COST OF LOCATION MAINTENANCE RELATED SIGNALING IN IP MICRO MOBILITY NETWORKS

László PIROSKA, Miklós A. RÓNAI, Zoltán R. TURÁNYI, András G. VALKÓ
{piroska, ronai, turanyi, valko}@ttt-atm.ttt.bme.hu

1. Introduction

Wireless mobile communication networks need more signalling to maintain routing state than traditional wired networks [1][2]. As a trivial solution, maintenance of routing state can be provided by continuous location tracking of mobile hosts. In this case mobile hosts update their location every time, when they move to a new base station; however, there can be extended periods of time when a mobile host has neither incoming nor outgoing traffic. During that time it is not essential for the network to know the mobile host's exact location. Since location updating consumes bandwidth and battery power, mobility related signalling should be minimized.

If a mobile host does not update its location at every new base station, the amount of signalling is decreased. On the other hand, in this case the network has to search for the mobile host when it has incoming traffic. To find the host, a group of base stations broadcast special *paging* messages over the air interface. If the host receives one of these messages, it updates its location information by answering the paging message. By combining continuous location tracking and paging, mobility related signalling can be effectively decreased.

In this paper we investigate, how the number of paging and location update messages can be minimized by controlling a mobile host's transitions between active and inactive states. In this discussion active state corresponds to the time periods where a mobile host updates the network at every movement and inactive state refers to the period where its location is only approximately known. We define *mobility cost* as an abstract value, which corresponds to the cost of mobility related signaling messages. Using a combination of analytical tools and simulation methods we analyze how mobility costs depend on key system parameters, such as the size of radio cells and paging areas, the speed of mobile hosts, the rules that govern a mobile host's transitions between active and inactive states, and the packet arrival process.

2. Location Maintenance, Active and Inactive States

To reduce mobility cost an algorithm is needed, which manages the duration of active and inactive states. The rules of location maintenance in inactive state have to be also defined.

One way is to approach the location maintenance of inactive mobile hosts is to group base stations into *location areas*. A mobile host in inactive state does not update its location, just when it changes location area; hence, the network is aware of only the location area of inactive mobile host, and not its exact base station. If the mobile host needs to be paged, every base station in the mobile host's location area broadcast the paging message. If the mobile host receives the paging, it updates its location, and becomes active.

Location maintenance incurs cost both in active and inactive states. In active state the frequent location update messaging results in significant mobility cost, while in inactive state the widely broadcast paging messages have high price To reduce mobility cost we have to carefully manage transitions between active and inactive states. If a mobile host sends or receives data packets, it shall be in active state to avoid frequent paging. In addition, we have found that it is worth to keep the mobile host in active state for a short time after data transmission. We call that time interval, while mobile host remains active without data traffic *additional active time*. If the additional active time is longer than the packet-inter-arrival time, one paging operation can be saved. Unfortunately the additional active time also contributes to mobility cost, as during that time the mobile host must update its location on each handover. If this results in higher mobility cost than a paging process, it is better to transit the mobile host into inactive state, and perform a paging if necessary. The good choice of the length of additional active time may significantly reduce mobility cost.

In our investigation we used a two level on-off traffic model to describe the data traffic of the mobile hosts. The upper level consists of *no-communication* and *communication* periods. During the no-communication periods mobile hosts do not transmit or receive packets. During the communication periods packet-groups folow each other with a certain amount of silence on the lower level of on-off activity. We assume thet the inter-arrival time between packets inside a packet group is negligibly short.

3. Analytical Tools and Results

To approximate the optimal additional active time length we used a mathematical model, which builds on our IP micro-mobility network model as well. In this paper we do not discuss our model in details due to space constraints, the interested reader is referred to [3]. Our aim is to describe our initial assumptions, and to give an overview of our model with measured results and conclusions.

For analytical tractability we assume, that packet-group length and inter-arrival time can be described by an exponential distribution. We assume also, that the cost of paging and location update messages are equal. The mean mobility cost within one second with various packet-group inter-arrival time and additional active time lengths can be seen on fig. 3.1. Here the cost of paging and location update messages is 1 credit, and one second of additional active time is 0.2 credit corresponding to one update message in five seconds for each active mobile host. The ‘Cost’ axis represents the mobility cost of a mobile host in 1 second, the ‘T_a’ axis stands for the length of additional active time in seconds, and ‘T_p’ represents the length of packet-group inter-arrival times. We took care that the absolute of the traffic on the long run is the same for all T_a T_p pair on fig. 3.1. The lowest mobility cost points for given packet-group inter-arrival times are marked by small circles on the figure.

