

Adaptive Admission Control for Mobile Ad Hoc Networks based on a Cross-layer Design^{*}

María Canales, José Ramón Gállego, Ángela Hernández-Solana, and Antonio Valdovinos

Institute of Engineering in Aragón, I3A, University of Zaragoza.

C\ María de Luna, 3, 50.018, Zaragoza (Spain).

mcanales@unizar.es, jrgalleg@unizar.es, anhersol@unizar.es, toni@unizar.es

Abstract. Radio resource management and QoS provision in Mobile Ad hoc NETWORKS (MANETs) require the cooperation among different nodes and the design of distributed control mechanisms, imposed by the self-configuring and dynamic nature of these networks. In this context, in order to solve the tradeoff between QoS provision and an efficient resource utilization, a distributed admission control is required. This paper presents an adaptive admission procedure based on a cross-layer QoS Routing supported by an efficient end-to-end available bandwidth estimation. The proposed scheme has been designed to perform a flexible parameters configuration that allows to adapt the system response to the observed grade of mobility in the environment. The performance evaluation has shown the capability of the proposal to guarantee a soft-QoS provision thanks to a flexible resource management adapted to different scenarios.

1 Introduction

Nowadays applications heavily demand the fulfillment of their Quality of Service (QoS) requirements. In Mobile Ad hoc NETWORKS (MANETs), characterized by a self-organizing capacity that leads to a distributed operation and a great dynamism, the complexity to solve the tradeoff between an efficient resource utilization and the required resource reservation according to the application demands points to an efficient design of distributed admission mechanisms.

The role of routing in a cooperative environment, where the establishment of multihop routes is required, is especially relevant in order to define the admission criteria [1]. A routing decision based on a metric that reflects the resources availability allows to identify the system ability to provide the demanded quality of service. However, this measurement highly depends on the resource management performed by the Medium Access Control (MAC) level, which suggests the collaboration between both layers (cross-layer [2], [3]) as a promising solution.

^{*} This work was financed by the Spanish Government (Project TEC2004-04529/TCM from MEC and FEDER), Gobierno de Aragón for WALQA Technology Park and the European IST Project PULSERS Phase II (IST - 027142)

In addition, the inherent dynamism of mobile ad hoc networks involves new difficulties that require more flexible admission mechanisms in order to provide a resource management adapted to the topology-variant scenarios.

In this context, this paper describes an adaptive cross-layer proposal that allows to perform a distributed admission procedure with a flexible reallocation that responds to the different grades of mobility of the environment, triggering accordingly the appropriate configuration mode. Therefore, the proposed architecture guarantees an end-to-end quality of service provision with an appropriate adaptation to the variability of MANETs.

The remaining of the paper is organized as follows. Section 2 presents the cross-layer architecture, detailing the basis of the MAC protocol and the designed admission procedure based on a proposed QoS Routing. The mechanisms to adapt the operation to a mobility scenario are shown in Section 3. The proposal scheme has been evaluated through simulations and results are shown in Section 4. Finally, some conclusions are provided in Section 5.

2 The cross-layer proposal

2.1 Efficient MAC protocol

In this proposal, a MAC TDMA layer based on the ADHOC MAC protocol has been considered [4]. ADHOC MAC works on a slot synchronous physical layer and implements a distributed access technique capable of dynamically establishing a reliable single-hop Basic broadcast CHannel (BCH) for each active terminal. Each BCH carries signaling information, including priorities, which distributes layer-two connectivity information and resources occupation to all the terminals. In response to the QoS demanded by multimedia applications, the MAC level efficiently allocates resources by exploiting the in-band signaling provided by the protocol. The reservation strategy is based on the use of the BCH capabilities to signal the request before the access, in such a way that collisions can be theoretically avoided (Book In Advance Scheme - BIAS). According to the demanded resources, a user signals its attempts to reserve in addition to the occupied resources. Upon reception of different requests for the same resource, each node signals the decision about the possible reservation (even if it is not the destination) according to a common criterion. The soliciting node receives this information in the corresponding BCHs in the next frame. If there are not collisions, the resource is effectively reserved – Fig. 1. Preemption can be also carried out in order to allocate resources for high priority services despite the lower priority ones. The rules used to resolve the conflicts are explained in detail in [4]. In order to overcome the problems stemmed from the dynamic nature of a realistic environment, where the actual interference of all the active users is considered, some modifications to the basic operation of the protocol have been carried out [5]. As a result, the designed MAC protocol becomes a stable and reliable support for the cross-layer architecture that provides a suitable estimation of the available bandwidth and allows to efficiently reserve the required resources.

