Optimization Scheme for Mobile Users Performing Vertical Handoffs between IEEE 802.11 and GPRS/EDGE networks

M.Ylianttila, M. Pande, J. Mäkelä, P. Mähönen
Centre for Wireless Communications at Oulu University
Tutkijantie 2E, P.O.Box. 4500 FIN-90014, Oulu, Finland

Abstract - **The next generation wireless networks are characterised by anywhere, anytime connectivity, enhanced data services and higher data rates to enduser. New technologies such as IEEE 802.11 WLAN, Bluetooth, HIPERLAN/2, GPRS/EDGE, cdma2000 and WCDMA aim to achieve this. To facilitate new services, and make them flexible and bandwidth efficient, vertical roaming of mobile nodes is a tempting possibility for operators. Benchmarks and metrics are needed to assess these issues. The need for qualitative and quantitative results for these parameters in a real time situation is critical. One such scenario is the effect on the network performance by means of effective throughput and handoff latency perceived by the mobile user, with increasing number of active users or network load. This paper presents simulation results for mean throughput and handoff delay obtained in vertical handoff and horizontal handoff in IEEE 802.11 and GPRS/EDGE networks. An optimization scheme for mobile users performing vertical handoffs is presented with analysis.**

I. INTRODUCTION

The next generation networks will be heterogeneous. Operators and service providers can build new service models by combining different technologies, such as IEEE 802.11 WLAN, Bluetooth, HIPERLAN/2, GPRS/EDGE, cdma2000 and WCDMA. These macro-, micro- and picocell networks often have overlapping areas of coverage. The mobile user may want to roam among these heterogeneous networks by seamlessly switching between the serving access nodes [1,2]. Internet Protocol's (IP) most recent version (v6) added with mobility support (Mobile IP) [3] has recently gained a lot of interest as a solution for global mobility. Related IETF (Internet Engineering Task Force) working groups try to enable fast and scaleable handoffs also for heterogeneous networking environment.

In Europe, the GSM based system evolution is progressing with GPRS [4] and EDGE [5] (Enhanced Data rates for GPRS Evolution). The data rates offered by these systems will depend on how many timeslots are selected by the mobile user for data communication and also the radio channel condition. Wireless local and picoarea network standardization has been carried out in IEEE 802.11 [6] and 802.15, ETSI BRAN HIPERLAN, Bluetooth and HomeRF. At the network layer level, IP assumes that a node's IP address signifies its location/connection point to the Internet. When a node moves from one subnet to another, it either has to change its IP address so that IP can route the packets destined for the mobile node correctly. Or it has to have host specific routes so that every individual node is routed separately based on their current location. Both these alternatives have their drawbacks. Host-specific routes are not scalable and have severe problems with respect to robustness and security. On the other hand each time an IP address is changed, the communication at the TCP level has to restart, thus adding delays. Mobile IP solves these problems with a reasonable amount of security, robustness and scalability and allows nodes to maintain their transport layer connections when they move from network to network.

An open issue is to optimize handoff process in proportion to several factors. Generally, in the handoff it is preferred to have negligible delay and maximum throughput with low packet loss ratio. A very important issue is to provide a smoothing function for the ping-pong effect. Since heterogeneous wireless networks can provide different data rates, ping-pong can be profitable with certain constraints. On the other hand, usually ping-pong is not desired since it generates additional signaling traffic and delay. To achieve an optimal balance between these contradicting aims, a new scheme is presented in this paper. Vertical handoff performance is analyzed with simulations in order to understand different aspects of the optimization scheme. The simulation scenario and assumptions are given in Section 2. Optimization scheme for throughput is given in Section 3. Simulation results are given in Section 4 with analysis. Conclusions and future work are outlined in Section 5.

II. SIMULATION SCENARIO

The scenario considered for handoff is as shown in Fig. 1. We consider IEEE 802.11 as the Home Network (HN) and GPRS/EDGE as the Foreign Network (FN), each having a particular subnet prefix. When evaluating the handoff performance of Mobile IPv6 [7] enabled wireless networks, the two important metrics are throughput and handoff delay. Certain features of Mobile IPv6 will have an effect on throughput and delay. The use of destination options, such as binding update, binding acknowledgement and binding request, sent as a separate packet or included in the TCP/UDP payload, increase network load and will affect the end-to-end delay. The other factor is the number of active users seeking a connection. An increasing number of users mean that individual share of bandwidth and data rate for each user will decrease, while the network load will increase.

