of attributes like consumer products, objects or events into clusters or groups, so that the degree of correlation is strong between members of the same cluster and weak between members of different clusters [11]. The *k*-means is one type of non hierarchical clustering which is applicable only if the mean of a cluster is known. In addition, it is more sensitive to noise and other outlier data points because small number of such data can considerably affect the mean value. Equation 1 presents the equation for the basic objective function of *k*-means algorithm.

$$\min \sum J = \sum_{n=1}^{N} \sum_{m=1}^{M} \left| x_{nm} - c_n \right| \qquad (1)$$

The goal of this process is to come up with a reasonable initial partition and to move samples from one group to another if such move improves the value of criterion function. Mostly, the objective of these clustering is to minimize the compactness of the group with a constraint from the number of centers. However, the procedure consumes more time, if there are more data to be process on the iteration method of the objective function. In our work, we proposed a filtering scheme to solve this problem.

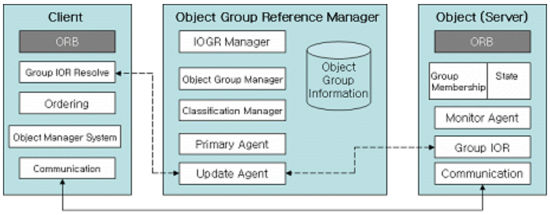
3. Intelligent Distributed Platform based on Object Group Management

This study uses the intelligent distributed platform from the previous research [5] which addresses the necessity of an intelligent approach in distributed object environment. The efficiency of data acquisition and rule extraction is considered in the design. Multi-agent approach is a popular scheme to implement information retrieval to process intelligent schemes [12,13,14,15]. The global view of the proposed intelligent model is shown in Figure 1 which consists of three layers; physical, logical and application layers. The physical layer represents networks of different computers like PC, laptop and PDA. The logical layer acts as the middleware where services are transparently executing by the request of clients to the objects. Also, the interaction of clients and object services are handled by the logic layer. The physical layer and application layer do not need to know the configuration on how to find, where to find and how to manage the resources but transparently executes the services to work with the clients and servers. This paper focuses on the object group management, shown in the figure, that is mainly interacting to the dynamic replication and adaptive load balancing scheme (DRALB) and intelligent grouping scheme (IGS) to provide the efficient management to the services like the replication service, load balancing service, grouping service and locator service. The application layer is consisted of objects and agents utilizing the distributed environment.



(Fig. 1) Architecture of the intelligent distributed platform

This paper mainly focuses on solving the problem of fault tolerance in the intelligent distributed platform. We propose an object management system for managing large number of objects based on the FT-CORBA to improved fault tolerance and to aim an efficient service for clients. In addition, the system uses multi agent approach for clustering the object group, notifying members of the group and gets the information from objects. Also, the object group management system is integrated in the grouping service of the intelligent distribute platform shown in Figure 1. The proposed system uses the object control technique to process grouping of objects and the mobile agent to automate the classification procedure. The object group management enables a search of the objects and controls the systems in managing objects that exists between application servers. Figure 2 presents the over all design of the object group management based on multi agent system.



(Fig. 2) The object group management using multi agent approach

3.1 Components of the Object Group Management

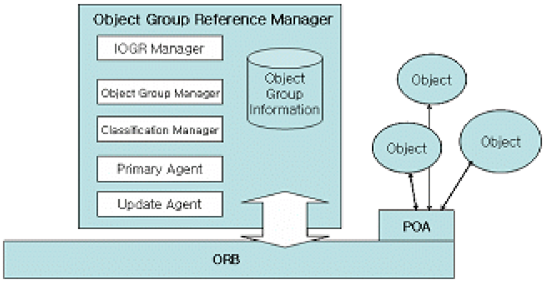
Client. CORBA is the underlying technology that is used by the client where the request method uses the Interoperable Object Reference (IOR) in requesting an object and the procedure is performed through object group management by communicating within the servers. The request is classified through the servers and searches for the appropriate object to serve the request. After the finding the appropriate object for the request, the

IOR is analyzed to find the replica that is possible to process the request.

Object Group Reference Manager. A client manages the request based on the IOR which can determine the appropriate object group where Interoperable Object Group Reference (IOGR) produces IOR of the object that is a replica according to FT CORBA. The tasks of object group reference manager are adding and removing IOR members, object generation, decides priority tasks and characteristics validation between objects. The administrator gets the object information which is processed from clustering method where the characteristic of the data from objects is used for analysis and generation of rules for filtering the objects. The object is looked up through the servers based on the client information to manage the appropriate object to be accessed by the client.

