Evaluating Rate-Estimation for a Mobility and QoS-Aware Network Architecture

Nuno Vasco Lopes* Maria João Nicolau** Alexandre Santos*
*Department of Informatics, **Department of Information Systems, University of Minho, Braga, Portugal
E-mail: {vascolopes, alex}@di.uminho.pt, joao@dsi.uminho.pt

***Abstract*—In a nearby future wireless networks will run applications with special QoS requirements. FHMIP is an effective scheme to reduce Mobile IPv6 handover disruption but it does not deal with any other specific QoS requirement. Therefore new traffic management schemes are needed in order to provide QoS guarantees to real-time applications and this implies network mobility optimizations and congestion control support. Traffic management schemes should deal with QoS requirements during handover and should use some resource management strategy in order to achieve this. In this article a new resource management scheme for DiffServ QoS model is proposed, to be used by access routers as an extension to FHMIP micromobility protocol. In order to prevent QoS deterioration, access routers pre-evaluate the impact of accepting all traffic from a mobile node, previous to the handover. This pre-evaluation and post decision on whether or not to accept any, or all, of this new traffic is based on a measurement based admission control procedure. This mobility and QoS-aware network architecture, integrating a simple signaling protocol, a traffic descriptor, and exhibiting adaptive behavior has been implemented and tested using ns-2. All measurements and decisions are based on DiffServ class-of-service aggregations, thus avoiding large flow state information maintenance. Rate estimators are essential mechanisms to the efficiency of this QoS-aware overall architecture. Therefore, in order to be able to choose the rate estimator that better fits this global architecture, two rate estimators - Time Sliding Window (TSW) and Exponential Moving Average (EMA) - have been studied and evaluated by means of ns-2 simulations in QoS-aware wireless mobility scenarios.**

1. Introduction

Being able to provide QoS levels suited to real-time applications needs is, in itself, a big challenge for research community. Currently, Mobile IP standard lacks on QoS provisions, scalability, robustness and on a unified resource management function. Mobile IP is a macro-mobility solution, and generally poorly suited to micro-mobility scenarios where cell size is small and high frequency handovers are common. There are a few proposals for micro-mobility problem, such as Hierarchical Mobile IP [1], Fast Handover [2], Cellular IP [3]. However, micro-mobility proposals and Mobile IP are best-effort and do not provide QoS guarantees i.e., currently mobility management and QoS models work apart. The unpredictable behavior nature of wireless links associated with the mobile node point of attachment leads to high dynamics on link utilization. Thus, an important issue during mobile handover is to provide information about network status to the resource management. This way, when the mobile node moves to a new access router, active applications on mobile could be negotiated in the new access router, as part of the handover procedure. Applications should also be provisioned with suited QoS, ensuring that packets reach the mobile in accordance with QoS contract. This work integrates Fast Hierarchical Handovers [2] to enhance MIPv6, a signaling protocol to request services, the DiffServ model, for traffic differentiation, and a new admission control scheme, to prevent QoS class deterioration. This combination of components will be optimized to work together in order to support seamless handovers for mobile users running real-time applications.

Nuno V. Lopes was supported by an FCT Grant (SFRH/BD/35245/2007)

The operation of this architecture is based on FHMIP mobility management messages that carry QoS context values to the admission control algorithm function in the new access router. The admission control decision is taken before L2 handover using both the QoS context information and the estimated needs for class bandwidth on the new access router. Rate estimators, running on access routers, perform measurements in order to extract current QoS context. Operationally, this QoS framework also preserves the QoS on mobile nodes already being served by the access router because admission control decision is taken prior to the mobile node handover.

In addition this framework does not exhibit scalability problems because the routers are nearly stateless and provides a seamless mobility control capability by adjusting the mobile class flows according to resource availability on the new access router.

This paper is organized into six sections. This section presented the motivation for the development of a new architecture. A presentation of micromobility protocol and associated signaling process follows, in section two. In section three we present and analyze related works. Then, in section four, an overview of the whole proposed architecture is presented, with associated resource management functions and rate estimators. Section five describes the implementation, the ns-2 simulations carried out and the discussion of experimental results both with TSW and EMA rate-estimators, in several mobility scenarios. Finally, paper concludes by summarizing both the advantages of this new architecture and rate-estimators simulation results.