0-7803-7208-5/01/$17.00 (C) 2001 IEEE

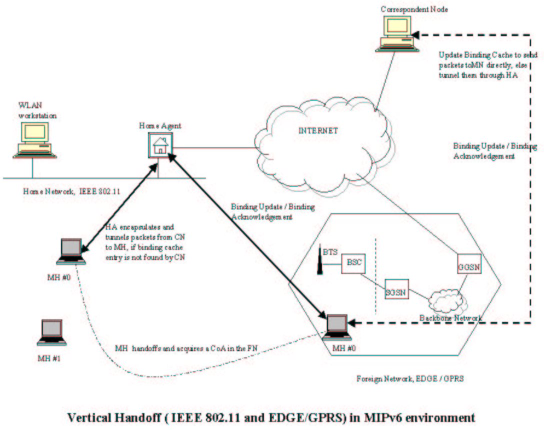

Fig. 1. Simulation Scenario for Vertical handoff in IEEE 802.11 and EDGE/GPRS using Mobile IPv6.

In our simulation we choose to model only certain features of Mobile IPv6 protocol. In the following assumptions for OPNET simulations are given.

1. We assume Mobile IPv6 enabled multiuser environment with two wireless subnetworks.
2. IEEE 802.11 is the Home Network and EDGE/GPRS is the visited or Foreign Network for one or more mobile nodes.
3. Movement of one or more mobile node is considered between two subnets only. This is done for both Horizontal handoff and Vertical handoff.
4. The data rates are constant, so the self-similar nature of the Internet traffic [8] is not considered in this paper.
5. Limited functionality of Mobile IPv6 protocol is modelled. Features pertinent to mobility issues are considered, such as header extensions (Binding Updates, Binding Acknowledgements, Home bit option etc.), packet size, addressing etc.
6. Binding updates are included in the TCP/UDP payload.
7. During handoff FIFO buffering mechanism is used.
8. A fixed packet size is used.
9. Overheads due to Authentication, Subscription of more than one mobile user are not considered.
10. Upward and downward vertical handoffs [1] are not distinguished for measuring the handoff latency.
11. Data rates for WLAN is nominally 1 Mbps, but is in practice 10-50 % of that at the edge of the cell
12. GPRS/EDGE: Always connected, wide area coverage. Protocol-payload performance 80%.

On the other hand we used MATLAB simulation framework as in [9] to analyze some aspects of the optimization scheme.
Table 1 shows the different modes of operation for GPRS and EDGE. For practical reasons, however, the first GPRS network implementation will only support CS-1 and CS-2 (using the original GMSK modulation), and first terminal implementation support maximum of 3+1 timeslots usage (3 for downlink, 1 for uplink). Therefore the maximum system performance in ideal conditions by means of throughput perceived by the user is in GPRS 40.2 kbps. In EDGE, the channel coding rates are enhanced with higher capacity modulation and coding schemes. With three timeslots and ECS-2 channel-coding scheme the data rata in EDGE would be 123 kbps. However, there are factors such as the number of active users in the same cell, protocol overhead and the number of lost packets that degrade the nominal throughput.

TABLE I

GPRS DATA RATES (PHY)

Channel Coding Scheme/ Transmit Rate (kbps)	**GPRS CS-1**	**GPRS CS-2**	**GPRS CS-3**	**GPRS CS-4**	**EDGE ECS-1**	**EDGE ECS-2**	**EDGE ECS-3**	**EDGE ECS-4**
Modulation	GMSK	GMSK	GMSK	GMSK	O16Q-AM	O16-QAM	O16-QAM	O16-QAM
1 timeslot	9.05	13.4	15.6	21.4	33.0	41.0	48.0	65.2
2 timeslots	18.1	26.8	31.2	42.8	66.0	82.0	96.0	130.2
3 timeslots (unidirectional max. of first terminals)	27.2	40.2	46.8	64.2	99.0	123.0	144.0	195.6
4 timeslots	36.2	53.6	62.4	85.6	132.0	164.0	192.0	260.8
5 timeslots	45.2	67.0	78.0	107.0	165.0	205.0	240.0	326.0
6 timeslots	54.3	80.4	93.6	128.4	198.0	246.0	288.0	391.2
7 timeslots (1 for signalling)	63.4	93.8	109.2	149.8	231.0	287.0	336.0	456.4
8 timeslots (theoretical maximum of carrier capacity)	72.4	107.2	124.8	171.2	264.0	328.0	384.0	521.6