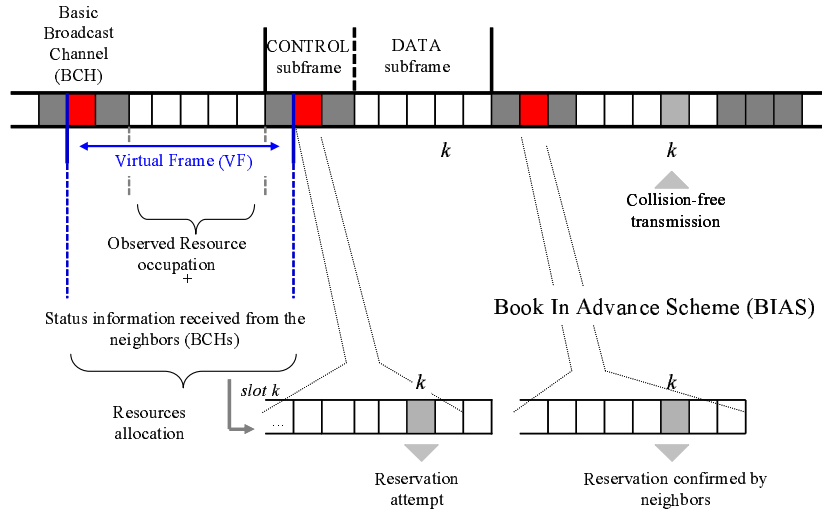


Fig. 1. Book in Advance Scheme (BIAS). Example of basic operation.

2.2 Admission control based on QoS Routing

A Distributed Admission Control (DAC) mechanism has been developed by defining a bandwidth metric required to establish a reliable path when trying to allocate a new application. According to this criterion, the routing protocol needs to find a connection path capable of satisfying the QoS requirements in terms of the demanded resources. The tradeoff between an efficient resource utilization and the required reservation that provides the desired QoS has a promising solution in the collaboration between the MAC level, responsible for the network resource management, and the routing protocol. The definition of a QoS metric that properly reflects the resource availability needs a correct estimation provided by the MAC level. In addition, in a TDMA structure, the end-to-end bandwidth cannot be estimated with the bottle-neck link (minimum local bandwidth), since links are not independent [6]. The hidden terminal problem imposes the disjoint property in three consecutive links requiring a global network perspective instead of a local vision [3], which can be achieved by including the end-to-end available bandwidth estimation in the routing process.

In this work, the Ad hoc On-demand Distance Vector (AODV) routing protocol [7] has been adapted to include BWC-FA (BandWidth Calculation - Forward Algorithm), a modified version of the proposal in [3], in order to measure the available bandwidth considering the whole path. The main features of this algorithm are described next. Additional information regarding slot availability is appended to the routing messages (request-RREQ, reply-RREP). During the discovery process, the proposed algorithm finds the available TDMA slots that can be used for transmitting in every link along the path so that these slots, if reserved, would be interference-free. The distributed bandwidth calculation

is supported by the measurements in the MAC level reflected in the QoS routing metric, which defines the available end-to-end bandwidth in the partial path from the source to each intermediate node, that updates the metric and evaluates if the QoS constraints are met in order to forward the RREQ. The final value calculated in the destination node represents the maximum available bandwidth in the whole path. Repeated RREQs are not directly dropped in the destination in order to select the best path according to this final metric. The decision considers the path with the minimum number of hops that supports the highest bandwidth. During the reply phase, information regarding the specific available resources is appended in the RREP in order to appropriately select the necessary slots to cover the QoS demands. Upon selection, the BIAS mechanism of the ADHOC MAC protocol is performed to ensure their reliable reservation. A more detailed explanation of the algorithm can be found in [5].