2. Micromobility Protocol

To achieve QoS enhancements in MIPv6 an optimized mobility management scheme is mandatory with Fast (with

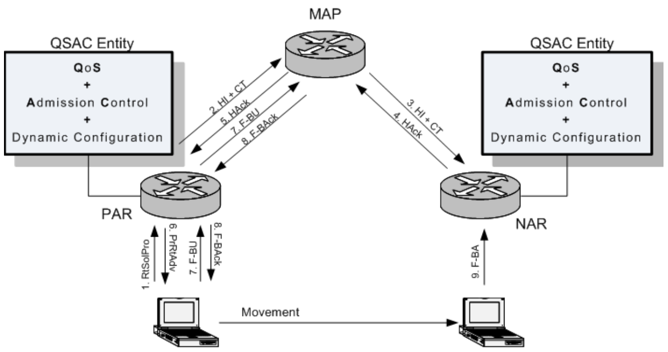

Figure 1: Architecture Signaling Process

strict delay bounds) and Smooth (with minimum losses of packets) Handovers. The combination of Fast Handover and HMIPv6 enables the anticipation of layer 3 handover, so that data traffic can be efficiently redirected to the mobile node's new location before it moves there. The hierarchical mobility management model allows the performance enhancement of Mobile IPv6 with local bindings, while using Fast Handovers helps Mobile Nodes to achieve seamless mobility. Therefore, the use of both HMIPv6 and FAST Handover is crucial to pursuit this goal.

The adopted strategy to integrate HMIPv6 and Fast Handover mechanisms was to place the MAP as the aggregation router, i.e., the first node of convergence or divergence depending on the direction of data-path. Using MAP on the aggregation router may improve the efficiency because, being the MAP the first point of divergence, it is the best place to redirect traffic to new path, thus saving delay and bandwidth between the aggregation router and the pAR. However, this is not enough to have quality of service within a domain. Regardless of fast and hierarchical handover could improve the IP connectivity, there is also a need to establish the QoS context that MN had on the previous router whenever a handover occurs. Therefore, transferring QoS contexts would facilitate other protocols to operate without the need for context reestablishment. Figure 1 shows all the signaling process steps before MN handover takes place; context data strutures are encapsulated into the FHMIP mobility signaling messages and the QoS control decision is made before the handover occurs. Notice that the QoS context is derived from measurements extracted by rate estimators. Hence, in order to have a precise admission control it is necessary to have a precise rate estimator, since the former uses information given by the latter.

Architecture details such as the key components of context transfer, data structure representation of context for interoperability across access routers and the encapsulation of context data structures are explained in [4] and overviewed in section 4.

3. Related Work

Literature shows that there have been some attempts to implement and enhance QoS frameworks proposed for fixed networks in wireless networks. In [5] the authors present the QoS-Conditionality Handover for Mobile IPv6 to eliminate the need of signaling protocols. It uses the QoS Option in the hop-by-hop extension header of BU message to carry QoS context. This solution has the disadvantage that all nodes needed to be changed in order to implement this functionality. Studies [6], [7] integrate RSVP and HMIP micro mobility protocols to provide QoS guarantees on UMTS environment. In [8] RMD has been used in UMTS access network and performance of measurement-based admission control algorithms on interior nodes have been evaluated.

Also, rate estimators are essential mechanisms to the efficiency of a QoS-aware architecture. Several approaches on rate estimators for packet networks [9], [10] have been proposed and analyzed. However, determining the kind of rate estimator that is well suited for a QoS-aware architecture is not trivial, as the quality of the estimation depends on overhead costs, stability, accuracy and even on its time-responsiveness [11], [12].

Concluding, despite of these improvements, signaling overhead and processing load problems are not completely solved. Therefore, our approach effort has been to solve this problem with more relaxed QoS requirements i.e., soft real-time services, in order to avoid the signaling overhead and Bandwidth Brokers.

Moreover, as admission control is based on measurements, signaling and processing load are minimized. Finally, the establishment of QoS context on new access routers before handover takes place enables to reduce the number of renegotiations with the user.

