Design and Analysis of a Denial-of-Service-Resistant Quality-of-Service Signaling Protocol for MANETs

Marek Hejmo, Brian L. Mark, *Member, IEEE*, Charikleia Zouridaki, *Student Member, IEEE*, and Roshan K. Thomas

***Abstract*—Quality-of-service (QoS) signaling protocols for mobile *ad hoc* networks (MANETs) are highly vulnerable to attacks. In particular, a class of denial-of-service (DoS) attacks can severely cripple network performance with relatively little effort expended by the attacker. A distributed QoS signaling protocol that is resistant to a class of DoS attacks on signaling is proposed. The signaling protocol provides QoS for real-time traffic and employs mechanisms at the medium access control (MAC) layer, which serve to avoid potential attacks on network resource usage. The key MAC layer mechanisms that provide support for the QoS signaling scheme include sensing of available bandwidth, traffic policing, and rate monitoring, all of which are performed in a distributed manner by the mobile nodes. The proposed signaling scheme achieves a compromise between signaling protocols that require the maintenance of per-flow state and those that are completely stateless. The signaling scheme scales gracefully in terms of the number of nodes and/or traffic flows in the MANET. The authors analyze the security properties of the protocol and present simulation results to demonstrate its resistance to DoS attacks.**

***Index Terms*—Cross-layer design, denial-of-service (DoS), mobile *ad hoc* networks (MANETs), quality-of-service (QoS) signaling.**

I. Introduction

QUALITY-OF-SERVICE (QoS) provisioning for wireless networks is becoming increasingly important as more real-time applications migrate to the wireless environment. Providing QoS in a mobile *ad hoc* network (MANET) is especially challenging due to the node mobility, the lack of a fixed infrastructure, the limitations of the wireless channel, and the limited resources of the mobile nodes. Recently, several QoS signaling protocols for MANETs have been proposed in the research literature [1], [2]. However, these schemes were not designed with security in mind and are highly vulnerable to attacks, in particular, denial-of-service (DoS) attacks.

QoS signaling mechanisms can be categorized as reservation-based or reservation-less, depending on whether or not the mechanism makes explicit reservations of network resources for traffic flows. Reservation-based mechanisms typically require the maintenance of per-flow state, which limits their scalability and makes them vulnerable to state table exhaustion, a well-known DoS attack. Reservation-less schemes are scalable, but can be more vulnerable to other types of DoS attacks such as flooding and overreservation of resources.

The main contribution of this paper is a cross-layer architecture for QoS signaling in MANETs, which provides resistance to a class of DoS attacks. The proposed DoS-resistant QoS (DRQoS) signaling scheme employs distributed rate control to manage the bandwidth resources of the network, but does not rely on the maintenance of per-flow state.[1] In the DRQoS scheme, each mobile node maintains a state table of bandwidth reservations, which grows as a function of the number of neighbor nodes rather than the number of traffic flows traversing the node. The DRQoS protocol provides QoS signaling on top of an arbitrary MANET routing protocol and employs mechanisms at the medium access control (MAC) layer for QoS provisioning and resistance to attacks in conjunction with the signaling protocol. The key MAC layer elements of the scheme consist of estimating the available wireless bandwidth, traffic policing, and rate monitoring, all of which are performed in a distributed manner in the network.

The remainder of the paper is organized as follows. Section II provides an overview of QoS signaling mechanisms for MANETs and discusses the vulnerabilities of current QoS schemes to DoS attacks. Section III describes the operation of the proposed DRQoS signaling scheme. Section IV analyzes the DoS resistance and scalability properties of the DRQoS scheme. Section V presents ns-2 simulation results demonstrating the key properties of DRQoS. Finally, Section VI concludes the paper.

