New Genetic Algorithm to Predict Data In Mobile Databases

Leila Alaei Sheini
Department of Computer
Engineer, Khomein Branch,
Islamic Azad University
Khomein, Iran

*Hamid Paygozarh
Department of Computer
Engineer, Khomein Branch,
Islamic Azad University
Khomein, Iran

Mohammad Khalily Dermany
Department of Computer
Engineer, Khomein Branch,
Islamic Azad University
Khomein, Iran

Abstract: Modern enhancement in wireless network technology, satellite services and cell concoctions has been provided the possibility to use mobile computing systems. This movement for conducting mobile computations which needs receiving data from databases is problematic. System movement while performing transaction will lead to disconnection of mobile system with wireless network server. This disconnection will lead to fall running transaction and mobile system will be obliged to run transaction again from the beginning. One of the techniques to prevent from this transaction falling is to predict till to connect to the next mobile network. Our purpose is to present a solution to predict information to send to the mobile system in order to be used in the memory using Memetic algorithm. Generally, Memetic algorithms in optimal selection and design in many technical and scientific problems causes producing the best product or acceptable solution. The main purpose to use these kinds of optimization algorithms is to choose the best information to be sent to the mobile system while disconnection to prevent from transaction falling. This research has concentrated only on information sent to the mobile system.

Keywords: Mobile Database, Cache Operation, Mobile Client, Cell Connection, Memetic Algorithm..

1. INTRODUCTION

Recently, technical enhancement relevant to the portable PCs development and also development of wireless technologies causes emerging portable PCs with capability to carry wireless connections and gives users a possibility to do actions in distributed calculated environment even while moving. Mobile users are willing to access private or corporate databases which have been saved in mobile or fixed hosts and present their information update and queries in wire and wireless networks [1].

In this kind of service giving, limitation such as mobile system power supply, expensive service and limitation of mobile systems' bandwidth and limitation of service giving space will lead to disconnection of mobile system with service giving center. This disconnection will lead to reject running transaction and mobile system will be obliged to run transaction again from the beginning. This rerunning requires costs such as connection to the network and time [2]. One of the methods to prevent from this transaction rejection is to use cache operation in mobile system that mobile service giver center predicts extra information more than what have been requested while receiving request from mobile system and sends these information to the mobile system. Mobile system stores extra information in Cache memory and as a following of transaction running or next transactions, if there is requested information in local memory, it uses Cache information and there is no need to connect and send request and receive information. This issue prevents from transaction rejection due to disconnection while running and helps with decreasing costs and increasing speed to response to mobile system user. The most important issue in this technique is correct and optimal prediction of information in order to be sent which needs a kind of data mining in service giver database table. In this paper, Memetic algorithm has been applied for data mining [3].

2. MOBILE DATABASE

Along with achieved enhancement in the context of wireless communication technology and portable computation devices called mobile computation. So that, users move their portable devices and meanwhile these users have a possibility to access information and services without considering physical place and behavioral movements. In fact, with appearing computer networks especially internet network and simultaneously emerging portable calculative tools such as qualitative PCs and PDAs and new need of users to access internet network, this possibility has been created that people in anywhere and anytime can do their actions through wireless networks. Due to growth speed of wireless devices and mobile calculative devices, in a very close future, millions users will be using portable computers and wireless communication devices and their main demand is to connect World Wide Web and resolve calculative and information needs. Wireless technology gives this opportunity to the users to keep their connection with network while moving [4]. In such architecture due to wireless nature and environment and mobile users, distributed environment is proposed in a new way. In such environment, mobile users due to mobile environment are more affected by disconnection. Mobile calculations are more affected by disconnection. Mobile calculations due to mobile users and their computers have been distinguished from fixed connections. Mobile service giver/taker database system is as [5]:

•   A set of weak or strong reliable information services which have been connected via a fast-paced immobile network.

•   A set of unreliable service takers is defined on wireless network.

Therefore, architecture includes mobile and immobile components. Mobile sector is known as Mobile Unit (MU).

