

Using Predictive Prefetching to Improve Location Awareness of Mobile Information Service

Gihwan Cho

Division of Electronic and Information Engineering, Chonbuk University,
664-14 Duckjin-Dong, Duckjin-Gu, Chonju, Chonbuk, 561-756, S. Korea
ghcho@dcs.chonbuk.ac.kr

Abstract. Mobile information services have to provide some degree of information adaptability onto the current service context, such as mobile user's location, network status. This paper deals with a predictive prefetching scheme for reducing the latency to get refreshed information appropriated on current location. It makes use of the velocity mobility model to exploit location knowledge about the terminal and/or user's mobility behavior. With considering the user's moving speed and direction, the prefetching zone has been proposed to effectively limit the prefetched information into the most likely future location context whilst to preserve the prefetching benefits. The proposed scheme has been evaluated with a simulator in the context adaptability point of view.

1 Introduction

With prosperity of hardware technology, computing devices are getting small and higher performance. Moreover, by cooperation of the wireless communication with hand-held computer, computing users can access their own information anytime, anywhere, even they move around. Even if many existing computing services may be logically extended into this new coming computing paradigm, its inherent characteristics, such as user, terminal and/or service mobility, require new service methodologies. That is, mobile information services have to provide some degree of information adaptability onto the current service context, where context is defined as the characteristics of the user's situation like as current location, and the user supporting environment like as network status (bandwidth, broadcast etc.) and mobile terminal capabilities (size, color etc.). This feature is generally formalized as context-aware mobile information services [1].

In the user and/or terminal mobility environment, location context may be changed whilst the user is still interested in the corresponding answer. Therefore, it is required to be adaptively provided a response appropriate to the current location: a previously provided response must be refreshed, if the corresponding request is still active after a

This work was supported in part by University Research Program of Ministry of Information and Communication in South Korea, and Center for Advanced Image and Information of Chonbuk National University in South Korea.

change of location that invalidates the old response [2]. Thus, the latency to get refreshed information appropriated on the current context is the most critical service parameter, especially in the frequently context changed environment. With the long latency, the user may suffer the long time duration provided with information not appropriate for the current location. To reduce the latency, a well-established technique, prefetching, can be adapted to fetch information in advance, usually based on what contexts are highly utilized [3]. Thus, it is very important to make use of some degree of knowledge to limit the amount of prefetched information; otherwise, unnecessarily prefetched information can overwhelm the benefit of latency reduction with communication and memory cost. Therefore, prefetching scheme is usually based on some sort of prediction strategy about the information that will be needed in the near future.

The predictive prefetching concept has been applied and evaluated to reduce the latency in retrieving a Web document on the wired environment, based on the user's navigation pattern [4]. It shows that the predictive prefetching reduces significantly in the average access time. The work in [3] provides the effectiveness of prefetching policies in the location-aware mobile information services. The work considered only the uniform mobility pattern based on the Markov model, and did not consider how to effectively confine the prefetching information to reduce communication cost. In the reference [5], a semantic based caching scheme has been presented to access location dependent data. The work is just centric to the cache information replacement strategy to utilize the semantic locality in terms of location.

In this paper, a predictive prefetching scheme is proposed by exploiting user movement knowledge to limit the prefetching to the most likely future contexts. Because the location is the most distinguished context for the context aware mobile information system, our discussion is limited in location context after this. The rest of the paper is organized as follows. In Section 2, a general architecture of mobile information service is presented, along with the problem definition. In Section 3, a velocity based mobility model is presented and a predictive prefetching scheme is formalized with the mobility model. Using a simulation, a numerical result is shown in Section 4. Finally a conclusion is given in Section 5.