For cellular data connection users, it would be preferable for a mobile user to seamlessly and automatically switch to a higher data rate network connection whenever it is available. On the other hand, WLAN can be the primary network e.g. for people who work in the office environment. In this scenario the user has a laptop or PDA in the office connected to the companies WLAN. When one leaves the office, one may want to maintain some network applications running, such as www download or telnet/ssh session. At the edge of WLAN cell, a handoff algorithm should decide when to trigger the handoff procedure and switch seamlessly to overlaying cellular data connection.

III. OPTIMIZATION SCHEME

Optimizing the handoff process brings up the issue of using some specific algorithm. An efficient handoff algorithm will try to minimize the delay and maximize the throughput. The most critical area is the very edge of the WLAN cell, where the received signal strength (RSS) varies around the sensitivity threshold of the WLAN receiver. This area is referred as *transition region* as in [9]. Smoothing of the ping-pong effect can be implemented in various ways. One feasible implementation of such a scheme is to employ a dwell-timer. From Fig. 2 we can see the dwell-timer functionality. It can be seen as a smoothing technique for ping-pong effect, cutting too frequent sequential handoffs. In Fig. 2 the handoff process is illustrated with a handoff signal (handoff from WLAN to GPRS/EDGE) as a Boolean function. When handoff from WLAN to GPRS/EDGE takes place, handoff signal goes up (true).

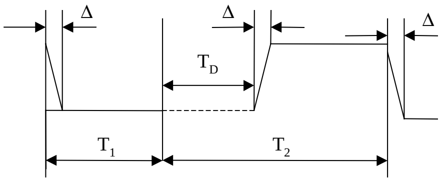

Fig. 2. Dwell-timer functionality.

Here, MN will take the samples of the RSS from the AP and compare it with a predefined threshold. If consecutive samples during predefined dwell time are below the threshold then MN initiates the handoff to GPRS/EDGE. Otherwise it will persist with WLAN. As in [9], we use the following variables to deduce an optimization scheme for throughput in vertical handoff:

χ: WLAN receiver sensitivity threshold
T_t: Region where RSS falls below the threshold first time and falls below it for good (transition region)
T_1: Each contiguous stretch of time where $P > \chi$ within T_t
T_2: Each contiguous stretch of time where $P < \chi$ within T_t
N: Number of times RSS crosses the value of χ
Δ : Handoff delay
T_D: Dwell time
R_1: Data rate available over the air in WLAN
η_1: Throughput reduction coefficient for WLAN
R_2: Data rate available over the air in GPRS/EDGE
η_2: Throughput reduction coefficient for GPRS/EDGE

In addition to handoff delay, several other factors may degrade the throughput. Such factors are packet losses (which cause retransmission), packet encapsulation delay (such as in IP-in-IP encapsulation in Mobile IP), protocol-payload ratio and the number of active users in the cell. For these reasons the throughput perceived by the mobile user can degrade even down to 1-10 % of the nominal data rates. In the simulations we evaluated the combined effect of various factors degrading the throughput by using throughput reduction coefficient η. It is a system specific parameter. Thus the effective throughput via WLAN and GPRS/EDGE is:

$$s_1 = \eta_1 R_1 \qquad (1)$$

$$s_2 = \eta_2 R_2 \qquad (2)$$

Additionally we define parameter Ω as the effective throughput ratio:

$$\Omega = s_1 / s_2 = \eta_1 R_1 / \eta_2 R_2 \qquad (3)$$

Handoff from WLAN to GPRS/EDGE is profitable for optimizing the throughput when:

$$s_2(T_1 - T_D - \Delta) > s_1 T_1 \qquad (4)$$

$$T_1 > \frac{s_2}{s_2 - s_1}(T_D + \Delta) \qquad (5)$$

$$T_1 > \frac{T_D + \Delta}{1 - \Omega} \qquad (6)$$

In practice this means that handoff is profitable only when $\Omega < 1$, and $s_2 > s_1$. This kind of a situation can occur for example when WLAN performance is degraded down to e.g. 10-30% and MN has a multislot GPRS/EDGE connection available.