4. Proposed QoS Model

A. *Overview*

A QoS model should define: the behavior of the resource management function; the inputs; the way QoS information is used to require resources; and the control information needed for resource management function. It must also describes a minimum set of parameters that should be used in the request message when signaling occurs. In order to implement a QSAC (QoS Admission Control) entity for the proposed architecture, the major design issues were: to use DiffServ mechanism as the quality of service model; to select the access router as a critical point in the end-to-end path; to define QSAC entities as stateless, only handling QSAC messages that contain QoS parameters. Requested QoS parameters are to be handled by a resource management function, which coordinates the activities required to grant and configure resources, e.g. admission control determines whether node has sufficient resources to support the requested QoS. If QoS availability checks succeed, the mobile traffic is accepted.

Interiors nodes need not be QSAC-aware because it is assumed that the whole domain is DiffServ so constrains are essentially in the wireless link. An important aspect of this model is that it uses a scalable QoS signaling and supports seamless mobility during handovers, provided by the resource management function in access network nodes. Resource management functions are based on a method that is able to provide MBAC (Measurement-Based Admission Control) for flows entering a DiffServ domain and describe a system that can dynamically adjust the load of classes in access networks, in order to maintain the QoS levels on the new access router in a proactive manner. The measurement-based algorithm estimates traffic levels (predicted resource utilization) and admits flows whose needs are within its availability at the time of the request. Once an admission decision is made, no record of the decision needs to be stored, it does not require pre-reservation state nor explicit release of reservations. The admission decision will be negative if the currently carried traffic, as characterized by the estimator, plus the requested resources, in the traffic descriptor, for the new flow, exceeds the DiffServ class capacity. The resource management function stores the values of estimated classes bandwidth on each access router and measures the class bandwidth in use by mobile nodes when a handover takes place.

B. *Resource Management Function Behavior*

Admission control will determine whether available resources can support the requirements of Mobile Node when it moves to the new access router. Case there are available resources, it admits the incoming mobile node flows and, if necessary, adapts the scheduler parameters to maintain the QoS classes requirements. Otherwise, it only rejects those flows that belong to the refused class and admits others. Therefore, before any handover effectively takes place, the new access router must evaluate the impact of admitting the new mobile node, in order to prevent possible QoS deterioration of already associated mobile nodes. The goal is to make admission control decisions based on the network-status reported and on MN per class QoS requirements at the new access router.

Presented work is focused in a measurement-based admission control that enables congestion avoidance in wireless IP networks. The admission control algorithms, located on access routers, are stateless, offer a simple traffic descriptor, high levels of adaptability and a good estimation of the aggregated traffic behavior. The measurement-based admission control algorithm has two distinct logical entities: estimator and policer (see Figure 2). Estimators implement measurement mechanisms in order to determine the current network load. The policer runs an algorithm to decide whether to admit, or reject, the new flows in the case of handover. For new flows the decision is made at DiffServ edge, based on inputs

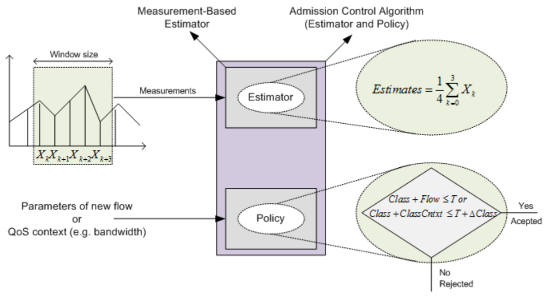

Figure 2: Measurement-based Admission Control

from traffic descriptors; for a handover the decision is based on QoS context and resource estimator measurements at the time of handover. Thus, the decision relies on inputs from requesting flow or handover, which typically includes its QoS requirements within each class.

Rate estimators: Time Sliding Window (TSW) estimator provides a running bandwidth average of traffic classes over a window of time. It uses all packets in order to determine the rate. It takes into account burstiness and smooths out its estimate to the long-term measures within each class. The Exponential Moving Average (EMA) applies weighting factors which decrease exponentially, providing a greater importance to recent observations. An interesting discussion on appropriate parameter values to be used for estimators can be found in [9] and [10].

5. Simulation Results

A. *Simulation Setup*

Network simulator, with FHMIP and NIST patch, has been used in order to carry out simulations. Simulation is restricted to a single DiffServ domain where the FHMIP micromobility protocol is implemented. A simple topology has been set up: two access routers (pAR and nAR) and two mobile nodes. Initially one mobile node is located in the pAR and the other mobile node is to be served via the nAR and both mobile nodes are receiving traffic. Mobile nodes are all receiving CBR flows, marked within different DiffServ Class, being originated in fixed correspondent nodes, somewhere within a DiffServ domain. All flows start at different instants of time, within the period 0-80 seconds.