2 Mobile Information Service

For mobile information service, a general architecture has been utilized to describe the location-aware feature [2], [3], [5], [6], [7]. The architecture consists of 3 logical components; mobile device that possibly moves around with a user, information server that would be located somewhere in Internet, and location-aware service manager that acts as intermediary between the mobile device and the information server. The manager performs a sequence of operations to retrieve information associated with the current location from the server, and to load it on the user device. To do this, it periodically monitors the current user location, and checks whether it belongs to the current information scope. If an out-of-scope condition is detected, the service manager restarts to retrieve the corresponding information on current location. A mobile user is assumed to equipped with a certain mechanism to obtain its current location,

such as the GPS or indoor infrared sensors.

For a given mobile information service, the set of possible location contexts is assumed to be divided into separated subsets, such as rooms, areas. Then a separate piece of information is associated to each subset, such as equipment available in a room, map of a given area. A unit area, which a user may move around with the same information piece, is divided into adjacent geographical portions. Each geographical portion has corresponding portion information that must be provided by the information service as response to a user's request within that geographical portion. Geographical portions are assumed to be equal shape represented as quadrangles. A discrete time model is used for a user's movement. At the end of each time slot, the user can remain in the same quadrangle, or move to an adjacent quadrangle through one of the shared edges. At a certain point of time, the choice of remaining inside or moving outside the quadrangle of a user is just depended on the current location and the moving direction and speed.

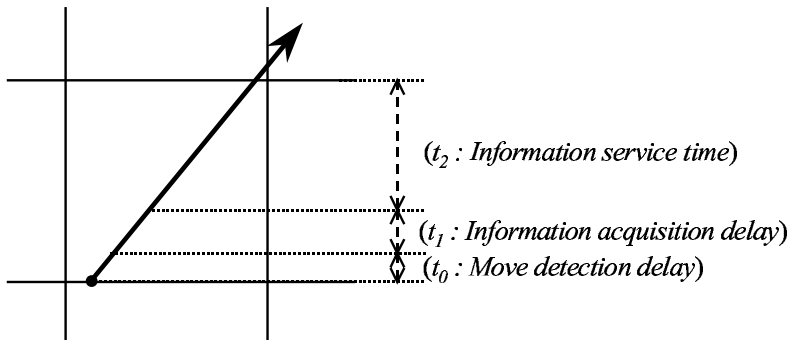


Fig. 1. Mobile information service delay on a given geographical portion

Fig. 1. shows the delay to get corresponding information for a given geographical portion (the quadrangle stands for a geographical portion). On crossing a geographical portion, the latency can be classified with two time factors; the delay needed to detect that the provided portion information is out of its current context, so t_0 : move detection delay, and the delay needed to retrieve and start loading the new information for current geographical portion, so t_1 : information acquisition delay. By reducing two delays, a mobile user can be blessed enough with the current context information, so t_2 : information service time. The effect of the first factor can be improved by reducing the time interval between two consecutive location checks. The most common technique to improve the second factor is to prefetch information that will be required in the near future. With the practical requirements such as communication and storage limitation of mobile device, the amount of prefetched information must be minimized as much as possible, whilst the prefetching amount reduction must not hurt the prefetching benefits.

3 Mobility Model and Predictive Prefetching

At a given location, the user's future location can be relatively represented by the moving direction and speed. Therefore, the moving behavior of mobile users can be now modeled from 3 parameters, current location, moving speed and direction.

3.1 Velocity based Mobility Model

To model both the direction and speed of a moving user, a velocity V is defined as a vector $\langle V_x, V_y \rangle$, where V_x and V_y are integers. A pair of values, V_x and V_y , is the projections of its speed vector on x axis and y axis respectively (in practically, each axis can stand for the east or the north respectively). The signs of V_x and V_y , and the ratio of the absolute value of V_x and that of V_y determine the directions of the movements. The model assumes that a user moves around in a 2-dimensional space at a speed that keeps constant during any observed time period and may change after a given unit time. So, at a certain point of time, when a user's current location has been given, the location in the future can be specified by considering the movement velocity. This concept has been mostly borrowed from [5].

