EFFECTIVE AND EFFICIENT MINING OF DATA IN MOBILE COMPUTING

C.Ashok BabuRaj, Senior Lecturer, Dept of MCA,
KLN College of Engineering, Sivagangai, India.
email – ashokbaburaj@yahoo.com

S.Thabasu Kannan, Head, Dept of Computer Science,
Sourashtra College, Madurai, India
email - thabasukannan@gmail.com

ABSTRACT

Data Mining consists of an evolving set of techniques that can be used to extract valuable Information and knowledge from massive volumes of data. Data Mining Research and tools have focused on commercial sector applications. This Research paper highlights the data mining techniques applied to mine for Location and Mobility Management. Location Management is a very important and complex problem in today's mobile computing environments. There is a need to develop algorithms that could capture this complexity, yet can be easily implemented and used to solve a wide range of Location Management scenarios. In the Reporting cell Location Management scheme, the mapping is done on the basis of classification. Some cells in the network are designated as Reporting cells; Mobile terminals update their positions (Location Update) upon entering one of these reporting cells. The remaining cells are designated as Non-reporting cells.

In the proposed scheme, an supervised learning technique is being introduced which maintains the history or mobility pattern (of size h) of the last visited reporting cell. The updation does not take place, when the user roams with in the reporting cells of his mobility pattern. The location management is updated when the user enters in to the new reporting cell, which is not in his history.As a result the updation cost is proportionately reduced with the value of h (the number of entries in the history) Artificial life techniques have been used to solve a wide range of complex problems in recent times. The power of these techniques stems from the capability in searching large search spaces, which arise in many combinatorial optimization problems very efficiently. To create such a planner, Genetic Algorithm has been implemented to show that the total cost is less when compared with the existing cost based updation and searching scheme.

Key Words – *Location Management, Genetic Algorithm, Mobility Management*

INTRODUCTION

The goal of mobility tracking or location management is to balance the registration and search operation, to minimize the cost of mobile terminal location tracking. Two simple location management strategies known are the *always-update strategy and the never-update strategies*. In the *always-update strategy* each mobile terminal performs a location update whenever it enters a new cell. As such the resources used for location update could be high. However no search operation could be required for incoming calls. On the other hand in the *never-update strategy*, no location update is ever performed. Instead, when

(Advance online publication: 12 November 2006)

a call comes in, a search operation is conducted to find the intended user.

One of the common location management strategy used in existing systems today is the *location area scheme*. In this scheme, the network is partitioned into regions or location areas (LA), with each region consisting of one or more cells (figure.1). For example , in Figure 1 , if a call arrives for user X, search is confined to the 16 cells of that Location Area.

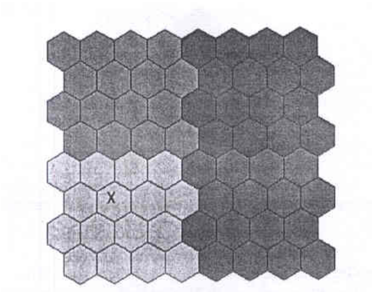


Figure 1

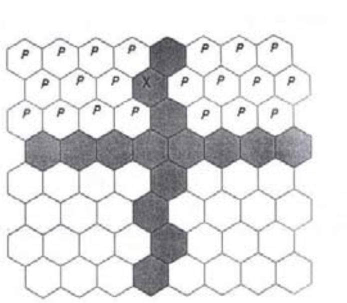


Figure 2

In the above Figure 2, it represents the network with reporting cells(shaded areas represent reporting cell). **R**egions representing Location Areas(LA) and individual cells(in this figure there are four LAs each consisting of 16 cells). Another location management scheme is a subset of cells in the network is designated as the reporting cells(figure.2).

Location Management Cost

The total cost of the two cost components (location update and cell paging) over a period of time T, as determined by simulation can be averaged to give the average cost of a location management strategy. The following simple equation can be used to calculate the total cost of a location management strategy [3].

$$Total\ Cost = C.N_{LU} + N_P$$

Where N_{LU} - the number of location updates performed during time T

N_P - the number of paging performed during time T and C is a constant representing the cost ratio of location update and paging. It is recognized that the cost of a location update is usually much higher than the cost of paging- several times higher.