MU is a mobile device which is able to connect with immobile network through wireless system. In figure 1, architecture of mobile database has been shown.
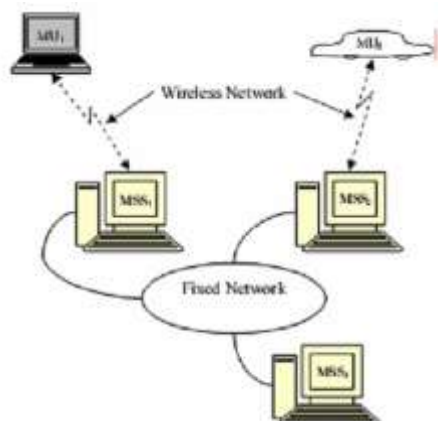

Figure 1: Mobile database architecture [1]

In fact, computer is a communication tool which is working with battery and it can move freely in a certain and limited geographical area (area covered by workstation). Restriction of this geographical area is due to restriction of bandwidth in wireless communication channels.

Immobile components of the architecture are connected through fast-paced immobile network (speed of this network can be in Mbps/Gbps range). Also, data and databases have been distributed throughout components of this fixed network. These fixed components are in 2 kinds:

- Fixed Host (FH)
- Workstation

In this environment, fixed hosts don't have capability to connect mobile unit, but they can propose their requests to the network like MU. Each workstation is capable to connect MU is equipped with a wireless communication to make a connection. Workstation is also known as a Mobile Support Station (MSS), because they act as a connector between mobile computers and fixed hosts. Geographical area which has been covered by is called cell. In workstations, information like users' identification file, network log files and data access rules are kept.

Mobile architecture due the cell which is now located, it only can connect to one MSS. Based on each MU; MU can change its connective and location position with network moving in wireless network. During this movement, MU might be running an application. Running this application/ transaction requires data which have been obtained by MU through connection with MSS. As a result of MU movement, it might be disconnected with the network for a moment. After reconnection, it might be inserted into a cell belongings to MSS. Running an application /transaction by mobile unit can be done in 2 ways:

- Doing transaction on mobile unit
- Doing transaction on fixed network

3. CACHE OPERATION

In mobile database systems, caching can play a considerable role in decreasing connection with services givers and takers. In a mobile computing environment, data are kept on mobile support stations (MSS). Primary control of data and doing any kind of operation over data are investigated by MSSs and after verification, it is recorded in database. Caching data

items which are repeatedly and highly being used by MUs can be an effective method to decrease connective costs related between MSSs and MUs and also to use mobile units' resource better. MUs can cache information in their database which is similar to MSSs' database. In this way, accessibility to data required in MU is being done faster and number of requests compared with data receiving in network is decreasing. When a mobile unit requests for a data, if this data is available, it can be used and results will be sent to the service giver, otherwise if required data is unavailable or a part of it is available, mobile unit sends request for required data to service giver [6].

4. MEMETIC ALGORITHM

Gene is a part of biological information which is transferred from a generation to another generation. Genes determine physical features such as face, body shape and all of features which have been inherited from their parents. Meme has been introduced in 1979 by Dowkins with a gene comparison in the context of genetic evolution and in the context of cultural evolution [7, 8]. Meme is a cultural and behavioral element which is being transferred from a generation to another generation via non-genetic factors. In fact, Meme includes each feature and characteristic which has been learnt through experience and imitation during a life of a creature and it is propagated among organisms. This propagation has never had a genetic nature and reproduction operators don't have impact on it. Lingual definition of Meme is: apart of civilization which are not inherited through genes. As biologists consider gene as a unit to transfer physiologic features such as eyes color, hair color and etcetera from parents to children, psychologists also know gene as a unit to transfer behavioral characteristics such as Irritability, traditionalism and etcetera from parents to children. Based on psychologists' viewpoints, a person who is born in an illiterate family, doesn't have to stay illiterate to the end of his life and he can be promoted with gaining some of skills form his environment while biologists consider chromosomes genes unchanged from the birth to the death. Basis of Memetic algorithm (MA) is based on this theory. Unlike genetic algorithm which considers people unchanged from birth to death (participating in reproduction process for the next generation), a person in Memetic algorithm can promote his competency in a generation with behavior known mimic [9].