3.2 Predictive Prefetching Strategy

Generally, a natural prefetching strategy is to prefetch the portion information corresponding to a set of geographical portions that are within a certain degree of outbound distance d from the current location. These geographical portions consist of a *prefetching zone* (shortly PZ). To form a PZ, one possible strategy is to include all the quadrangles within outbound distance d fraction (so, radius d) of rings 0, 1, 2, ... n [3]. No remote loading is necessary unless the user moves around within its current PZ. When the user enters a portion outside its current PZ, the geographical portion becomes the new starting position to build a new PZ. The new PZ may include the portion information for quadrangles within distance d ring.

Here, our concern is to exploit some degree of prediction about the user behavior, in order to be prefetched only the information for most likely future contexts. When a user crosses a PZ, the moving speed is applied to estimate the distance d , and the moving direction is utilized to decide the width w of PZ. Now, a PZ shapes as a rectangle [7], rather than a ring. This is based on that, if a user moves with high speed, it would be very reasonable to set the distance higher than that of low speed movement. Similarly, if a user's behavior tends to move with high degree of obliquity, it would be very reasonable to set the width wider than that of low degree of obliquity.

Our strategy to predictive prefetching is as follows. The distance d of a rectangle PZ is to get the upper bound integer of the user's speed, that is, the squared root of the sum of squares of two speed values V_x and V_y , as $\lceil \sqrt{V_x^2 + V_y^2} \rceil$. Thus, the signs of speed values should be ignored because the distance considers the speed characteristic only. In addition, if one of speed values is zero, it must be replaced with the other

non-zero value. Similarly, the width w of a rectangle PZ can be obtained with the upper bound integer of the mean value of the sum of two speed values, V_x and V_y , that is $\left\lceil \frac{|V_x| + |V_y|}{2} \right\rceil$. Note that the signs of speed values should be ignored because the direction reflects the degree of obliquity rather than that of dimension. Differently with the distance, zero valued speed will be ignored in the width. Fig. 2. shows an example for the predictive prefetching reconstruction idea for three different moves. The dark rectangles depict the new PZ that is constructed on each crossing of the PZ. For each moving with the velocity $\langle 2, 1 \rangle$, $\langle 3, 5 \rangle$ and $\langle 5, 0 \rangle$, the size of PZs is calculated as the distance $d = 3, 6$ and 8 , the width $w = 2, 4$ and 3 respectively.

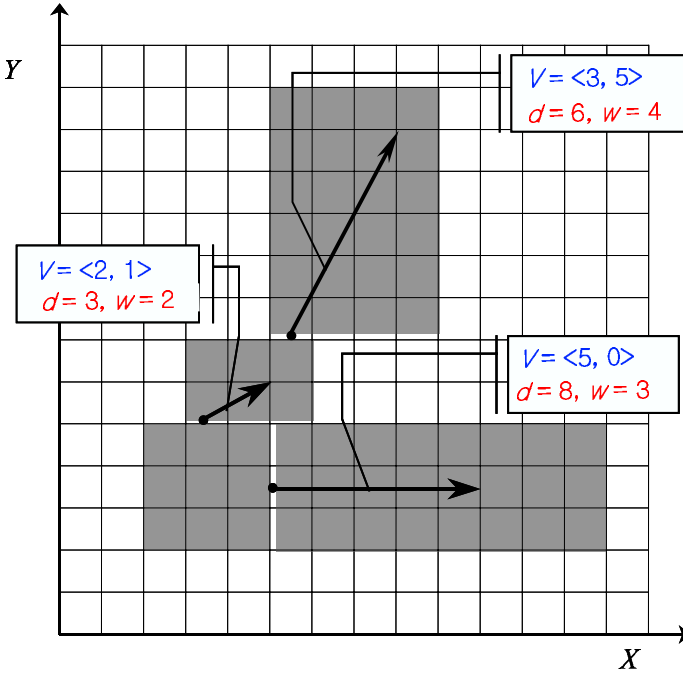


Fig. 2. A predictive prefetching reconstruction example

As shown in the Fig. 2., the gap between the PZs is inevitable. The location aware manager may start to build a new PZ and to prefetch the new PZ, just after it finds out the fact of crossing the current PZ, so the move detection delay in Fig. 1. Then, until the device is refreshed with information for the new PZ, so the information acquisition delay in Fig. 1., the mobile user therefore does not have any information appropriate to its current context. Due to the mobility feature, the user is in the out-of-scope condition of an effective information service. This situation can be improved with defining the edge geographical portions of a PZ into the *crossing zone*; it should be also defined with considering its moving speed and direction. When a user meets a geographical portion in the crossing zone, it starts to prefetch the new PZ in advance.