The specific identification of the available resources in every link makes it possible an estimation of the end-to-end available bandwidth closer to the real resource occupation than the classical local link measurement of the bottle-neck. An admission decision according to this measurement and the reservation of the specified resources increase the probability of admitting connections with the guarantee of a fulfillment of the QoS requirements [5]. The improvement in the effective capacity is the result of the establishment of guaranteed virtual circuits.

3 Operation in a mobility scenario

3.1 Reallocation and readmission procedures

The DAC mechanism allows to better allocate resources for the demanding applications without congesting the network over its capacity. However, the variability in the network conditions leads to a different scenario from that when the active connections were admitted. The established paths in the new topology may be unviable to provide the expected QoS, degrading the corresponding admitted connections. This degradation can be due to two differentiated effects: the initially non-colliding applications can interfere each other leading to unexpected collisions reducing the experienced bandwidth, or the lost of connectivity between neighbors may affect active paths, which will suffer from broken links and the consequent packet dropping. In any case, it is required to initiate a readmission capable of ensuring the QoS provision demanded by the admitted applications. The reallocation criterion, however, should depend on the nature of the suffered degradation. Fig. 2 shows the proposed admission scheme. The implemented procedures are summarized in this paper, but they are fully described in [8]. In this figure, P_b represents the probability of blocking a new admission and P_d is the dropping probability of admitted connections. In addition, due to the distributed nature of the developed procedures, some degraded applications are not correctly identified and they are maintained in the system. This, in fact, is equivalent to dropping for the perspective of the end user. According to this assumption, the failure probability P_f also includes the completed connections that do not experience the required QoS.

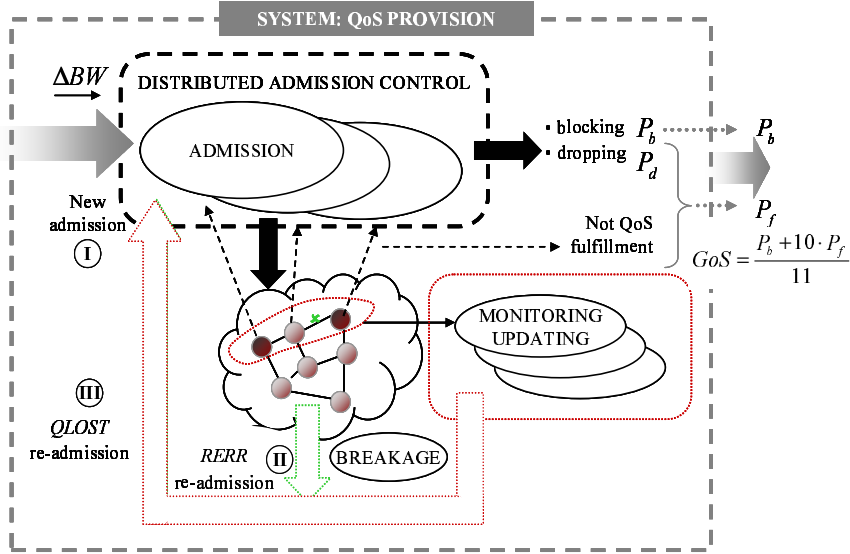


Fig. 2. Admission procedures. (I: new admissions, II-III: re-admission after breakage and QLOST)

As in the normal operation of the AODV, nodes react to broken links sending error messages that trigger a new discovery in the source of the connection. However, if the affected path requires certain QoS, a new attempt to allocate resources may fail due to the new topology which would lead to drop the connection increasing P_d . To reduce this probability, the QoS constraints are relaxed during the discovery including best-effort paths. If the connection is finally readmitted, a flexible resource reallocation or a new QoS readmission could finally provide the expected QoS, as it is explained next.