Handoff from GPRS back to WLAN (i.e. ping-pong) is profitable when:

$$s_1(T_2 - T_D - \Delta) > s_2 T_2 \qquad (7)$$

$$T_2 > \frac{s_1}{s_1 - s_2}(T_D + \Delta) \qquad (8)$$

$$T_2 > \frac{T_D + \Delta}{1 - \Omega^{-1}} \qquad (9)$$

In practice this means that handoff is profitable only when $\Omega > 1$, and $s_1 > s_2$. This is the case for example when WLAN is not heavily loaded and/or MN is not using a multislot GPRS/EDGE connection that would provide a higher data rate (and would have higher cost). Thus we can infer the rules for optimizing the throughput:

1) Handoff from WLAN to GPRS/EDGE
   If (6) is true, then persist in WLAN, else make handoff to GPRS/EDGE.

2) Handoff back from GPRS/EDGE to WLAN
   If (9) is true, then make handoff to WLAN, else stay in GPRS/EDGE.

As the duration of T_1 and T_2 are unknown at the time when decision for handoff has to be made, a robust algorithm could try to predict them each time RSS crosses. A simpler strategy is to use a predefined dwell-timer for different operating modes (data rates).

IV. SIMULATION RESULTS

The objective of this simulation was to analyse the two critical parameters, mean throughput and handoff delay for different values of η_1 and the number of active users in the transition region, based on this optimization scheme. We consider three simulation scenarios and compare the results obtained from these. We vary the number of users as 1, 5, 10, 15, 20, 50, 100 and 200, and see the effect on mean throughput and handoff delay. In the graphs, we mainly show the relative performance for GPRS class CS-1 and EDGE 3 slot ECS 2.

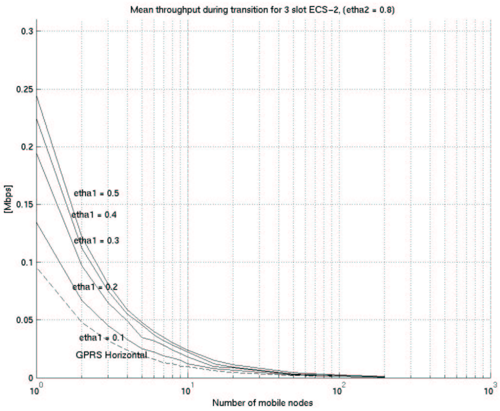

Fig. 3. Comparison of mean throughput in vertical handoff and in horizontal handoff for 3 slot ECS-2.

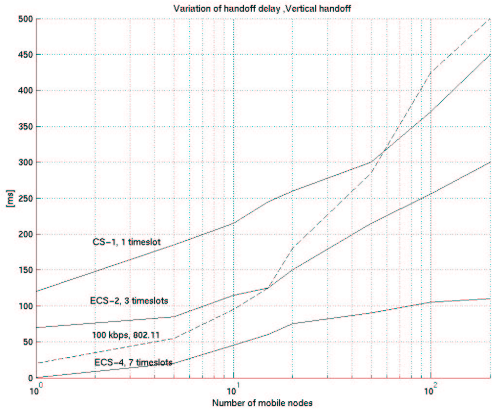

Fig. 4. Comparison of variation of handoff delay in vertical handoff with 3 slot ECS-2, and horizontal handoff in 802.11.

In Fig. 3 we see that the mean throughput achieved for Horizontal Handoff in 3 slot ECS-2, is almost coincident with that achieved in Vertical Handoff with value of η_1 for 802.11 being 0.1. As the number of mobile nodes increases up to 10 or more, the mean throughput collapses to an unacceptable level. Thus a network operator may want to limit the number of mobile nodes requesting vertical handoff based on the network environment and application specific requirements and expected latencies and mean throughput.