4 Numerical Result

In order to provide the effectiveness of proposed scheme, a comparison has been made between the ring PZ approach and the rectangle PZ one, through a simulation study. For simulation runs, a set of system parameters has been borrowed from other works, as shown in Table 1. Some numerical result has been measured for the amount of prefetched portion information, the ratio of utilization of the prefetched portion information, and the service time for end user. The first is significant to reduce the communication cost, the second stands for how much the prefetched portion information is made use of real information service, and the last is meaningful to configure the user's satisfaction for the service. The simulator has been implemented in C language using an event-based simulator CSIM [8].

Table 1. System parameters for simulation

Parameter	Description	Value
Wired bandwidth	bandwidth of the wired line	800 kbps - 1.2 Mbps
Wireles bandwidth	bandwidth of the wireless medium	80 kbps - 120 kbps
Move detection time	time to detect the fact a user accrosses a gegraphycal portion	100 ms
Residentiary time	mean time for a user staying in a gegraphycal portion	10 sec. (exponential distribution)
Portion size	information size which stands for a gegraphycal portion	1 MByte
Packet size	size of a unit data packet which is passing through the network	1440 byte

4.1 User Move Scenario

In the velocity based mobility model, a user moves around in a 2-dmenstional portion's geographical space at a speed that keeps constant during any given unit time period. Then it changes its speed and direction, and moves constantly within the unit time. Therefore, the model is much significant when a user's moving pattern is uniform. Practically, this is analogous to many real life cases, such as driving a car on highways or walking a street. A typical user move scenario has been utilized for the simulation. A user moves around 20 by 20 geographical portions, as shaped in Fig. 2. It starts from the position (10, 10), and moves 20 times, sequenced as (2, -4), (3, 2), (-3, 4), (-5, 0), (2, -5), (-3, -3), (5, -2), (0, 3), (4, -2), (-1, 3), (3, 3), (0, 5), (-2, 5), (-3, 0), (-5, -2), (5, -2), (-5, 0), (-2, 4), (-3, -6), (4, -5).

4.2 Simulation Experiments

With the given user mobility scenario, the number of prefetched portion information has been evaluated to show the effectiveness of prefetching strategies by comparing three different cases; prefetching with ring PZ, prefetching with rectangle PZ, pre-

fetching is not applied. The same simulation parameters are used for each case, whilst different strategies are used to build a PZ. For the prefetching with ring PZ, radius of the ring is a critical simulation parameter. In order to provide some degree of fairness with that of prefetching with rectangle PZ, the speed value $\sqrt{V_x^2 + V_y^2}$ is adapted for calculating the radius of ring PZ.

Fig. 3. shows the number of prefetched portion information for each cases. The prefetching with ring PZ has relatively large variation depending on the speed value. Thus, some of portion information would be overlapped between that of the previous PZ and that of the current PZ. After excluding these overlapped portions, 1,105 portions are totally prefetched for the ring PZ strategy, and mean number of prefetched portion information is 55.25 for each move. In the rectangle PZ scheme, the number of prefetched portion information is stable; total number of prefetched portion information is 335, mean number of prefetched portion information is 16.75. In the case of prefetching is not applied, each portion information has to be acquired for every geographical portions in turn, on the path of the user moves through.