II. QoS Signaling in MANETs

A number of approaches to providing quality-of-service in MANETs have been proposed in the literature. Several QoS schemes are designed as QoS extensions to MANET routing protocols such as *Ad Hoc* On Demand Distance Vector (AODV) and Optimized Link State Routing Protocol (OLSR) over a best effort MAC layer such as IEEE 802.11 DCF (cf. [4]–[6]). Other QoS routing protocols are designed explicitly with some form of QoS support (cf. [7], [8]). Several QoS routing protocols for MANETs assume a MAC layer based on Time-Division Multiple Access (TDMA) (cf. [9], [10]). The DRQoS protocol proposed in the present paper is closest in

Manuscript received October 24, 2005; revised November 9, 2005. This work was supported in part by the National Science Foundation under Grant CCR-0209049. An earlier version of this paper was presented at the IEEE/ACM QShine'05 Conference. The review of this paper was coordinated by Prof. X. Shen.

M. Hejmo, B. L. Mark, and C. Zouridaki are with the Department of Electrical and Computer Engineering, George Mason University, Fairfax, VA 22030 USA (e-mail: mhejmo@gmu.edu; bmark@gmu.edu; czourida@gmu.edu).

R. K. Thomas is with the SPARTA, Inc., Centreville, VA 20120 USA (e-mail: roshan.thomas@sparta.com).

Digital Object Identifier 10.1109/TVT.2006.873834

[1] A preliminary version of the DRQoS scheme was first proposed in [3].

spirit to the INSIGNIA [1] and Stateless Wireless *Ad Hoc* Networks (SWAN) [2] protocols, which are QoS signaling protocols designed to operate above an arbitrary MANET routing protocol with an underlying best effort MAC layer.

A. Stateful Versus Stateless QoS Signaling

The INSIGNIA protocol is a representative stateful or "reservation-based" signaling scheme, whereas SWAN is a stateless or "reservation-less" scheme. The INSIGNIA scheme uses in-band signaling, whereby control information is piggybacked in the IP options field of the IP datagram. By combining soft state with in-band signaling, INSIGNIA can respond quickly to route breakages. However, the maintenance of per-flow state information does not scale well with network size and mobility and may not be feasible for typical mobile devices, which are limited in terms of storage, battery life, and computational power.

In the SWAN scheme, the source node probes a given route to determine whether sufficient resources are available to support a new real-time flow. If enough bandwidth resources are available along the path to the destination node, the source node initiates data transmission for the real-time flow. Otherwise, the source may probe another route, or it may reduce its own transmission rate to accommodate the amount of resources reported by the bandwidth probe. In either case, no resources are explicitly reserved for the real-time flow. The source simply transmits at the rate determined by the bandwidth probing phase.

B. Vulnerabilities of QoS Signaling

QoS signaling in MANETs introduces new vulnerabilities that are not addressed by secure routing primitives (cf. [11]). Attacks on routing are generally directed toward disrupting network connectivity, whereas attacks targeted at QoS signaling need not affect connectivity. For example, a route that is established by means of a secure routing protocol can still be susceptible to attacks on QoS. If an attacker manages to compromise the key needed for network authentication, it can become part of a "secure" route. Such a node may comply with a secure routing protocol, but at the same time attack and exploit the signaling protocol.

Securing QoS signaling is challenging because some attacks against signaling may be difficult to distinguish from legitimate network congestion conditions or loss of connectivity. Attacks on confidentiality, integrity, and accountability can be mitigated by appropriate cryptographic protection on QoS signaling messages such as those based on digital signatures, message authentication codes, etc. In this paper, we focus on attacks that impact the availability objective. Attacks that target the availability objective lead to DoS [12] by exploiting limited link resources such as bandwidth and node resources such as energy, memory, and CPU.

C. DoS Attacks on QoS Signaling

The proposed DRQoS protocol specifically addresses the class of DoS attacks comprising flooding, overreservation, and state table exhaustion. In general, these attacks are more damaging and are capable of being launched more easily in MANETs than in wired networks. In the "flooding attack," the attacker sends traffic into the network at a rate higher than a "negotiated rate." For example, in INSIGNIA, the negotiated rate is the reserved rate for the given traffic flow. In SWAN, the negotiated rate is the rate returned by the network, which represents the available bandwidth along a path in response to a bandwidth request probe. A flooding attack expends the resources of the network on illegitimate traffic, resulting in a DoS condition for legitimate sources. One technique to mitigate flooding that is used in wired networks is to trace back the attacker and cut off the attack traffic as close to the source as possible. However, tracing back an attacker in a MANET is not typically feasible due to the node mobility.