Fig. 4 depicts the variation in handoff delay with increasing number of active mobile users. As expected handoff delay increases with increasing number of users. Handoff delay in horizontal handoff in 802.11 increases faster than in vertical handoff due poor performance of CSMA/CA protocol with increasing number of active users, as has been shown in numerous studies.

Fig. 5 shows MATLAB simulation results how the length of dwell time effects the throughput for different GPRS/EDGE modes (R_2). Most drastic gains can be achieved for the throughput when the difference in the data rates of the two systems is high (Ω is high). Fig. 5 shows that the optimum value of dwell time is specific for each GPRS mode. For example, throughput in the transition region for CS-2 with 1 timeslot is optimized with a 1sec dwell timer. The gain is significant (17,8 times the throughput of CS-2 with 1 timeslot) compared to the case where handoff is executed only when RSS falls below χ for the first time. Similarly, throughput in the transition region for CS-2 with 4 timeslots is optimized with a 200 ms dwell timer. The gain is still 4,7 times the throughput of CS-2 with 4 timeslots. Consequently, a robust handoff algorithm using a dwell timer could have different values of dwell time defined for each GPRS/EDGE mode thus optimizing the throughput.

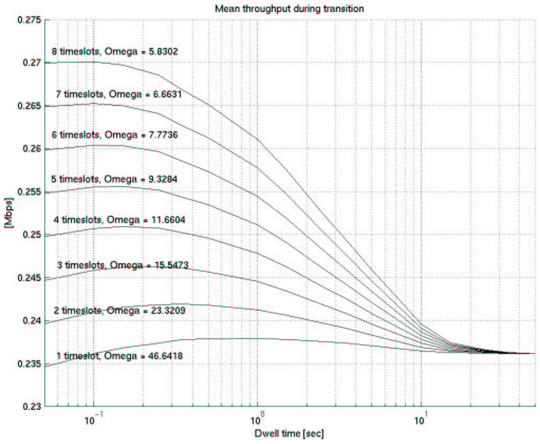

Fig. 5. Mean throughput with GPRS for CS-2 during transition as a function of dwell time.

In general, higher the difference in data rates of the two systems (when Ω is high) the longer dwell time is preferable. This is due the fact that with very low GPRS/EDGE data rates, MH does not benefit in practice at all from a temporary handoff to a lower data rate system. In these cases it is preferable to persist in the WLAN as long as RSS falls below χ for good and make only one handoff at the end of the transition region. From Fig. 5 we can approximately see that when $\Omega > 5$ then some amount of dwell time is profitable. When $\Omega < 5$, dwell timer does not bring any added value as a part of the handoff algorithm, but is degrading the throughput.

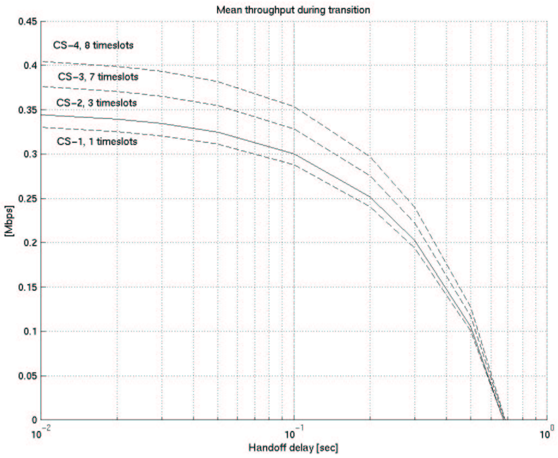

Fig. 6. Mean throughput for four GPRS modes during transition as a function of handoff delay.

A very important issue to consider in the context of vertical handoff is the handoff delay. It is not generally well known how much handoff delay is affordable in vertical handoff. Also it has not been earlier shown how handoff delay effects to the throughput during the transition. Fig. 6 shows how handoff delay effects to the performance of vertical handoff. We can see that when handoff delay exceeds 650 ms continuous handoffs (ping-pong effect) suffocate the throughput. Only one handoff may not be the optimal since the mobile user may gain from making handoffs back and forth during the transition region where the signal level of WLAN goes up and down around the receiver sensitivity threshold. However, the throughput should not drop below the throughput of GPRS/EDGE. Therefore the upper bound for handoff delay for e.g. CS-3 with 7 timeslots would be less than 500 ms.