Idea relevant to use of Meme conception to design a Meta-heuristics algorithm was firstly introduced by Moscato [10]. He recommended that with applying a local searching operator in the genetic algorithm body (after mutation operator), an opportunity is given to each organism (child) to live. Memetic algorithm has lots of similarities with genetic algorithm. The underlying difference of Memetic algorithm with Genetic algorithm is optimization of population of each generation after doing mutation and combination operators. To achieve this purpose, for each person who is living in the population, a local search with predetermined neighborhood radius around chromosome relevant to the produced child is done in problem state space.

Memetic algorithm like genetic algorithm has been employed to solve problems related to continuous and discrete optimization and a wide range of real world issues [11].

5. MYTHOLOGY

In database architecture, each MU in order to use data existed in database must be connected to one of MSSs and gives it its request and receives its required information from it. Of

course, sending and receiving data by MU is done via wireless network. In this architecture, database has been dispensed to MSSs. Of course, each of mobile units has a structure similar to the main database which is being used for information caching and local data reception. Each MSS is responsible to manage connection with MUs. Also, MSSs are responsible of accuracy of conducted transactions and recording transactions' results in database. When MU is connected with MSS, it is responsible to create an agent to control caching and managing MU's relocation. In figure 2, agents within MSS have been shown.
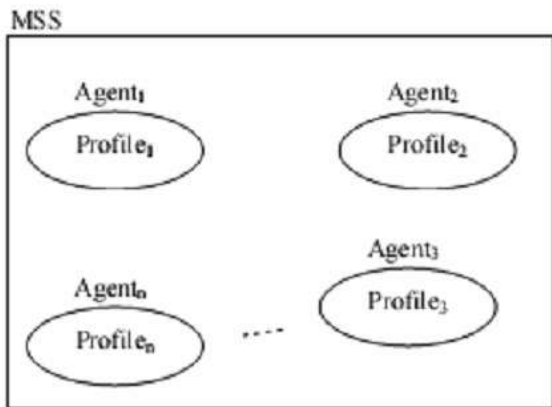

Figure 2: Agent within MSS [2]

In fact, for each MU, an agent is created. Having each agent within MSS shows MUs which have been connected to that MSS. Agent is applied to manage MU mobility. That is, with moving MU from a MSS to another MSS, all of information of agent related to this MU is transferred from origin MSS to the new MSS and new MSS is controlled via this agent and will take authority to manage MU.

Each agent is the maintenance of a profile of information which is being used to predict data which have to be cached. Agent can share their information to other agents. Of course, this action is being done to predict information required for the future. Profile information of each agent is created and completed by agent and based on data which has been used by MU so far and agent per se makes decision to eliminate these data.

Agent has to response each request from MU and record information relevant to profile completion in profile. Then, agent takes an effort to anticipate information required for MU in the future and after completing anticipation, sends obtained information to MU. In this process, 2 following stages are being done:

- Determining MU neighbors
- Findings information category related to the data which is being followed by user more and then choosing data having reasonable relation with used information.

How to perform these two steps will be described as following

5.1 Operation of Mobile site removing

Each MU when is connected to MSS, if it is a first time which connects to MSS, MSS creates a new agent for this MU and prepares agent to control and manage MU. While creating agent, in order to help MU identify its agent, MSS gives each agent a unique number. This number is unique in other MSSs. This number is seen in relation (1).

Agent ID=MSS Number+(Count Agent+1)   (1)

Here, MSS Number is service giver number, because each service giver is assigned a unique number. Count Agent determines number of MSS agents. This number is incremented one unit for new agent. After determining agent number, a unit is added. After determining agent number, this number is sent to MU and MU is able to identify its agent based on this number. From now on, while sending request to MSS, request is given to the respective agent and agent makes decision to send information to MU and with each request, it completes its profile to make better decision. If MU is disconnected with MSS, after reconnection, first MU must identify MSS and its agent. For this purpose, MU sends identification number of its agent to MSS. If MU is not evicted from scope of previous MSS, MSS identifies number and connection is established. If number is unknown for MSS, new MSS is responsible to determine previous MSS via identification number and sends a request based on receiving information related to that agent to the previous MSS. New MSS, after receiving agent information, changes agent identification number based on its own number and sends this number to the MU entitled new number and makes connection between MU and its agent.