Network Structure

The Data set that has mined for N X N Network in Cell planning problem are movement weight, Call arrival weight, Vicinity, Vicinity value of a cell. When calls arrive for a user, the user has to be located. Some cells in the network, however, may not need to be searched at all, if there is no path from the last location of the user to that cell, without entering a new reporting cell (a reporting cell i.e. not the last reporting cell the user reported in).That is, the reporting cells form a “solid line” barrier, which means a user will have to enter one of these reporting cells to get to the other side

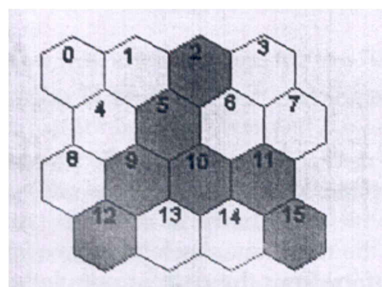


Figure.3 Network with reporting cells (shaded areas represent reporting cells)

For example in Figure. 3, a user moving from cell 4 to cell 6 would have to enter a reporting cell. As such, for location management cost evaluation purposes, the cells that are in bounded areas are first identified, and the maximum area to be searched for each cell is calculated which is described below.

We can define the *vicinity* of reporting cell i as *the collection of all the cells that are reachable from a reporting cell i without entering another reporting cell as the vicinity of reporting cell i.* We can define the *vicinity value* of reporting cell i that as *the number of cells in the vicinity of a reporting cell i is the maximum number of cells to be searched, when a call arrives for a user whose last location is known to be cell i.*

As an example, in figure.4, the vicinity of reporting cell 9 includes the cells 0,1,4,8,13,14 and cell 9 itself .The vicinity value is then 7, as there are seven cells in the vicinity. Each non reporting cell can also be assigned a vicinity value. For example, in figure. 4 , a cell 4 belongs to the vicinity of reporting cells 2,5,9 and 12, with vicinity values 8,8,7 and 7 respectively.

History based Mobility Management system

In the proposed scheme, an supervised learning technique is being introduced which maintains history or mobility pattern (of size h) of the last visited reporting cell.

The updating does not take place when the user roams within the reporting cells of his mobility pattern. That is the location information is updated when the user enters to a new reporting cell, which is not in his history. As a result, the updation cost is proportionately reduced with the value of h (no. of entries in the history). When we increase the number of reporting cells in the history, the location update cost is proportionately reduced. Hence the cost equation can be modified as follows:

$$N_{LU} = \sum_{i \in s} NW_{mi}$$

Where NW_{mi} - the new movement weight.

$$NW_{mi} = W_{mi} * \frac{(S-h)}{(S-1)}$$

Where h - the number of reporting cells maintained in the history.

Here if we keep h=1, the NW_{mi} tends to W_{mi}.
By increasing h value the NW_{mi} will be reduced, as a result the updating cost is reduced.

Consequently, the paging cost gets increased proportionately to the h value. But it can be kept under control by using the following technique. Whenever the user enters into the reporting cells of his previous history, the mobility pattern in the mobile is modified and does not leads to the location update.

The *new paging* cost equation is obtained from the new call arrival weight.

$$NW_{cj} = \left[\frac{W_{cj}}{W_{mi}} * NW_{mi} \right]$$

Search Cost for the Location updated users: N-1

$$N_{P1} = \sum_{j=0} (NW_{cj}) * v(j)$$

Search Cost for the non-updated users from the same reporting cell:

$$N_{P2} = \sum_{j=0}^{N-1} (W_{cj} - NW_{cj}) * v(j) * 1/S$$

Search Cost for non-updated users from different reporting cell: (first call)

$$N_{P3} = \sum_{j=0}^{N-1} \frac{(W_{cj} - NW_{cj})*(S-1)/S*v(j)*h/2}{\text{Call factor}}$$

(Subsequent calls)

$$N_{P4} = \sum_{J=0}^{N-1} (W_{cj} - NW_{cj})* (S-1)/S * (1-1/\text{Call factor}) * v(j)$$

The *call factor* can be calculated as follows, if $(W_{cj} / W_{mi}) < 1$ then call factor =1 else call factor = (W_{cj} / W_{mi}).

Total Paging Cost :

$$Np' = N_{p1} + N_{p2} + N_{p3} + N_{p4}$$

$$\text{Total Cost} = C.N_{LU'} + N_{P'}$$

The total paging cost is divided into four sub components and except the third component (Np3) the other component costs are similar to the old method. As a result the increase in the paging cost is under control; hence it improves the total cost for the reporting cell configuration. The cost function described above shows that by varying the size of the mobility pattern (h), the total cost can be reduced to some extent. If h=1 then this cost is equivalent to the old cost. So in the worst case it behaves like the old scheme and in best case, we can introduce h value so that the entire cost is reduced.