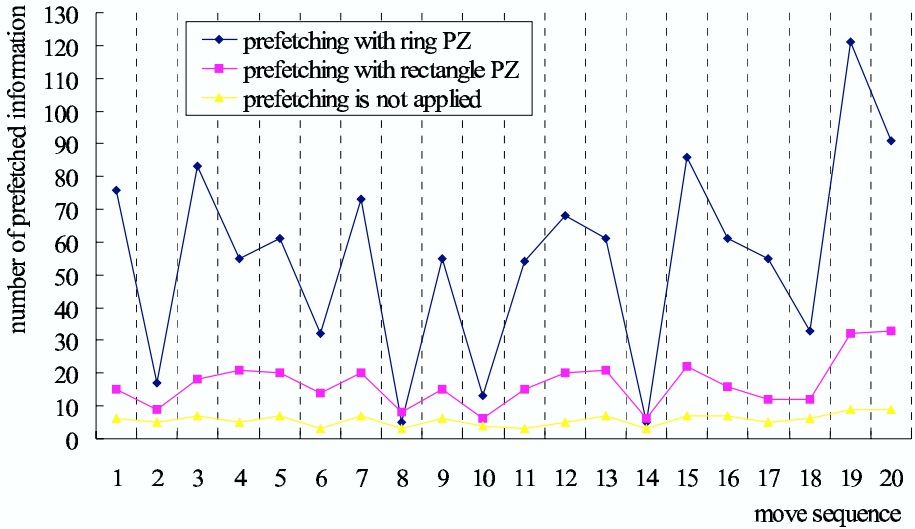


Fig. 3. The number of prefetched portion information

Now, it is meaningful to figure out how much the prefetched portion information would be actually participated for the user’s location-aware service. Fig. 4. shows the utilization ratio of prefetched portion information for two prefetching strategies, so the ring PZ and the rectangle PZ. Note that the utilization ratio is just depended on the user’s mobility pattern. With the given user move scenario, the utilization of rectangle PZ is much better than that of ring PZ; the former shows that over 30% of prefetched portion information is utilized for real service, whilst the later figures around 10%. We believe that the predictive prefetching with considering the user’s move behavior plays great role in building a prefetching method for mobile information service.

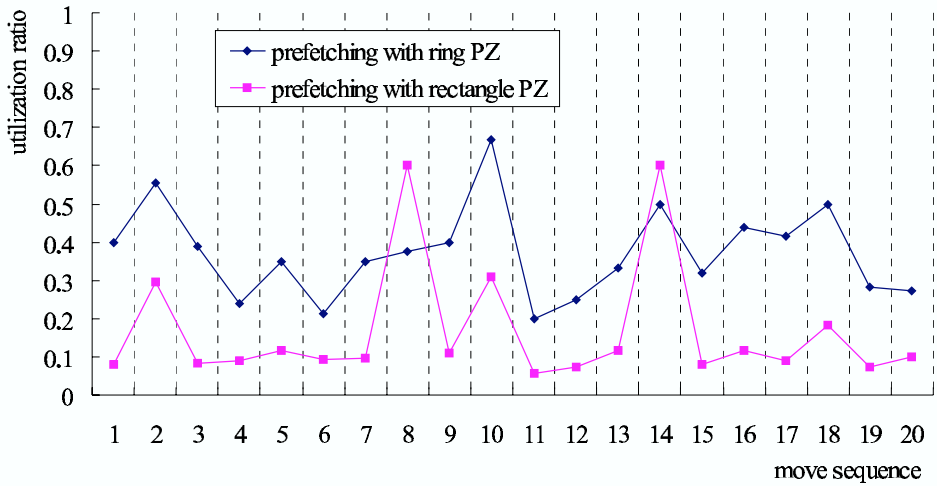


Fig. 4. Utilization ratio of the prefetched portion information

Fig. 5. shows the service time which has been spent to provide the corresponding information on each move. In the case of the prefetching is not applied, the time can be obtained by adding the service delay for getting each of portion information that the user has visited. So, mean service time is 570 sec. for each PZ. With the rectangle PZ strategy, the service delay is applied for only when a user crosses the current PZ. Once the service delay has been given on the portion, the other portion information within current PZ would be delivered in concurrently during the first portion information is serviced to end user. So the delay is the same for getting one of portion information on crossing PZ, and can be reduced as about 95 sec.