In the "overreservation attack," the sender reserves a transmission rate with the network that is much higher than the rate at which it generates traffic. We remark that the overreservation attack is specific to reservation-based protocols such as INSIGNIA. The SWAN protocol is not vulnerable to this type of attack. An overreservation attack does not consume the resources of the network, but locks out legitimate sources that could make use of the unused bandwidth that has been reserved by the attacker. This is a DoS condition that requires relatively little effort for the attacker to create. The "state table exhaustion attack" affects only reservation-based signaling schemes such as INSIGNIA. The attacker causes the state table to be exhausted by issuing a large number of reservation requests to the victim node. In MANETs, mobile devices are typically highly constrained in terms of memory and can store only a limited amount of state information. By consuming the memory and computational resources of the victim node, the attacker causes a DoS condition for other nodes that would otherwise use the victim node in multihop routes.

III. Specification of DRQoS

A. Overview

The DoS-resistant QoS signaling scheme aims to provide QoS for real-time traffic while providing protection against DoS attacks. The basic mechanism for DoS protection is a rate control scheme that polices traffic flows in a distributed manner. The DRQoS scheme avoids the storage of per-flow state. In the DRQoS scheme, each node maintains state for each active aggregate traffic stream between an input/output port pair. The aggregate "in-out" traffic stream through a node may consist of many individual traffic flows. However, a given node is responsible only for policing the in-out traffic streams that traverse the node. Therefore, the amount of state information stored at each node is a function of the number of neighbor nodes, rather than the number of flows traversing the node. If an individual flow transmits above its assigned rate, it may experience traffic policing from at least one of the intermediate nodes on the associated path as a side effect of the control mechanisms operating on an in-out stream basis.

Similar to SWAN [2], real-time traffic flows are established by a protocol involving a bandwidth probing phase followed by a data sending phase. We shall assume that real-time packets

are scheduled with priority over best effort packets at a given node. In addition, a mechanism must be in place to isolate real-time traffic from the effects of best effort cross-traffic. This could be achieved with a QoS-based MAC [13], [14] or with a local rate control mechanism for best effort traffic [2], [15] if the underlying MAC provides only best effort service. In the remainder of the paper, we shall focus on QoS provisioning and DoS protection for real-time traffic.

B. State Table

DRQoS is a stateful protocol in the sense that a given node x maintains state for aggregate traffic streams traversing node x on an in-hop/out-hop basis, i.e., a state table entry is maintained for each pair (i, j), where i and j denote one-hop neighbors of the node x. The aggregate traffic generated by a set of flows traversing the subpath $\{i, x, j\}$ is referred to as an in-out stream corresponding to (i, j). Node x maintains a record of the traffic rate corresponding to each in-out stream traversing it. The total number of state table entries is determined by the number of active in-out streams, which is at most $N(N-1)$, where N is the number of neighbors of the given node x. The number of active in-out streams will typically be significantly smaller than the number of individual flows traversing node x. By avoiding the maintenance of per-flow state, DRQoS is much less vulnerable to state table exhaustion attacks than protocols such as INSIGNIA. On the other hand, by maintaining state on an in-hop/out-hop basis, DRQoS is scalable with respect to flows and can offer resistance to a class of DoS attacks that would incapacitate other signaling protocols.

In the DRQoS state table, the (i, j)th entry records the following information: 1) assigned rate R^*_{ij} corresponding to in-out stream (i, j), 2) counter X_{ij} for the number of bits that have arrived in the current measurement window, and 3) measured rate $\hat{R}_{ij}$ from the previous measurement window. Each DRQoS node is responsible for policing the in-out stream (i, j) to the assigned rate R^*_{ij}. The measured rate $\hat{R}_{ij}$ is used to perform rate adjustment, as will be discussed shortly.

C. Control Packets

Similar to the SWAN protocol, DRQoS consists of two phases: 1) bandwidth probing phase and 2) data transmission phase. Two control messages are sent during the bandwidth probing phase: 1) bandwidth probe request (BPReq) and 2) bandwidth probe reply (BPRep). The BPReq packet contains the source IP address, destination IP address, type of the message, flow ID, and requested data rate stored in the bottleneck bandwidth (BB) field.