Mobile node 1 traffic includes one flow of 13Kbps in class 1, two flows of 15Kbps each in class 2, five flows of 30Kbps each in class 3 and three flows of 60kbps each in class 4, with a traffic grand total (all classes aggregated) of 373Kbps.

Mobile node 2 traffic includes one flow of 18Kbps in the class 1, one flow of 30Kbps in class 2, two flows of 20Kbps in class 3 and two flows of 40kbps in class 4, with a traffic grand total (all classes aggregated) of 168Kbps.

As soon as 80 seconds have passed, mobile node 1 starts to move, each time at a different speed, towards a region within the nAR scope and so all its traffic may eventually (if

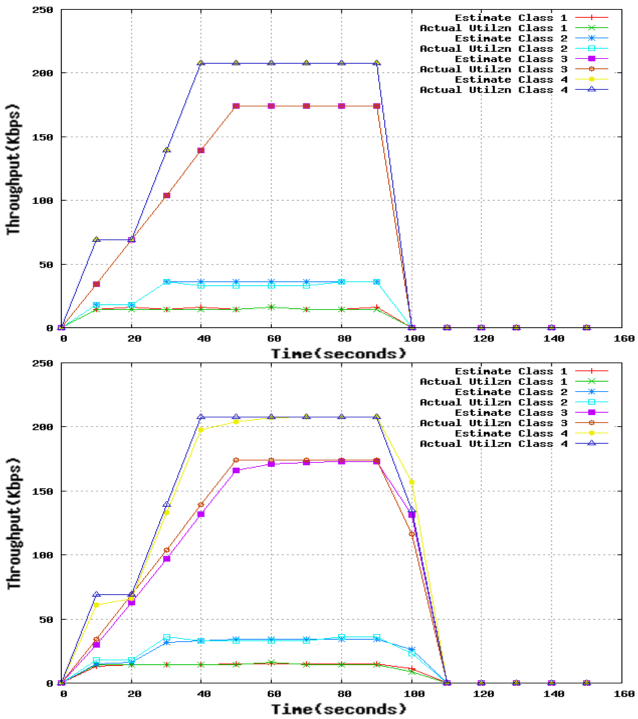

Figure 3: Estimated vs Real Throughput on Mobile Node (moving at $1m/s$) - (a) TSW (b) EMA

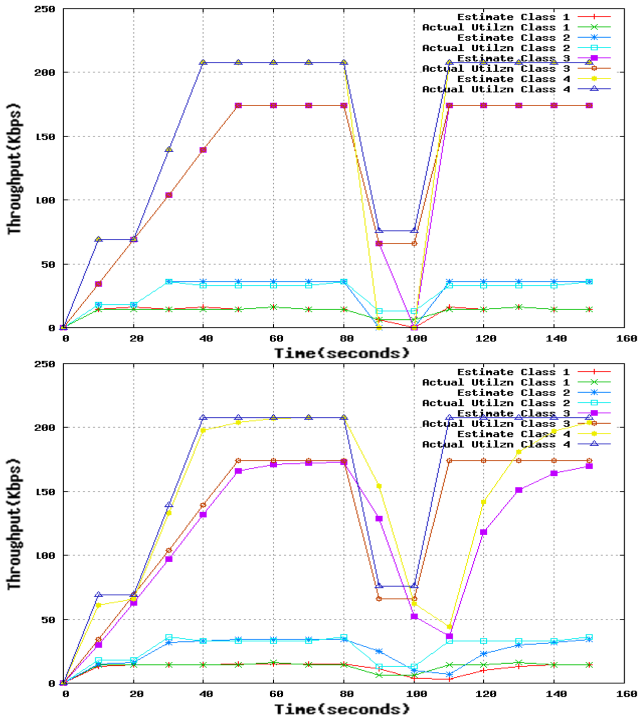

Figure 4: Estimated vs Real Throughput on Mobile Node (moving at $3m/s$): (a) TSW (b) EMA

accepted) move from the pAR to the nAR router. Simulation tests have been carried out using this topology and with two different rate estimators: Time Sliding Window (TSW) and Exponential Moving Average (EMA). For each experiment with each rate estimator we vary the mobile node velocity during handover from 1m/s, to 3m/s, to 10 m/s.