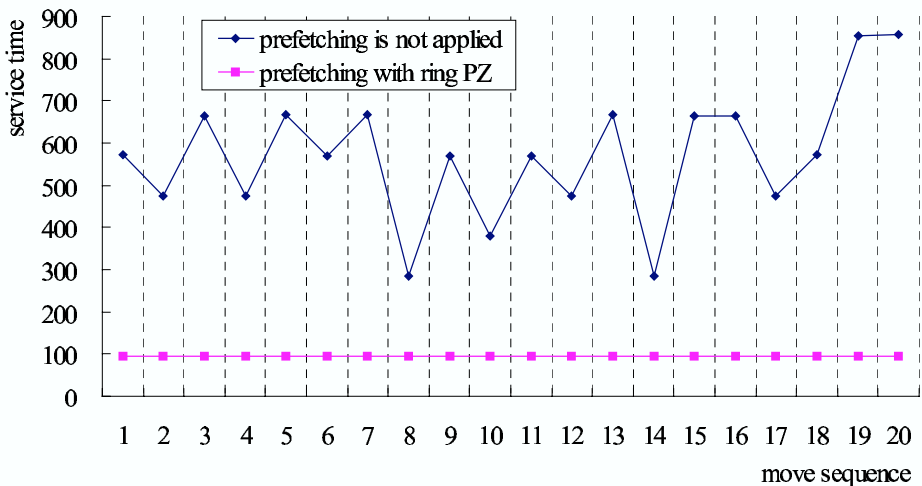


Fig. 5. Information service Time

5 Conclusion

In the location aware mobile information service, prefetching has a tradeoff between the amount of information transferred and the time delay to get the current context. In this paper we proposed a predictive prefetching scheme to limit the prefetched portion information to the most likely future contexts, whilst to preserve the prefetching benefits. This has been achieved by defining the prefetching zone with distance and width based on the user's moving behavior, that is, speed and direction. Using a simulator, the effectiveness of the proposed prefetching strategy has been measured with respect to the service effectiveness. The proposed prefetching with rectangle PZ was highly beneficial for both the prefetched data amount as well as the latency reduction. The idea in this paper is expected to be able to be extended to the other mobile server contexts, such as network bandwidth, mobile device type.

References

1. Brown, P. J., Bovey, J. D., Chen, X.: Context-aware Applications: From the Laboratory to the Marketplace. *IEEE Personal Communications*, Vol. 4, No. 5 (1997) 55-65
2. Couderc, P., Kernmarrec, A. -M.: Improving Level of Service for Mobile Users Using Context-Awareness. *Proc. 18th IEEE Symp. on Reliable Distributed Systems (1999)* 24-33
3. Persone, V. N., Grassi, V., Morlupi, A.: Modeling and Evaluation of Prefetching Policies for Context-Aware Information Services. *Proc. ACM Conf. on Mobile Computing and Networking (1998)* 55-65
4. Padmanabhan, V. N., Mogul, J. C.: Using Predictive Prefetching to improve World Wide Web Latency. *ACM Computer Communication Review*, Vol. 26, No. 3 (1998) 22-36
5. Ren, Q., Dunham, M. H.: Using Semantic Caching to Manage Location Dependent Data in Mobile Computing. *Proc. ACM Conf. on Mobile Computing and Networking (2000)* 210-221
6. Kovacs, E., Rohrlé, K., Schiemann, B.: Adaptive Mobile Access to Context-aware Services. *Proc. 1st Symp. on Agent Systems and Applications (1999)* 190-201
7. Kim, M. J., Cho, G. H., Cho, I. J.: A Prefetching Scheme for Context-Aware Mobile Information Services, *Proc. 3rd Conf. on Advanced Communication Technology (2001)* 123-125
8. Mesquite Software Inc.: CSIM18 Simulation Engine (1997)