As in SWAN, to initiate a real-time flow along a given route, the source node sends a BPReq packet to the destination. Upon receiving a BPReq packet, an intermediate node along the path from source to destination determines the "available bandwidth" on its outgoing link (see Section III-E). If the available bandwidth A_j on the outgoing link to the next hop j is greater than the BB value stored in the BPReq packet, the node forwards the packet to the next node (i.e., node j) on the path. Otherwise, the node replaces the BB field of the BPReq packet with the available bandwidth A_j and forwards the packet to the next node. When the destination node receives the BPReq packet, it copies the value of the BB field to the BB field of a new BPRep packet. The BPRep packet is then sent back to the source node using the reverse path.

Our scheme departs from SWAN in the processing of a BPReq on the reverse path. Unlike SWAN, bandwidth probing in DRQoS involves the manipulation of node state table information along a path. Upon receiving a BPRep message on the reverse path, an intermediate node updates its state table using the BB value stored in the BPRep packet and then forwards the BPRep to the next node. The state table is updated in the following way. Define the "in-hop" node i to be the next node to which the BPRep will be sent. The "out-hop" node j is the node from which the BPRep packet was received. First, the available bandwidth A_j is checked. If the value of A_j is greater than or equal to the BB value in the BPRep packet, the reservation of bandwidth for the flow can proceed. Otherwise, the BB value in the BPRep packet is overwritten with the (smaller) value A_j. Next, if a state table entry for in-out stream (i, j) already exists, i.e., the stream is active, the current BB value in the BPRep packet is added to the reserved rate R_{ij}, associated with the in-out stream. If the stream (i, j) was previously inactive, a state table entry is created with an assigned rate value R_{ij}, set equal to the BB value of the BPRep packet. Then, the BPRep packet is forwarded to the next node on the reverse path (i.e., node i). Finally, when the BPRep packet reaches the source node, the source establishes the real-time flow based on the value of the BB field.

D. Distributed Rate Control

The DRQoS protocol includes a distributed rate control mechanism consisting of two components: 1) traffic policing and 2) rate monitoring and adjustment. Traffic policing ensures that a real-time in-out stream traversing a given node does not exceed the rate recorded in the state table. Rate monitoring and adjustment implements a "use it or lose it" policy for real-time in-out streams, whereby the rate of an in-out stream is measured and compared with the assigned rate recorded in the state table. If the measured rate is lower than the reserved rate by a sufficient margin, the reserved rate is decreased by a certain factor.

Traffic policing for an in-out flow can be accomplished by means of a sliding window or a leaky bucket mechanism (cf. [16]). As defined above, R^*_{ij} denotes the assigned rate for in-out stream (i, j). The actual rate that is used to police the traffic stream is defined by $\tilde{R}_{ij} \triangleq \gamma R^*_{ij}$, where $0 < \gamma \leq 1$ is a reduction factor defined by

$$\gamma \triangleq \min\left\{\frac{C_j}{R^*_j}, 1\right\} \tag{1}$$

where C_j is the estimated link capacity (see Section III-E) and R^*_j is the aggregate assigned rate for out-hop j defined by

$$R^*_j \triangleq \sum_i R^*_{ij}. \tag{2}$$

Reducing the assigned rate R_{ij} by the factor γ ensures that the sum of the policed rates R^*_{ij} over all in-hops i does not exceed the estimated link capacity C_j. The in-out stream (i, j) can be policed by a leaky bucket with leak rate R^*_{ij} and a bucket size B, which allows for some delay variation tolerance. Packets that are in violation of the parameters (R^*_{ij}, B) would either be dropped immediately, marked as low priority (i.e., best effort), or delayed to force conformance to the leaky bucket parameters.