5.2 Profile structure

Based on aforementioned contents, agent has profile which includes information to predict data required for the future. Information of this profile is created based on earlier used data. Each profile has a structure as table 1.

Table 1: profile structure (gene) of A chromosome agent

Agent(A)	
T_C	Table code
R_C	Record code
T_R	Request time
W_R	Request weight
W_U	Usage weight
Vote	Vote

So that:

A: From here, mobile unit a place where anticipation is done for that is determined with A.

Table code (TC): Here, for each table in database structure, an address is considered and TC shows a table which its record has been used.

Record code (RC): Based on definitions of database, each record (or entity sample) must have unique address named key. Here, is RC being used.

Last request time (TR): It is the last time which record agent sends RC from TC to MU. This time for the record determines which record has been newly used or hasn't newly

used and agent eliminates records in the profile which have a certain difference with the current time, because these records are probably not interesting to A and are not being used.

Usage weight (WU): It shows number of times which MU uses RC from TC and changes/ updates this record. Difference between WU and WR is that a user requests for a record, but doesn't change it.

Vote: Agent, with having its profile information, can determine records which have been used more than others (based on request time and usage weight). To reach this purpose, a coefficient as an importance is assigned to each of above values as following:

- Sending time $\times$ 0.6
- Request weight $\times$ 0.25
- Usage weight$\times$ 0.15

Now, vote relevant to profile's ith record is calculated as following:

$$Vote_{(i)} = (Tr \times 0.6) + (Wr \times 0.25) + (Wu \times 0.15)$$

In fact, each time that information is sent to MU, some of information of profile (Wu, Wr, Tr) are being changed and value of vote relevant to each profile's line with changes of that line are again obtained. Value of vote can display opinion/ interest of users about RC. With creating agent and profile relevant to the A mobile unit, agent completes receiving of each request from MU profile and tries to determine working scope of MU.

5.3 Defining and implementing Memetic space:

To use Memetic algorithm, the most important part is to define and implement genetic space which has been expressed as following:

1- Population: All of agents in all of MSSs connected to the wireless and fixed network are considered population.
2- Gene: Records which are located in tables of each agent are called gene which is equivalent to profile structure.
3- Chromosomes: Each agent is considered as chromosome.
4- Crossover: Finding a packet of information (records), the most optimal information to anticipate and send to MUs, is called crossover.
5- Fitness function: A function which is used to fit the best records to anticipate.
6- Local search: It causes that problem is explored sooner.

5.4 Prediction

As it was expressed, prediction algorithm is two-step algorithm. How to do these 2 steps and prediction algorithm have been illustrated as following

5.5 Neighbors selection

Here, neighbor doesn't mean near MUs in physical and place terms, but also it means closeness and adjacency in terms of information and data which have been used by respective MUs and other similar MUs. As it was discussed before, chromosomes can give their gens' information to other chromosomes. This issue is also true for chromosomes in MSSs. This can be done through MSSs. To determine

neighbors, genes of other chromosomes have to be investigated and choose the nearest neighbors, because with the number of chromosomes in MSSs, comparison is being done. After selecting chromosomes, comparison between 2 genes can be done using Euclidean distance (12) (of course something that is changed) which is working on profile records. Euclidean (A, B) is similarity of active user A and user B.

$$Euclidean(A,B) = \sqrt{\sum_{i=1}^{k} TR_{A,i} \times (Vote_{A,B} - Vote_{B,i})^2}$$

So that:

K: It determines number of similar records which have been used by A, B chromosomes.

$TR_{A,i}$: It is a time that chromosomes A has sent record i for MU.

$Vote_{A,B}$: It is a vote which has been computed based on gene of ratio of chromosome A usage from gene B.