As it is shown on fig. 3.1. the mobility cost extremely increases, if both packet-inter-arrival time and additional active time are low. If the additional active time is longer, the mobility cost first decreases, and after a local minimum it increases again. It means, that it is worth keeping mobile host in active state for a limited time after the end of a packet group, but not for long. The optimal additional active time (marked by circles) depends on the inter-arrival time. Since the inter-arrival time is unknown in a real network, fixed additional active time value would result in sub-optimal operation. We have developed an algorithm, which is to decrease mobility costs by measuring inter-arrival times, and adaptively adjusting the additional active time to the appropriate value. We have developed a simulator, with which we checked the effectiveness of our algorithm. Using the simulator we also investigated, how mobility cost is affected, of the distribution of packet-group length and inter-arrival times are not exponentially distributed, but heavy tailed.

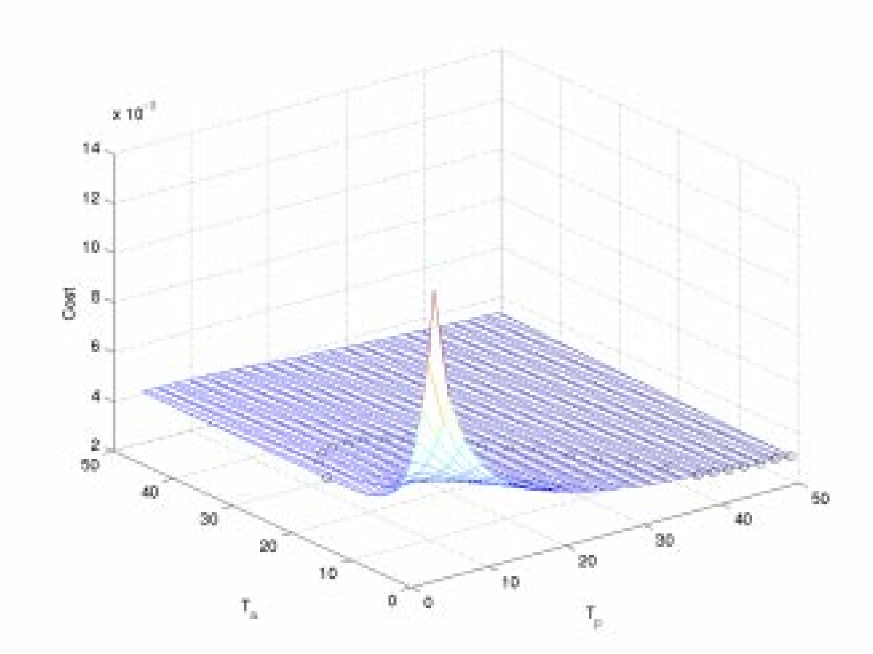

Fig. 3.1. Mobility Cost in Function of Additional Active Time and Packet-Inter-Arrival Time, Counted Results

4. Simulation Results

Our simulator is a discrete, event triggered, object oriented simulator. Three kind of object represent the parts of our IP micro mobility network model: the BS, the MH and the Net objects. In the simulator 81 piece of BS represents an area, in which the MH objects can simulate the movements of abaut 100 mobile hosts. The Net object simulates core network of the IP local mobility network. Between the objects packet-groups are transmitted, and the caused mobility cost is measured, logged and monitored.