The rate monitoring function measures the traffic rate of a given in-out stream over a time interval $\hat{T}$. Rate monitoring could be accomplished by keeping a counter of the total number of bits arriving on an in-out stream over the period $\hat{T}$. As each new packet arrives on a given in-out stream (i, j), a counter X_{ij} is incremented by the size of the packet in bits. After the time period of $\hat{T}$ elapses, as indicated by expiry of a time, the measured rate $\hat{R}_{ij}$ is simply computed as $\hat{R}_{ij} = X_{ij}/\hat{T}$. If the measured traffic rate $\hat{R}_{ij}$ is less than the assigned in-out rate R^*_{ij} by more than a certain percentage p_d, then the assigned in-out rate R^*_{ij} is decreased by a factor $1 - \alpha p_d$, where $\alpha \in (0, 1)$ is a design parameter. This is one aspect of the "rate adjustment" step. If the assigned traffic rate for an in-out stream is decreased below a threshold $r_{\min}$, then the in-out stream is removed from the state table, i.e., it is treated as inactive.

Congestion occurring on an outgoing link is indicated by a large queue size associated with an output port. Such congestion can be alleviated by decreasing the assigned rates for all in-out streams destined to the output port by a factor $1 - \beta$, where $\beta \in (0, 1)$ is a design parameter. This is another aspect of rate adjustment, whereby congestion is alleviated by an explicit adjustment of the assigned rates for in-out streams destined to the congested output port.

E. Available Bandwidth Estimation

In the DRQoS scheme, each node is responsible for estimating the available bandwidth on its local outgoing link. For a given node, let A_j and C_j denote, respectively, the available bandwidth and link capacity on the outgoing link associated with out-hop j. Let R^*_j be the aggregate assigned rate for all in-out streams destined for out-hop j, as defined in (2). Ideally, the sum of the assigned in-out stream rates and the available bandwidth on link j should equal the capacity of link j; i.e., the following equation should hold in principle:

$$R^*_j + A_j = C_j. \tag{3}$$

In practice, the quantities A_j and C_j can be estimated through traffic measurements. Unlike a wired network, link capacities in an *ad hoc* network are not known *a priori* and frequently change dynamically due to node mobility and the time-varying conditions of the wireless channel.

Recall that during the bandwidth probing phase (of both SWAN and DRQoS), the source node of a flow sends a BPReq packet along a given path to the destination node. The BPReq packet contains a requested bandwidth value stored in the BB field. Each intermediate node is then responsible for determining whether or not sufficient bandwidth is available on the local outgoing link to support the new flow request. In SWAN, the available bandwidth A_j on an outgoing link j is measured directly. However, in DRQoS, the link capacity C_j is measured and the "available bandwidth" is defined by

$$A_j \triangleq \max\{0, C_j - R_j\}. \tag{4}$$

In DRQoS, (4) is used by the intermediate node to determine the value that should be recorded in the BB field of a BPReq packet in transit. This definition of A_j explicitly takes into account the amount of bandwidth that has been reserved for (and used to police) the in-out streams passing through out-hop j. This approach eliminates the problem of "false admission" in the SWAN scheme and avoids the need for the explicit congestion notification (ECN) and the timer-based regulation mechanisms employed in SWAN [2].

In the context of DRQoS, the link capacity C_j represents the total amount of consumed and available bandwidth for transmission over link j, taking into account medium access contention on the wireless channel. Estimation of C_j depends on the type of MAC layer used in the network. For the IEEE 802.11 DCF MAC layer, the link capacity can be estimated by considering the throughput for a successful packet transmission defined by (cf. [17], [18])

$$T = \frac{S}{t_r - t_s} \tag{5}$$

where S is the size of the packet in bits, t_s is the time at which the packet enters the MAC layer queue, and t_r is the time at which the corresponding ACK is received. Clearly, the per-packet throughput T is an increasing function of the packet size S. To make the per-packet throughput measurement independent of packet size, it can be normalized with respect to a predefined standard packet size as proposed in [18]. To obtain meaningful estimates of link capacity, per-packet throughput measurements should be smoothed over a suitably defined packet window [17], [18]. An alternative approach to estimating link capacity in an *ad hoc* network is proposed in [15], where the concept of fraction of air time (FAT) is introduced. In the context of DRQoS, the link capacity C_j is equivalent to the sum of the consumed and residual FAT for link j as defined in [15].