Object (Server). The servers provide group membership information, location information and group IOR information to process the client request on the proposed system. IOR information is managed by object group reference administrator. Object Group Reference Manager uses the multi agent approach shown in Figure 3.

IORG Manager. The IOGR manager manages the object group management where it generates IORG, filters objects and classifies objects. The IOGR manager decides the ID and name of an object by the help of naming service. Also, the IORG filtering procedure refines the members of the group selected from IORG lists that groups the most relevant objects. This procedure verifies if there is strong a relationship between the objects and decides if the object belongs to the group.



(Fig. 3) Interaction of object group management

Object Group Manager. The object group manager manages the adding and removal of object members from a group. The data inputs of an object property are the following; ft_domain_id, object_group_id, and object_group_ref_Property and these are encapsulated into information of TAG_FT_GROUP to determine the member and add the object if it is a new member.

Classification Manager. The classification manager classifies objects in the group. The grouping method used is based on Fk means that is an improvement of the existing k means. The proposed Fk means is proposed to improve the speed of the k means to group the objects. The procedure is done dynamically by the classification manager.

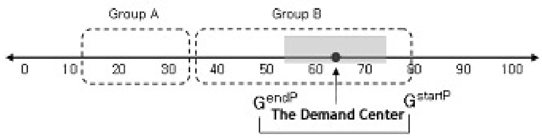
Primary Agent. The primary agent searches the status information of an object given by the monitor agent and decides the priority order of objects to be accessed by clients. At this point, faulty objects are determined and objects are prioritized based on the characteristic that the client is requesting.

Update Agent. The IORG has an IIOP (Internet Inter ORB Protocol) profile about all groups member from the current group membership. However, all members of a group must be known to the IORG where the object participates in a group. As an object joins a group, it adds its new group reference address to the system to guarantee the transparency of object. Therefore, the IORG must communicate to all members of the groups every time a join or leave of object occurred. The agent dispatches IOGR information of object group management and updates IORG to the members of each group. These procedures occur to provide the new information to all members of the group.

Monitor Agent. Monitor agent collects the state of object, communication cost, and relevance of the information and return to the mobile agent. This information provides necessary information to decide the primary IOR.

4. Filtered k means for Fault Tolerant Object Group

Managing the next object that will be serving clients in case of disconnection of the original object is the goal of the proposed system. However, the object that was selected may not be the most right object because there is no profile verification of objects. The object group management uses the proposed filtered k means algorithm which is based on k means algorithm. The k means is one type clustering which is applicable only if the mean of a cluster is known. It uses the Euclidean distance to calculate the membership of each data. In addition, it is more sensitive to noise and other outlier data points because small number of such data can considerably affect the mean value. The propose algorithm uses filtering value to remove the less relevant data from the center value of the k means. In this method, it speeds up the performance in finding the appropriate object compared not filtering the search. In the first run, the k means is processed to group the objects according to the order where the values are transformed in a vector shown in Figure 4.



(Fig. 4) Object groups ordered in a vector values

Figure 4 shows the boundary of the group indicated by doted lines and filtered objects indicated by the shaded box. G^{endP} and G^{startP} are the values that determines the range where to search an object. The selected group to process the object search is decided by the k means in Equation 1. The filtering range of a group is based on the users' preferences in searching the object and processes this in classification method and provides the center value Q where Q is pointed inside the groups' vector range. First, G^{endP} and G^{startP} are calculated to get the range length in Equation 2 where P1 is the first vector value and P2 is the last vector value of the group. The L1 and L2 indicates the range or length of the vector and then multiplied by the filtering rate or χ. χ is the range from 0 to 1.0 number of object before filtering. If filtering value is large then the grouping of objects also expands. This implies a correct setting of values from the filtering value. In Equation 3 calculates the value of the group starting point indicated by T1 and value of the group ending point T2.

$$L_1 = Q - P_1 \text{ and } L_2 = P_2 - Q \tag{2}$$

$$T_1 = L_1 \times \chi \text{ and } T_2 = L_2 \times \chi \tag{3}$$

T_1 and T_2 are indicators filtering range to search the object presented in Equation 4 and 5 which illustrated in Figure 5.

$$G^{startP} = Q - T_1 \tag{4}$$

$$G^{endP} = Q + T_2 \tag{5}$$

As mentioned above, the filtering range is acquired from the request of information from Q T_1 to Q+T_2 to each server that has objects specified in the filtering range.