V. CONCLUSIONS AND FUTURE WORK

In this paper an optimization scheme for mobile users performing vertical handoff was presented with analysis. Simulation results for mean throughput and handoff delay obtained in vertical handoff and horizontal handoff in IEEE 802.11 and GPRS/EDGE networks was presented. Results show how the number of active users in the transition region, desiring a handoff, affect the performance. A dwell timer can be used to optimize throughput. Maximum time bounds for handoff delay and dwell time, and analyze the effect of degraded throughput typical in multi-user and fading channel environments were shown. It was seen that some amount of dwell time is usually preferable (i.e. when $\Omega > 5$). However, the amount of optimal dwell time varies along with the used data rate (GPRS/EDGE mode, or to be more precise, with effective throughput ratio. Consequently, a robust handoff algorithm using a dwell-timer could have different values of dwell time defined for Ω thus optimizing the throughput. Based on the shown results and formulas, vertical handoff profitableness as a function of handoff delay and effective throughput ratio can be evaluated. Profitableness estimate for handoff could be used in a robust handoff algorithm as one of the input parameters. It should be noted that a successful optimization from one user's perspective must not decrease the system performance in large scale. Local optimization scheme presented in this paper is scaleable since it exploits otherwise unused bandwidth at the edge of WLAN cell. In this simulation Internet traffic is assumed to be constant bit rate and fixed packet size. Future work includes using more realistic traffic models such as self-similar or fractal traffic models. More work is also needed to further optimize the handoff performance.

ACKNOWLEDGMENTS

The acknowledgments are due to the project team of WINGIP/WINNER[1]. The authors would like to thank TEKES, Nokia and Finnish Defence Forces for supporting this project financially. Special acknowledgment is given to Prof. Kaveh Pahlavan from Worcester Polytechnic, USA for fruitful discussions and variety of help.

REFERENCES

[1] M. Stemm and R.Katz, "Vertical Handoffs in Wireless Overlay Networks", ACM Mobile Networking (MONET), Special Issue on Mobile Networking in the Internet, Summer 1998, pp. 335-350.

[2] K. Pahlavan, P. Krishnamurthy, A. Hatami, M. Ylianttila, J. Mäkelä, R. Pichna, J. Vallström, "Handoff in Hybrid Mobile Data Networks", IEEE Personal Communications Magazine, Vol.7 No.2, April 2000, pp. 34-47.

[3] C. Perkins, "Mobile IP: Design Principles and Practice", Addison-Wesley Longman, Reading, Mass., 1998.

[4] G. Brasche and B. Walke, "Concepts, Services, and protocols of the new GSM Phase 2+ General Packet Radio Service", IEEE Communications Magazine, August 1997, pp. 94-104.

[5] A. Furuskar, M. Frodigh, H. Olofsson, J. Skold, "System performance of EDGE, a proposal for enhanced data rates in existing digital cellular systems", Proceedings of 48th IEEE Vehicular Technology Conference, 1998 (VTC 98), volume 2, pp. 1284 -1289.

[6] ISO/IEC 8801-11, ANSI/IEEE Std 802.11, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, 530 pages.

[7] D. Johnson and C. Perkins, "Mobility Support in IPv6" (work in progress). Internet Draft, draft-ietf-mobileip-ipv6-14.txt, Internet Engineering Task Force, July 2000, 120 pages.

[8] B.Ryu, S.Lowen,"Fractal traffic models for Internet simulation", Proceedings of Fifth IEEE Symposium on Computers and Communications, 2000, pp. 200 -206.

[9] A. Hatami, P. Krishnamurthy, K. Pahlavan, M. Ylianttila, J. Mäkelä, R. Pichna, "Analytical Framework for Handoff in Non-Homogeneous Mobile Data Networks", Proceedings of PIMRC'99, Osaka, Japan, September 1999, pp. 760-764.

---

[1] http://www.cwc.oulu.fi/winner/