A proposed QoS monitoring process, based on the BWC-FA algorithm, is performed appending certain routing information piggy-backed in the DATA-ACK packets with a configurable updating period (t_{update}). The operation, similar to that in the RREQ-RREP exchange, allows to realize if the QoS constrains are not met anymore. If additional available bandwidth is found, new resources can be reserved without new discoveries, performing a flexible reallocation procedure. However, if the QoS cannot be maintained (after $n_{failupd}$ updates), a QLOST (QoS LOST) packet is sent to the source in order to trigger the discovery of a new path capable of satisfying the demanded bandwidth. This new discovery does not relax the routing constraints in order to find only a new QoS path. However, despite the degradation of the QoS, the former path is still viable to send traffic as best-effort, so it is maintained to avoid dropping packets in excess while discovering the new path. If it is found, the QoS is recovered. At the end, despite a temporary lost, certain soft-QoS is achieved allowing an acceptable performance. Nevertheless, if the changes in topology make unfeasible to reallocate

the connection with the demanded QoS ($t_{qlost,max}$ seconds without the required bandwidth), it is dropped in order to free resources for other applications. The configuration of the specified parameters can modulate this operation in order to define the correct response of the system according to the scenario (Table 1).

Table 1. Admission – Readmission procedure

	QoS flows (QoS Routing - BWC-FA)			best-effort
Admission criterion ^a	I	II	III	always
Discovery tries	$n_{disc,max}$	1	1	if $n_{pq} > 0^b$
RREQs dropping	✓	-	✓	-
Waiting in dest.	t_{wait}	0	t_{wait}	0
Guard bandwidth	ΔBW	0	0	0
Available priority ^c (preemptable resources ^d)	p_1 (p_2)	p_0 (p_2, p_1)	p_0 (p_2, p_1)	p_2 -

^a I: new connections, II: rediscovery after breakage y III: rediscovery after *QLOST*.

^b Queued packets.

^c Priority of the reserved resources if the connection is admitted. $p_2 < p_1 < p_0$

^d Reserved resources that can be reallocated for the connection to allow its admission.

In order to ensure a satisfactory Grade of Service ($GoS = (P_b + 10 \cdot P_d)/11$ [9]), a new attempt to allocate resources for an admitted connection must be preferential. Therefore, a flexible prioritization is performed in the MAC level in such a way that new applications cannot preempt resources reserved in virtual circuits, even during the best-effort operation (readmission after *QLOST*).

In addition to this basic readmission scheme, an additional parameter, ΔBW , can be configured to reduce the dropping probability. When a highly dynamic scenario leads to continuous degradations in the experienced QoS for the admitted connections, frequent new discoveries arise to readmit the disrupted ones. If the admission procedure allocates all the available resources, the variations in the admission conditions make it difficult to find sufficient alternative QoS paths to accommodate all these connections, which finally can increase the dropping probability. In order to provide extra free resources to flexibly reallocate the admitted connections when they are disrupted due to mobility, not only the demanded resources but also a guard bandwidth ΔBW are required during the initial admission. However, if admitted, only the necessary resources are effectively reserved. This ensures certain unused bandwidth available when required during a readmission.

3.2 Adaptive configuration

Despite the flexibility provided thanks to the reallocation and readmission procedures, the implemented mechanisms introduce control overhead that competes with the user data. In some cases, this occupation of priority resources can excessively reduce the bandwidth availability. In addition, when extra bandwidth

is underutilized (function of ΔBW), this reduction can be even more representative. Despite the better grade of service, a reduction in the effective capacity does not represent a global improvement in the network performance. The benefit of a flexible resource management becomes significant when a high grade of mobility leads to continuous route breakages or bandwidth degradations that make the maintenance of the admitted QoS connections unfeasible. According to the analysis of several scenarios, two different operation modes have been selected:

- **Quasy-Static Operation mode (QSOP):** In static or low mobility scenarios the control overhead degrades the effective capacity with a slight improvement in the GoS. Therefore, the QoS monitoring, the dropping capability and ΔBW are not configured.
- **Mobility Operation mode (MOP):** A flexible resource management due to the whole readmission procedure (including ΔBW) allows to improve the grade of service without a significant reduction in the effective capacity. The appropriate configuration has been selected thanks to an exhaustive analysis of the application scenarios (section 4): $n_{failupd} = 3$, $t_{update} = 0.3$ sec. and $t_{qlost,max} = 3$ sec.