5. Experimental Evaluation

The intelligent distributed platform of the previous research was used to simulate the proposed object group management in implementing the fault tolerance. This paper used the proposed object search method in case of object disconnection based on the Fk means for fast and accurate search of objects. First, the object grouping was done by determining each class using k means. Clustering the objects separates k groups where the range of the group is determined by the clustering method. After determining the number of clusters, each object is identified by calculating the membership. The class that has the highest membership from the classified data of object is the chosen group. Also, the structure that was built by clustering is used for the object search method in classification algorithm. In the proposed search, the redirection of the users to another object replica is faster because of filtering which eliminates the unnecessary data that are looked up through the procedure.

5.1 Result of the Experiment

The locator service registers the information of the objects in intelligent distributed platform. Also, locator service classifies the names of all object replication. If the object contains the same function from the replicated object then a different name is specified to the replicate object. The clustering procedure determines the class value of each object and groups the objects based on the proposed algorithm. The result of Fk means is compared to k means for the performance of grouping. The environment of the experiment generates a random of 200 objects and set the value of 10 for small groups and value of 7 for large groups shown in Table 1.

(Table 1) Data set for clustering

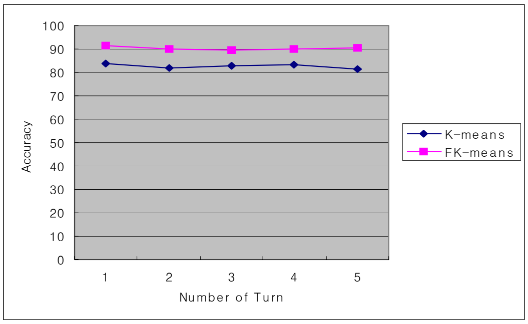
	Big Group	Small Group	Object
Computer	C	CA~CJ	28
Business	B	BA~BJ	29
Art	A	AA~AJ	25
Game	G	GA~GJ	29
Health	H	HA~HJ	30
Science	S	SA~SJ	28
Music	M	MA~MJ	31

The experiment finds the optimal value of the filtering of the Fk means by changing the range of χ from 0 to 1.0. Each result from changing the range of χ is shown in Table 2. We consider the most number of relevant data with high classification accuracy. As observed in the table, the 0.8 filter value has the high accuracy rate and has 150 data, compared to 0.3 filter value with the highest accuracy but has smaller data and 0.9 filter value with the highest number of data but lower accuracy. This also shows that the 0.8 filter value is the optimal value for filtering the data. This also means that the value is already enough to implement the filtering value because there will be still have irrelevant data in setting the highest filtering value and the relevant data can be deleted in setting the lowest filtering value.

(Table 2) Changing the filtering range value with number of filtered data and accuracy

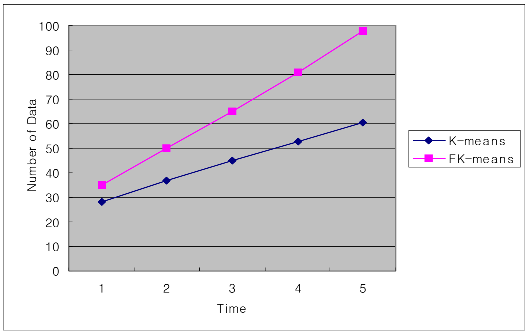
Filtering Range	Number of Data	Accuracy
0.2	119	0.91
0.3	122	0.92
0.4	122	0.91
0.5	125	0.91
0.6	130	0.88
0.7	138	0.90
0.8	150	0.90
0.9	175	0.83

The experiment has done a 5 cross-fold validation in classifying the filtered data shown in Figure 5. Usually, cross-validation is used to test the accuracy of classification in data mining methods. The filtering rate of 0.8 value based on the result from Table 2 was used before the classification procedure is applied. The data were processed to the proposed *Fk*-means and k-means and were compared shown in the Figures 5 and 6.



(Fig. 5) Result of accuracy of clustering

The accuracy of the Fk means is much higher compared to k means shown in Figure 5. The simulation of the search objects is processed within the group range. Using the 0.8 value of filtering range, a constant of almost 90 percent is achieved on classification accuracy. The filtering rate adjustment is necessary to speed up the search for objects which is shown in Figure 6.



(Fig. 6) Amount of data processed over time

Figure 6 shows the performance of amount of data to process over time using the two algorithms. In 5 seconds time, the Fk means has 38 more processed data than k means which is 63% more data was processed. As the time increase, the amount of the data processed from the Fk means increases dramatically more than the k means.