IV. DoS Resistance of DRQoS

In this section, we analyze the DRQoS protocol's resistance to the flooding, overreservation, and state table exhaustion attacks. We do not specifically address attacks directed against the lower layers of the protocol stack, e.g., routing protocol attacks, MAC layer attacks, and physical layer jamming. Further, we focus only on real-time sessions requiring QoS. In particular, we do not address the DoS problem for best effort traffic.

We define a node to be "DRQoS compliant" if it follows the DRQoS protocol as specified in Section III. Any DoS attack will be stopped at the closest DRQoS-compliant node downstream from the attacker. Hence, the DoS attack scenarios can be reduced to the situation shown in Fig. 1, which consists of a malicious node X and a one-hop neighbor node o that is

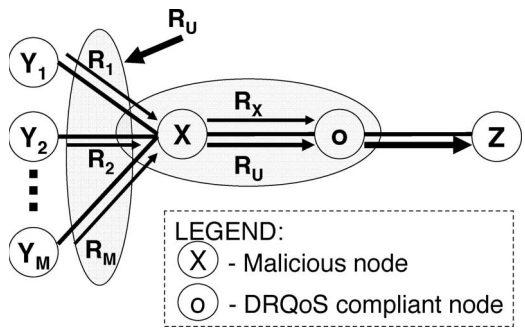

Fig. 1. General attack scenario.

DRQoS compliant. Node o maintains an entry (X, Z) in its state table for the aggregate traffic from its neighboring node X and destined for its neighbor node Z. The traffic originating from node X has an associated rate R_X. The traffic arriving from the nodes upstream from node X has an aggregate rate R_U. The aggregate traffic from node X has rate $R_U + R_X$.

A. Real-Time Flooding

Perhaps the simplest violation of the DRQoS signaling protocol is for the source node to send traffic without first initiating the bandwidth probing phase. In this scenario, the source node's DRQoS-compliant neighbor, say node I_1, simply drops (or marks as low priority) all traffic from the source node due to the lack of a state table entry corresponding to the source node. Note that the SWAN protocol is susceptible to this attack because no state information is stored at intermediate nodes. Under SWAN, the source node would cause a DoS condition for all other traffic flows passing through node I_1.

Even if the source node issues a proper BPReq packet, it could ignore the negotiated rate returned in the BPRep packet and send at a higher rate. In Fig. 2, the source node S_1 sends at a rate R_{S_1} that exceeds the negotiated rate $R^*_{S_1}$. The DRQoS-compliant node I_1 maintains a state table entry for the traffic stream originating at S_1 and polices the outgoing traffic at the negotiated rate. Thus, other nodes downstream from I_1 are insulated from the flooding attack of node S_1. Many variations of the flooding attack are possible, but in each case, the first DRQoS-compliant node on the path of a flooding attack will effectively quench the attack.

B. Overreservation Attack

The overreservation attack and the corresponding DRQoS response is summarized in Fig. 3. Here, the attacking node S_1 sends at a rate R_{S_1} that is far below its negotiated rate $R^*_{S_1}$. The DRQoS-compliant node I_1 detects the mismatch between the measured rate of the traffic stream and reduces the negotiated rate $R^*_{S_1}$ until it matches the actual traffic rate R_{S_1}. This frees the otherwise wasted bandwidth on the outgoing link from node I_1 for other flows to use in the future. Variations of the basic overreservation attack are possible, but as in the flooding attacks, the first DRQoS-compliant node downstream from the attacking node will prevent a DoS condition from arising due to the overreservation.

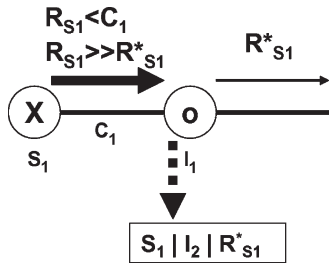

Fig. 2. Flooding attack.