Therefore, an appropriate configuration is highly dependent on the grade of mobility, which makes it difficult to define the adequate readmission parameters without a previous knowledge of the application scenario. In order to carry out a correct operation in a dynamic scenario it could be more interesting to adaptively configure the response of the system. To this purpose, an adaptation mechanism based on the estimation of the grade of mobility in the environment is proposed.

Estimation of the grade of mobility

The selected MAC protocol bases its operation on a correct measurement of the stable neighbors. The maintenance of this connectivity information allows to realize the existence of link disconnections, which are more likely to happen with a higher grade of mobility. Thus, the variability of the number of neighbors that the MAC level observes can be a good approximation of the dynamism in the surrounding area. This local measurement can represent the mobility in the network in deployment scenarios with geographically uniform movement, such as those modeled through the Random WayPoint (RWP). A metric according to this idea is defined as next:

$$var_{NB,i}^{filt}(t) = \alpha \cdot var_{NB,i}^{filt}(t - \Delta t) + (1 - \alpha) \cdot var_{NB,i}(t) \quad (1)$$

$$var_{NB,i}(t) = \left| \frac{NB_i^{est}(t) - NB_i^{est}(t - \Delta t)}{\Delta t} \right| \quad (2)$$

where NB_i^{est} is the number of stable neighbors for node i , $var_{NB,i}(t)$ is the temporal variation of this measurement and the metric $var_{NB,i}^{filt}$ is the result of filtering this variation in order to estimate the mean value of $var_{NB,i}$, which is

directly related to the grade of mobility. Δt has been selected to 1 sec. The parameter α should solve the tradeoff between determining an accurate estimation of the average value of $var_{NB,i}$ and a rapid response to its sudden variations. Due to the oscillations in the measurement, an stable configuration must be performed according to two defined thresholds (hysteresis):

- th_1 : $var_{NB,i}^{filt}$ to trigger the MOP configuration.
- th_2 : $var_{NB,i}^{filt}$ to change from the MOP to the QSOP configuration.

The value th_1 must be high enough to correctly identify the mobility environment avoiding the full readmission procedure in quasy-static scenarios, and it is related to the expected average $var_{NB,i}$ for the minimum mobility that requires a MOP configuration. The value th_2 tries to maintain an stable configuration of the readmission parameters in a mobility scenario without considering the metric variations, which must be absorbed by $(th_1 - th_2)$.

4 Performance evaluation

The proposed architecture (cross-layer QoS Routing - QoSR-CL) has been evaluated by means of simulation. With this purpose, we have built up an event-driven simulator in C++ which implements all the required functionalities. A set of 50 nodes are randomly positioned within a square area of 2 Km². Terminals follow a Modified Random WayPoint mobility model (MRWP), whose parameters are defined in Table 2. Speeds are measured in Km/h, time in sec. and distances in meters (topology grid: X_{max} , Y_{max} , equal to 1400 m, and $D_{max} = \sqrt{(X_{max}^2 + Y_{max}^2)}$). The additional parameters β_{max} and R_{max} define the maximum variation from the current direction and the maximum distance to the next destination for a new movement, which allows to follow a path closer to the natural movement. The simulation length has been selected in order to guarantee a stable topology distribution of the terminals according to the different mobility scenarios.

Table 2. Mobility scenarios. MRWP model.