The two rate estimators are evaluated and compared in order to analyze its behavior and determine which will be the best one to mimic the traffic dynamics in order to take part of the overall architecture. There has been a special concern in the estimator performance during handover periods, therefore estimations are analyzed and commented specially for mobile node 1 and the new access router (nAR). Results are analyzed both in terms of traffic dynamics and rate deviation between estimated and real used bandwidth.

B. Class Estimator for the Mobile Node 1

Chosen speeds for mobile node handover are equivalent to mobility speeds of nodes when walking (3.6 Km/h), running (10.8 Km/h) and cycling (36.0 Km/h).

1) Mobile Node velocity of 1m/s: Figure 3 shows the estimated and the real bandwidth usage for mobile node 1 when it moves towards new access router with velocity of 1 m/s. It can be seen that before handover the TSW estimator sightly under estimates classes with less traffic, whereas EMA estimator under estimates classes with more traffic.

2) Mobile Node velocity of 3m/s: Figure 4 shows the estimated and the real bandwidth usage for mobile node 1 when it moves towards new access router with velocity of 3 m/s. When moving at velocity of 3 m/s mobile node receives traffic from both access routers. In the case of TSW estimator, the estimation is almost equal to bandwidth utilization. There is only a little decrease of estimation values during handover but it rapidly recovers from this shifting.

In the case of EMA estimator, its estimation during handover is significantly delayed in relation to the current class load. In contrast with the former, that immediately follows the actual class bandwidth utilization, the latter only converges its estimation values to the real bandwidth usage by the end of simulation.

C. Class Estimator at the New Access Router

The estimation of bandwidth usage at the new access router, when mobile node velocity is 1 m/s, only takes into account mobile node 2 traffic since, at this velocity, simulation ends before mobile node 1 starts to receive traffic from new access router. For the experiment with mobile velocity of 3 m/s, it receives traffic from both access routers. As shown in Figure 5, at this velocity the TSW estimator exhibited a good accuracy in the estimation of the actual class bandwidth utilization, whereas the EMA estimator expressed a significant delay to follow the current class traffic load (see Figure 5b).

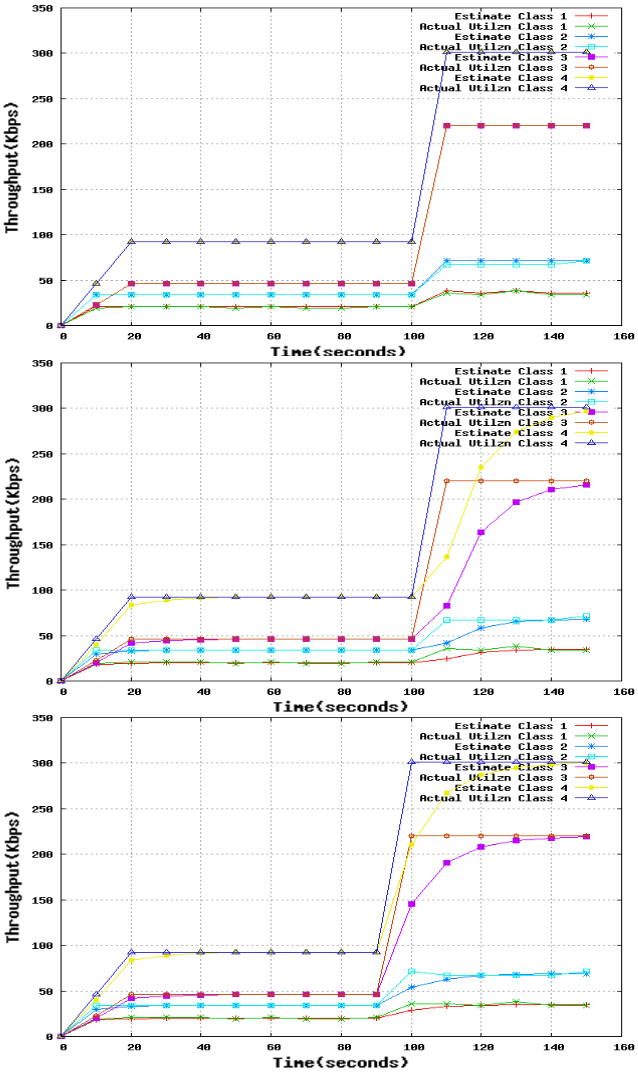

Figure 5: Estimated vs Real Throughput on nAR: (a) TSW at $3m/s$ (b) EMA at $3m/s$ (c) EMA at $10m/s$

Even in the case of a mobile velocity of 10 m/s it shows a significant delay in the achievement of reasonable accuracy (see Figure 5c).