The problem of marking the cells in the network as reporting cells and deciding the value of h can be seen and solved using one of the artificial life cycle Techniques like Genetic Algorithm. By generating the number of generations and history value the hidden information that has been derived for various data sets are the total cost (updates cost and search cost).

Genetic Algorithm

Each binary gene, which represents a cell in the network, can have a value of either “0”, representing a non reporting cell, and a “1”, representing a reporting cell. Thus, *high-fitness* (``good") individuals stand a better chance of ``reproducing", while *low-fitness* ones are more likely to disappear.

```
Begin G:=0 (generation counter)
    Initialize population P(g)
    while not done do
    Evaluate population P(g)
    Compute Objective function
    Compute Expected Count
    Reproduce P(g)
    Crossover P(g)
    Mutate P(g)
    g:=g+1;
  end while  end GA
```

RESULT – 1

DATA SET FOR 4 X 4 NETWORK

NUMBER OF GENERATIONS=500

HISTORY VALUE	DATA SET 1	DATA SET 2	DATA SET 3	DATA SET 4
1	6075.8	7695.8	9560.6	11210.6
2	5787.8	7282.6	8804.1	10342.8
3	6477.3	7721.7	9401.0	11130.6
4	6859.3	8002.8	9695.5	11431.9
5	6616.2	7959.9	9692.4	11527.3

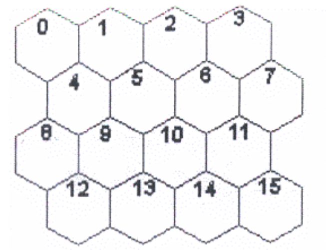
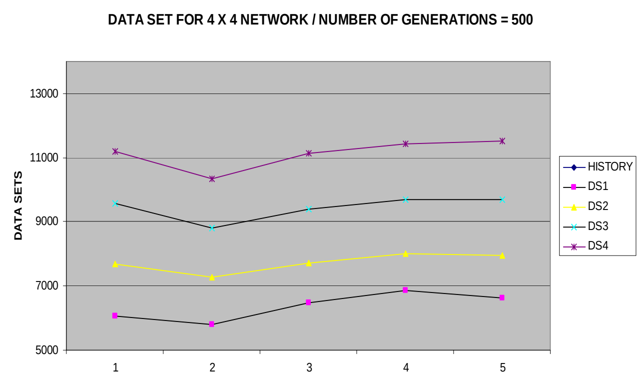


Figure 4 – Network Of Size 4 X 4 Cells

DATA SET 1 FOR A 4 X 4 NETWORK

CELL	WCI	WMI	CELL	WCI	WMI
0	317	318	8	51	245
1	373	574	9	24	1149
2	55	53	10	641	1458
3	107	1496	11	500	752
4	442	1417	12	25	107
5	751	272	13	340	185
6	326	450	14	495	1146
7	309	169	15	125	372



CONCLUSION

The paper *Effective and Efficient Mining of Data in Mobile Computing* suggests a modification in the existing location management scheme by maintaining a history of the user mobility pattern. The solution to the reporting cell planning problem can be improved by using the user mobility pattern and this concept is proved by implementing through Genetic Algorithm.

REFERENCES

1. R. Subrara & A.Y. Zomaya, ”A comparison of three artificial life techniques for reporting cell planning in mobile computing”,IEEE trans. On parallel and distributed systems, vol 14, no. 2, Feb 2003
2. R.L. Gondim, “Genetic Algorithms and the location area partitioning problem in cellular networks”, proc. IEEE 46th vehicular technology conf.,1996
3. R. Subrara & A.Y. Zomaya, “Location Management in mobile computing”, proc. ACI/IEEE int’l conf. computer systems and applications, 2001
4. Jochen Schiller “Mobile Communications”, Addison – Wesley publications.
5. William C.Y.Lee “ Mobile Cellular Telecommunications “ Second Edition MC graw hill International Editions.
6. D.E.Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning Mass :Addison – Wesley, 1989.
7. Histrory based location and mobility management, M.Phil dissertation
8. Intelligent optimized technique for reporting cell planning, JMC proceedings.
9. Genetic technique for reporting cell planning in Mobile Computing, GRI Proceedings.