6. Conclusion and Future Works

The design of the distributed system is continuously evolving to adapt the system requirements base on the current changes in system environments. Multi agent systems, a part of distributed system, are designed with intelligence which is useful to implement the adaptive system applications. The intelligent distributed framework was design to realize the integration of intelligence by multi agent approach to the classical distributed object system. This paper contributes on the fault tolerance of the previous work. The design of the object group management is based on multi agent system to implement the efficient fault management. Using the multi agent approach in the object management group reduces the burden of IOGR to manage the group. We proposed a clustering method to create the structure of the object searching and forwarding mechanism based on the proposed filtered k means in case of object disconnection. Using the given data, we determined the optimal value of the filtering range which was 0.8. The result using the proposed Fk means shows more accurate on searching by performing the cross validation and processed more data compared to the k means technique.

The proposed object group management system will be optimized by implementing prediction techniques to decide faster which object should be forwarded and the proposed framework will be applied in the various fields like the healthcare system, an electronic commerce system, and other environments such as ubiquitous systems.

References

[1] S. Maffeis, "The Object Group Design Pattern", The 1996 USENIX Conference of Object Oriented Technologies (1996)

[2] S. Maffeis, "Adding Group Communication and Fault Tolerance to CORBA", The 1995 USENIX Conference on Object Oriented Technologies (1995)

[3] R. Baldoni, C. Marchetti, R. Panella, and L. Verde, "Handling FT CORBA Compliant Interoperable Object Group References", 7th IEEE International Workshop on Object oriented Real time Dependable Systems

[4] D. Lee, D. Nam, H.Y. Youn, and C. Yu, "The Implementation and Analysis of OCI based Group Communication Support in CORBA", IEEE 2001 Pacific Rim International Symposium on Dependable Computing (2001)

[5] R. M. Mateo, J. Lee, "Intelligent Distributed Platform using Mobile Agent based on Dynamic Group Binding", Journal of KSII, Vol. 7, No. 3, pp 131-143

[6] D. Angelo, R. Dolores and L. Emilio, "An Intelligent Management of Fault Tolerance in Cluster using RADICMPI", LNCS Vol. 4192

[7] R.K. Iyer, Z. Kalbarczyk, S. Bagchi, "Chameleon: Adaptive Fault Tolerance Using Reliable, Mobile Agents", SRDS 97, p. 61

[8] Object Management Group : CORBA Specification available at http://www.omg.org

[9] J. Han and M. Kamber, "Data Mining Concepts & Techniques", USA: Morgan Kaufmann (2001)

[10] P. Bellot, M. El Beze, "A Clustering Method for Information Retrieval", Technical Report IR 0199 (1999)

[11] C. Fraley and Raftery, “Model based Clustering, Discriminant Analysis and Density Estimation”", Journal of American Statistic Association (2002)

[12] H. Tu, and J. Hsiang, “An Architecture and Category Knowledge for Intelligent Information Retrieval Agents”, Proceeding of the Thirty First Annual Hawaii International Conference on System Sciences, Vol. 4. (1998) pp. 0405

[13] S. Mukhopadhyay, S. Peng, R. Raje, J. Mostafa, and M. Palakal, “Distributed Multi agent Information Filtering A Comparative Study”, Journal of the American Society for Information Science and Technology, Vol. 56, No. 8, pp. 834-842

[14] S. Yi, B. D. Gerardo, Y.S. Lee, J. Lee, “Intelligent Information Search Mechanism Using Filtering and NFC Based on Multi agents in the Distributed Environment” LNCS Vol. 3982, pp. 144-153

[15] J. Ko, B. D. Gerardo, J. Lee, J.J. Hwang, “The Information Search System Using Neural Network and Fuzzy Clustering Based on Mobile Agent”, LNCS 3481, pp. 205-214

◐ 저 자 소 개 ◑

이 재 완(Jaewan Lee)

1984년 중앙대학교 학사 전자계산학
1987년 중앙대학교 석사 전자계산학
1992년 중앙대학교 공학박사 전자계산학
1996년 3월~ 1998년 1월 한국학술진흥재단 전문위원
1992 ~ 현재 군산대학교 교수
관심분야 : 분산 시스템, 운영체제,데이타마이닝, 컴퓨터 네트워크
E-mail : jwlee@kunsan.ac.kr

나 혜 영(Hyeyoung Na)

2005년 군산대학교 학사 정보통신전파
2007년 군산대학교 석사 정보통신전파
관심분야 : 분산 시스템, 컴퓨터네트워크,데이터마이닝
E-mail : jyunbba@hanmail.net

Romeo Mark A. Mateo

2004 West Visayas State University, Philippines
BS in Information Technology
2007 Kunsan National University, South Korea
Master of Engineering major in Information and Telecommunications
2007 ~ currently Kunsan National University, South Korea
Graduate student in Ph.D course
Research interest : Distributed systems, data mining, fuzzy systems, multi agents, mobile computing, ubiquitous sensor networks
E-mail : rmmateo@kunsan.ac.kr