6. Conclusion

This work proposes an add-on to FHMIP micromobility protocol enabling the support of Quality of Service. For this purpose a new resource management function for DiffServ model has been designed. The implemented resource management function is being tested and presents a scalable solution based on a class measurement-based admission control algorithm. The architecture has been conceptualized on the network layer-3 in order to provide a common framework across the different network access technologies. Our scheme aims to reduce the signaling overhead because it uses an in-band message, with mobility and QoS information, tries to avoid the congestion overload on the new access router by implementing a measure-based admission control to support the handover decision. The resource management function in the access router is able to evaluate the impact of admitting the incoming mobile node before it moves to the router, thus preventing QoS deterioration of the existing traffic.

Relating to rate estimators, several ns-2 simulations have been used and results show that TSW estimator provides a reasonably accurate estimation of traffic within each class for all mobile node velocities, whereas the EMA expressed a significant delay to converge its estimation to the actual class traffic load. Another relevant aspect illustrated by experiments was the slightly under-estimation of class traffic usage by the two estimators. However, this under-estimation is stronger when the EMA estimator is used.

Ongoing work deals with the handover process that will be policed by an admission control that reacts accordingly to handover requirements, avoiding network congestion at access routers. This probably will lead to QoS and QoE (Quality of Experience) improving from the customer's point of view and also from the network operator's point of view. The complete solution is currently being deployed in the ns-2 platform.

References

[1] H. Soliman, C. Castelluccia, K. E. Malki, and L. Bellier, "Hierarchical mobile ipv6 mobility management (hmipv6)," August 2005.

[2] R. Koodli, "Fast handovers for mobile ipv6," *IETF RFC 4068*, July 2005.

[3] Valkó and A. G., "Cellular ip: a new approach to internet host mobility," *SIGCOMM Comput. Commun. Rev.*, vol. 29, no. 1, pp. 50–65, 1999.

[4] N. V. Lopes, M. J. Nicolau, and A. Santos, "Qos-aware architecture for fhmip micromobility," in *ICUMT'2009 (IEEE Xplore)*, 12-14 October 2009.

[5] C. Liu, D. Qian, Y. Liu, and K. Xiao, "A framework for end-to-end differentiated services qos context transfer in mobile ipv6," in *Proc. First IEEE and IFIP International Conference in Central Asia on Internet* (D. Qian, ed.), pp. 466–475, 2005.

[6] L. Sanchez, L. Sanchez, and P. Ruiz, "Qos and micromobility coupling integrated with umts," in *Proc. 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications PIMRC 2004* (P. Ruiz, ed.), vol. 2, pp. 763–768 Vol.2, 2004.

[7] L.-A. Galindo-Sanchez and P.-M. Ruiz-Martinez, "Qos and micro-mobility coupling: Improving performance in integrated scenarios," *UPGRADE*, vol. VI, April 2005.

[8] H. E. Allali, G. J. Heijenk, A. Lo, and I. G. M. M. Niemegeers, "A measurement-based admission control algorithm for resource management in diffserv ip networks," *Proceedings 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06)*, pp. 1–5, September 2006.

[9] D. D. Clark and W. Fang, "Explicit allocation of best-effort packet delivery service," *IEEE/ACM Trans. Netw.*, vol. 6, no. 4, pp. 362–373, 1998.

[10] N. S. W. Fang and B. Nandy, "A time sliding window three colour marker (tswtcm)." IETF RFC 2859, June 2000.

[11] F. Agharebparast and V. Leung, "A new traffic rate estimation and monitoring algorithm for the qos-enabled internet," in *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, vol. 7, pp. 3883–3887 vol.7, Dec. 2003.

[12] K. Salah and F. Haidari, "On the performance of a simple packet rate estimator.," in *AICCSA*, pp. 392–395, IEEE, 2008.