Time to run this equation is for n O(nk) chromosomes which is K=Average($K_1,K_2,…,K_n$), here that K_i shows number of genes similar to chromosome A and chromosome i. Obtained distance is more or equal to zero. Chromosomes which have the least distance are required and out of N selected chromosomes, M chromosomes which have the least distance can be selected to be used to predict.

5.6 Prediction algorithm

Prediction is being done when MU proposes a request and information needed for this request are existed in MU. In this case, MU sends request to the respective MSS to obtain its required information. In this time, chromosomes existed in MSS has duty to predict other information in addition to information required for requested fixed genes and sends to MU. To predict, we can:

Based on gene information, record number (RV) and table number (TC) relevant to the line which has the most votes are obtained.

Among genes in selected chromosomes, those which are requested are eliminated, because they have previously located in new crossover chromosome.

After determining genes which don't have to participate in crossover (those which have been selected before), the nearest neighbor has to be determined as it was mentioned above. Now, genes which have been sent to MU have to be used by more number of nearest neighbors and also haven't been used by MU so far (they have been the best genes in fitness function).

After completing first selection operation, we have to select next gene which has the most value and perform Memetic algorithm for that gene. A certain number of Votes or just Votes which are more than a certain limit can be considered.

6. CONCLUSION

In researches which have been called above, only information are cached which are needed now and no more prediction is being done for information which will probably be used in the future. In this way, one of the best optimization

algorithm (Memetic algorithm) has been used which can guarantee an appropriate efficiency during a certain time

7. REFERENCES

[1] Tarafdar, M , Hagh joo Saniij, M. location prediction query processing in mobile database to discover patterns in timed sequence. Tabriz; 2011. p. 234-41.

[2] Saqayshy Mohammadi, M, Habibi Zadnovin, A., F. Mohammadi Saqayeshy, M. Using fuzzy database to predict and improve operations in mobile databases. Science and Research, East Azerbaijan; 2011. p. 65-71.

[3] Xuan K, Zhao G, Taniar D, Rahayu W, Safar M, Srinivasan B. Voronoi-based range and continuous range query processing in mobile databases. J Comput Syst Sci. 2011;77(4):637–51.

[4] Srikant R, Agrawal R. Mining Sequential Patterns: Generalizations and Performance Improvements [Internet]. Springer; 1996 [cited 2014 Oct 2]. Available from:
http://link.springer.com/chapter/10.1007/BFb0014140

[5] Fernando N, Loke SW, Rahayu W. Mobile Cloud Computing: A survey. Future Gener Comput Syst. 2013;29(1):84–106 .

[6] Yavaş G, Katsaros D, Ulusoy Ö, Manolopoulos Y. A Data Mining Approach for Location Prediction in Mobile Environments. Data Knowledge. 2005; 5 4(2):121–46 .

[7] Goldberg DE. Genetic Algorithms in Search, Optimization, and Machine Learning. 1 edition. Reading, Mass: Addison-Wesley Professional; 1989. 432 p .

[8] Mitchell M. An Introduction to Genetic Algorithms. Reprint edition. Cambridge, Mass.: A Bradford Book; 1998. 221 p .

[9] Memetic algorithm [Internet]. Wikipedia, the free encyclopedia. 2014 [cited 2014 Oct 28]. Available from: http://en.wikipedia.org/w/index.php?title=Memetic_algorithm&oldid=626348976

[10] Moscato P, Cotta C, Mendes A. Memetic algorithms. New Optimization Techniques in Engineering [Internet]. Springer; 2004 [cited 2014 Oct 28]. p. 53–85. Available from:http://link.springer.com/chapter/10.1007/978-3-540-39930-8_3

[11] Neri F, Cotta C, Moscato P, editors. Handbook of Memetic Algorithms. 2012 edition. S.l.: Springer; 2013. 370 p .

[12] Euclidean [Internet]. Wikipedia, the free encyclopedia. 2014 [cited 2014 Oct 28]. Available from: http://en.wikipedia.org/w/index.php?title=Euclidean&oldid=630468632