Random Waypoint				MRWP	
V_{min}	V_{max}	T_{pausa}	V_m	β_{max}	R_{max}
1	3	2	~ 2	$\frac{\pi}{12}$	$0.1 \times D_{max}$
3	20	2	~ 9	$\frac{\pi}{12}$	$0.1 \times D_{max}$
10	40	2	~ 20	$\frac{\pi}{12}$	$0.1 \times D_{max}$

The connectivity among terminals is determined by the ability of decoding the BCH transmissions according to the received SIR, considering a transmitted power of 20 dBm, a Kammerman propagation model [5] and a minimum decoding threshold SIR_{th} of 5 dB. The sensing threshold CS_{th} is set to 1 dB.

An additional margin of 3.5 dB is selected to define the stable neighbors from the MAC level perspective [5]. Connections between different pairs of nodes are generated according to a Poisson process with rate [connections/sec.] ranging according to the simulated offered load and the mean connection duration (50 sec). QoS flows are generated as CBR sources demanding a bandwidth of 128 kbps (2 TDMA slots). Packets are early discarded to avoid an excessive delay. The tolerable packet loss P_{loss} is set to 3%, which defines the correctly dispatched traffic as the occupied bandwidth by the number of connections that ends with a throughput higher than 97%.

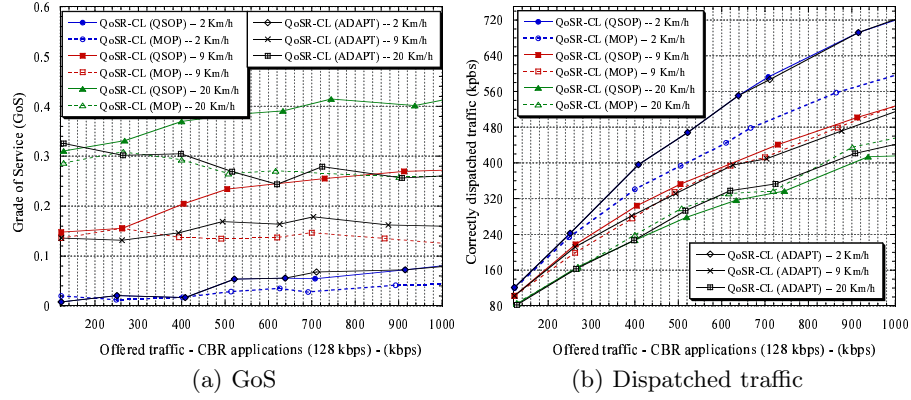


Fig. 3. Operation modes {QSOP, MOP and adaptive ($th_1 = 0.09$, $th_2 = 0.03$)} for different scenarios.

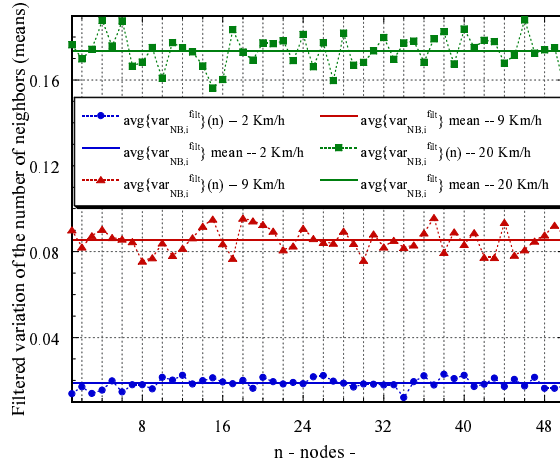


Fig. 4. Average filtered variation of the estimated number of neighbors. $\alpha = 0.99$.

The evaluation of the proposal in mobility conditions has shown the difficulties to guarantee the QoS in this dynamic scenario. As the average node speed increases, the number of correctly dispatched connections falls due to a growing packet dropping probability, since frequent route breakages or emerging interference disrupt the initially established virtual circuits. However, the correct configuration of the full readmission procedure (MOP operation) allows to improve the global performance in terms of grade of service (Fig. 3(a)) although only in scenarios with higher mobility (20 Km/h) the control overhead can be compensated with an increase in the effective capacity (Fig. 3(b)), thanks to a reduction of the failure probability (P_f). At lower speeds, P_f is not so significant and it cannot be extremely reduced, thus the increasing blocking probability due to both the overhead and the unused resources (ΔBW) has a negative effect over the effective capacity.