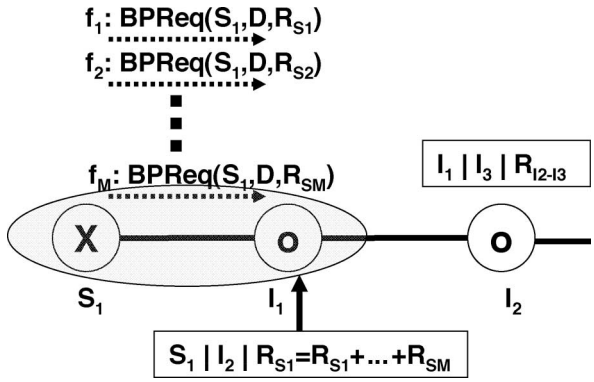

Fig. 3. Overreservation attack.

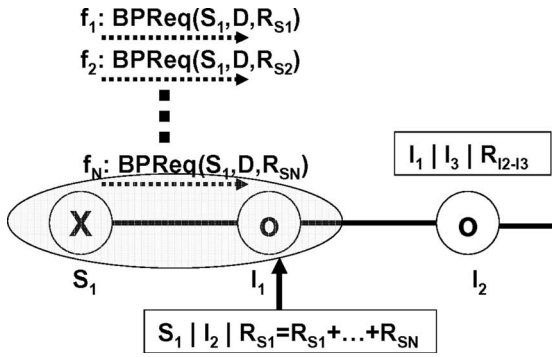

Fig. 4. State table exhaustion attack.

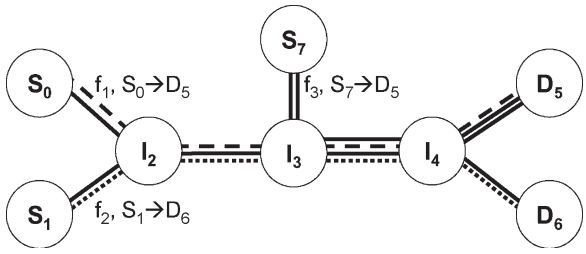

Fig. 5. Topology for flooding attack.

C. State Table Exhaustion Attack

Fig. 4 depicts the basic state table exhaustion attack. Here, the attacker node S_1 issues BPReqs for multiple flows $f_1, \ldots, f_M$. The DRQoS-compliant node I_1 creates only a single state table entry for the aggregate stream $(S_1, I_2) = f_1 + \cdots + f_M$. In the worst case, the attacking node S_1 can cause $N - 1$ state table entries to be created in node I_1, where N is the number of one-hop neighbors of I_1. The value of N

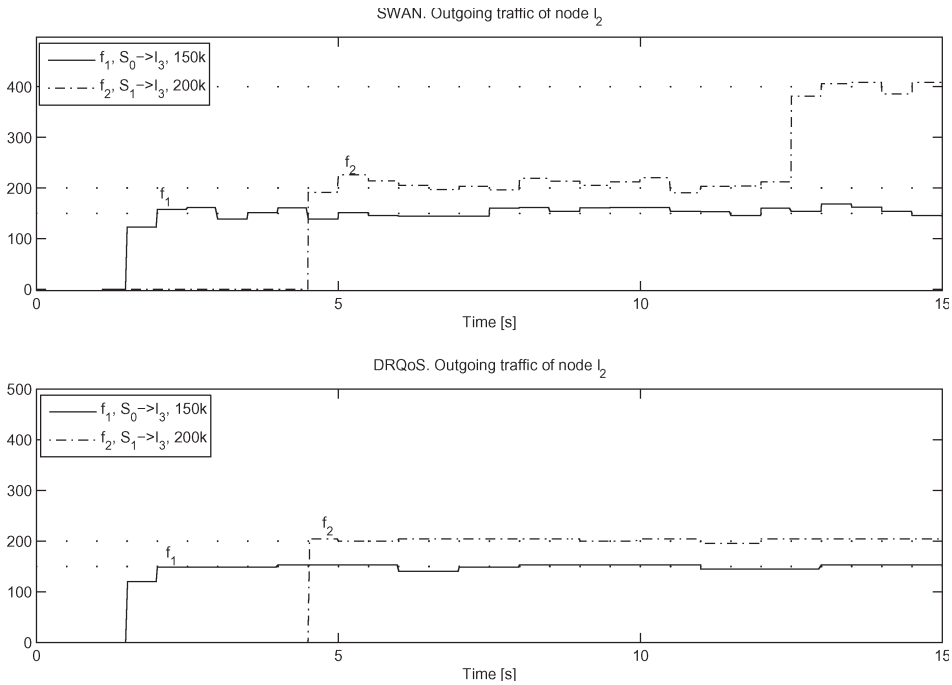

Fig. 6. Flooding attack that does not cause network congestion.