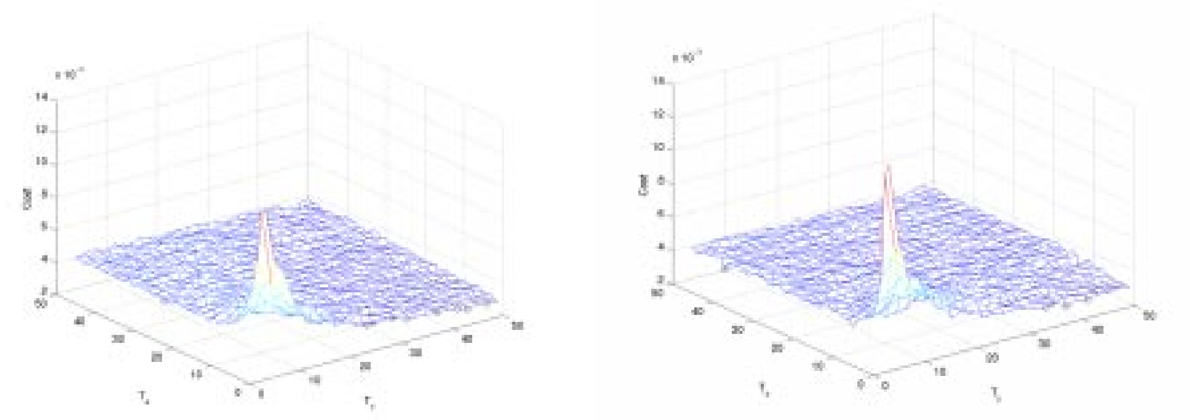

Fig. 4.1. Mobility Costs, Simulation Results

During simulation we used exponential and pareto distributions for describing packet-group length and inter-arrival times, and then we compared the results. On the left side of fig. 4.1, exponential distribution was used to describe additional active time, while on the right side pareto distribution was used. Our first observation is that the simulation results (fig 4.1) are quite similar to the analytical results (fig. 3.1). We can see the surfaces are rougher; it is caused by finite number of simulation runs. Comparing the two surfaces on fig. 4.1 it is obvious, that mobility costs are the same, irrespective to the exact distribution of packet-group length and inter-arrival times.

The packet arrival process on the Internet can be described by heavy tailed (e.g., pareto) distributions [4]. Such distributions are hard to tackle analytically and also present a challenge to simulations due to their high variance. By investigating two entirely different arrival distributions we found, however, that the mobility cost is unaffected by the exact shape of the distributions and depends only on the mean value.

5. Adaptive Algorithm

Our algorithm:

- Let the $T_a(T_p)$ function give the appropriate Ta value for a given T_p. This function can be obtained from the analytical work and is marked by circles on fig. 3.1.
- Measure each packet-group inter arrival time.
- Discard those inter-arrival time values that are lower than T_p^{min} or larger than T_p^{max}.
- Calculate an exponentially weighted moving average of the inter-arrival time values.
- Substitute the moving average into $T_a()$ and use the result as the actual length of additional active time.

6. Summary

In our investigation we examined the influence of mean time and distribution type of packet-group inter-arrival times to mobility cost. We used both an analytical formula and a simulator. We have shown that mobility cost does not depend on the type of inter-arrival time distribution, but only on its mean value. We developed a measurement-based algorithm, which can effectively decrease mobility cost by adjusting the length of additional active time periods.

References

[1] E. Gustaffson, A. Jonsson, C. Perkins, *"Mobile IP Regional Registration"*, Internet Draft, draft-ietf-mobileip-reg-tunnel-02, March 2000.

[2] C. Perkins, *"IP mobility support"*, Internet RFC 2002, IBM, October 1996. http://vega.icu.ac.kr/~iyhwang/mobile/mobileip-main.htm

[3] László PIROSKA, Miklós Aurél RÓNAI, Zoltán R. TURÁNYI, András G. VALKÓ, "Cellás IP hálózatok mobilitási költségelemzése", BUT Scientific Conference, November 2000. http://www.ronai.hu/tdk

[4] V. Paxon, S. Floyd, *"Difficulties in Simulating the Internet"*, Winter Simulation Conference, December 1997. http://www.aciri.org/floyd/papers.html

Zusammenfassung

Wir haben gezeigt, dass die Kosten der Mobilität betreffenden Nachrichtsignalen unabhängig von dem genauen Ablauf des Ankunftsprozesses der IP Paketen sind. Sie sind nur von dem Erwartungswert der Zeit zwischen den Ankünften der Paketen abhängig. So kann das Paging Prozess nach dem Verkehrsmuster ohne überflüssigen Messungen eingestellt werden. Schliesslich haben wir ein adaptives Algorithmus präsentiert, die aufgrund der Messungen der Erwartungswerten die freien Parametern einstellt, und damit die Kosten der Nachrichtsignalen reduziert