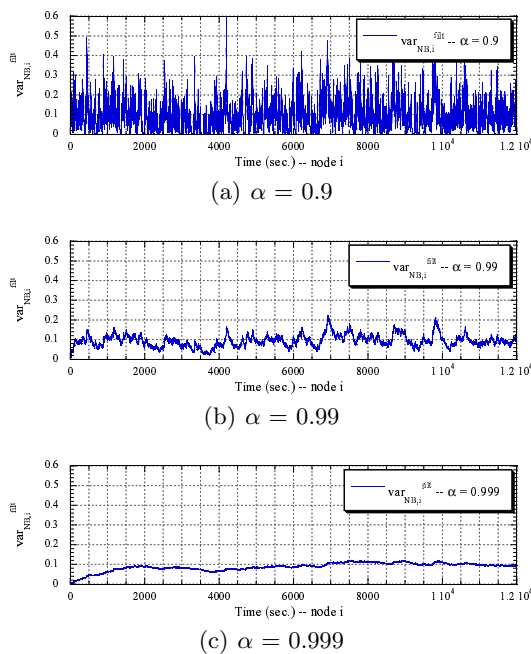


Fig. 5. Evolution of $var_{NB,i}^{filt}$ with different values of parameter α .

In order to guarantee a correct operation in any scenario, adaptively configuring the operation mode (QSOP or MOP), the mobility metric $var_{NB,i}^{filt}$ is defined. As it is shown in Fig. 4, this estimation provides a clear differentiation among scenarios. The selection of $\alpha = 0.99$ in Fig. 4 comes from the observed results represented in Fig. 5. Lower values (Fig. 5(a)) generate a more fluctuating metric that can complicate the desired differentiation. On the contrary, an excessive memory (Fig. 5(c)), despite the better average estimation, implies a slower convergence inadequate to adapt the system response to potential changes in the

mobility environment. The selected value provides a good approximation to the average $var_{NB,i}$ with a rapid adaptation, as it is observed in Fig. 5(b). Due to the metric oscillation, the adaptation is made according to a pair of thresholds (th_1 and th_2).

Results observed at 9 Km/h show this grade of mobility as an inflection point between both operation modes (QSOP, MOP). At lower speeds, the overhead of the MOP configuration cannot be assumed whereas, at a higher mobility, the global performance is improved thanks to the more flexible resource management. The mean and standard deviation of $var_{NB,i}^{filt}$ in the scenario of 9 Km/h (0.0865 and 0.035 respectively) have been considered as a reference. Finally, after the evaluation of different configurations, results have shown the convenience of a selection $\{th_1 = 0.09 - th_2 = 0.03\}$. The resulting ratio of time that corresponds to a MOP configuration is 0, 0.86 and 0.98 for 2, 9 and 20 Km/h respectively. As it is shown in Fig. 3, this provides an adaptation near the best performance in any mobility scenario.

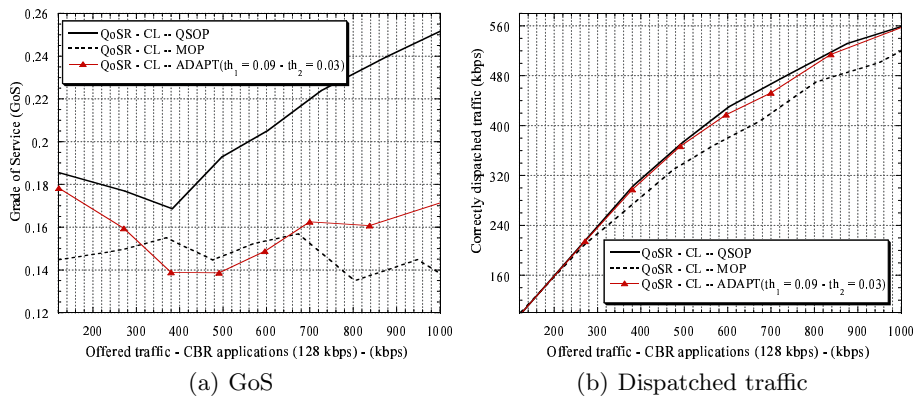


Fig. 6. Performance in a scenario with variable grade of mobility.

Obviously, a previous configuration adapted to the specific grade of mobility provides a more adjusted definition of the readmission parameters. However, an adaptive solution allows the system to auto-configure the procedure without a prior knowledge of the mobility scenario. In addition, variable mobility conditions can make it difficult to properly define the best operation mode. In order to evaluate the proposed solution in a mobility variant environment, a more dynamic scenario has been analyzed. The mobility pattern follows the MRWP model with three consecutive grades of mobility (2, 9 and 20 Km/h).

The proposed mobility pattern could suggest a MOP configuration as appropriate due to the highest percentage of high mobility (9 and 20 Km/h). However, despite the GoS improvement (Fig. 6(a)) the representative reduction on the effective capacity shows the impact of an incorrect configuration during the quasi-static mobility pattern (Fig. 6(b)). On the contrary, the adaptive readmis-

sion procedure allows to obtain a grade of service close to that provided by the MOP configuration thanks to the flexible reallocation (Fig. 6(a)) maintaining the effective capacity, evaluated as the correctly dispatched traffic, as it is shown in Fig. 6(b).

5 Conclusions

A distributed admission control designed for MANETs has been evaluated in different mobility scenarios. A cross-layer architecture based on a QoS Routing allows to allocate different applications according to their demanded end-to-end bandwidth establishing virtual circuits that guarantee the QoS provision. However, the dynamic nature of the environment leads to changes in the topology that modify the admission conditions, which has motivated the proposal of a flexible reallocation scheme. Results have revealed the difficulties to properly configure this mechanism in different mobility conditions. A proposed mobility estimation based on the measurement of the neighbors variability allows to implement an adaptive parametrization that performs a flexible management capable of guaranteeing a reasonable performance in different scenarios.

References

1. Chen, S., Nahrstedt, K.: An overview of quality-of-service routing for the next generation high-speed networks: Problems and solutions. *IEEE Network Magazine* (1998) 64–79
2. Zhou, B., Marshall, A., Wu, J., Lee, T.H., Liu, J.: A cross-layer route discovery framework for mobile ad hoc network. *EURASIP Journal on Wireless Communications and Networking* **5** (2005) 645–660
3. Zhu, C., Corson, M.: QoS routing for mobile ad hoc networks. In: *Proc. IEEE INFOCOM'02*. Volume 2., New York, USA (2002) 958–967
4. Gállego, J.R., Canales, M., Hernández-Solana, A., Campelli, L., Cesana, M., Valdovinos, A.: Performance evaluation of point-to-point scheduling strategies for the ADHOC MAC protocol. In: *Proc. of WPMC'05*, Aalborg, Denmark (2005) 1380–1384
5. Canales, M., Gállego, J.R., Hernández-Solana, A., Valdovinos, A.: Performance evaluation of cross-layer routing for QoS support in mobile ad hoc networks. *Springer LNCS (IFIP PWC'06)* **4217** (2006) 322–333
6. Chen, T.W., Tsai, J., Gerla, M.: QoS routing performance in multihop, multimedia, wireless networks. In: *Proc. of IEEE ICUPC'97*. Volume 2., San Diego, CA. USA (1997) 557–561
7. Perkins, C., Belding-Royer, E., Das, S.: Ad hoc on-demand distance vector (AODV) routing. *Experimental RFC 3561*, The IETF Network Working Group (2003) <http://www.ietf.org/rfc/rfc3561.txt>.
8. Canales, M., Gállego, J.R., Hernández-Solana, A., Valdovinos, A.: Performance analysis of cross-layer QoS routing for mobile ad hoc networks. In: *Proc. of WPMC'06*, San Diego, USA (2006) 946–950
9. Dimitriou, N., Tafazolli, R., Sfikas, G.: Quality of service for multimedia CDMA. *IEEE Communications Magazine* **38**(7) (2000) 88–94