is typically a small number in MANETs, e.g., $N = 10$. In a variation of this attack, node S_1 could spoof the identities of its neighbor nodes, thus causing the creation of $N(N-1)$ state table entries in node I_1, one for each possible in-out traffic stream traversing node I_1. If $N = 10$, then 90 state table entries would be created, which should be well within the storage and computational capabilities of modern mobile devices.

V. Simulation Results

The DRQoS protocol was implemented and evaluated in the ns-2 network simulation environment [19]. We present three simulation scenarios: two involving flooding attacks and the third involving an overreservation attack. In the flooding attack scenario, the performance of DRQoS is compared with that of the SWAN protocol.[2] The network topology shown in Fig. 5 has been used to simulate flooding attacks. Each wireless link consists of two nodes, which are 200 m away from each other. The wireless radio transmission range is set to 250 m.

A. Flooding Attacks

In the first flooding scenario, two traffic flows are established in the network: 1) flow f_1 along a path from $S_0 \rightarrow D_5$ and 2) flow f_2 along a path from $S_1 \rightarrow D_6$. Node S_0 negotiates a rate of 150 kbps for flow f_1 and begins sending traffic on flow f_1 at a rate of 150 kbps at time $t = 1.5$ s. Node S_1 negotiates a rate of 200 kbps for flow f_2 and starts sending traffic on flow f_2 at time $t = 4.5$ s with a rate of 200 kbps. At time $t = 12.5$ s, node S_1 doubles its transmission rate to 400 kbps. Under the SWAN protocol, all of the intermediate nodes I_2, I_3, and I_4 forward the traffic from node S_1, as shown in the top graph of Fig. 6. In this case, the flooding attack causes the intermediate nodes to waste their battery and bandwidth resources. As shown in the bottom graph of Fig. 6, under DRQoS, the flooding attack from flow f_2 is stopped at node I_2, which polices the flow to the original rate negotiated by source node S_1, i.e., 200 kbps. Here, the downstream nodes I_2, I_3, and I_4 are effectively insulated from the flooding attack.

In the second flooding scenario, we assume the network topology of Fig. 5, but the link capacities are set to 600 kbps, and a new flow, f_3, is added. Flow f_3 is established from node S_7 to node D_6 at a negotiated rate of 250 kbps and begins transmission at time $t = 9$ s. Under SWAN, as long as flow f_2 does not exceed its negotiated rate, no congestion occurs in the network. The total rate of the real-time traffic on link $I_3 - I_4$ equals the capacity of 600 kbps. However, when node S_1 doubles its transmission rate to 400 kbps, network congestion occurs. The total rate of the incoming traffic to link $I_3 - I_4$ equals 800 kbps, but the link capacity is only 600 kbps.

To deal with network congestion SWAN employs an ECN mechanism [2] wherein a node that experiences congestion (I_3 in this case) marks the congestion experienced (CE) bit in the IP header of every packet that belongs to a randomly chosen flow traversing the congested node. Once the destination node receives a packet with the CE bit set, it sends a special "regulate" control message to the source of the marked flow, which forces the source to reestablish its congested flow. In the simulated scenario, flows f_1 and f_3 are legitimate, whereas flow f_2 is malicious. Thus, the probability of forcing a legitimate flow to reestablish its session is 2/3, whereas the probability of forcing the malicious flow to reestablish is 1/3. As shown in the top graph of Fig. 7, the legitimate flow f_3 was forced to reestablish its real-time flow. Inasmuch as the link's capacity cannot support the new request of S_7, f_3 cannot be reestablished as a

[2]An ns-2 code implementation of the SWAN protocol is available at http://www.comet.columbia.edu/swan/sourcecode.html.

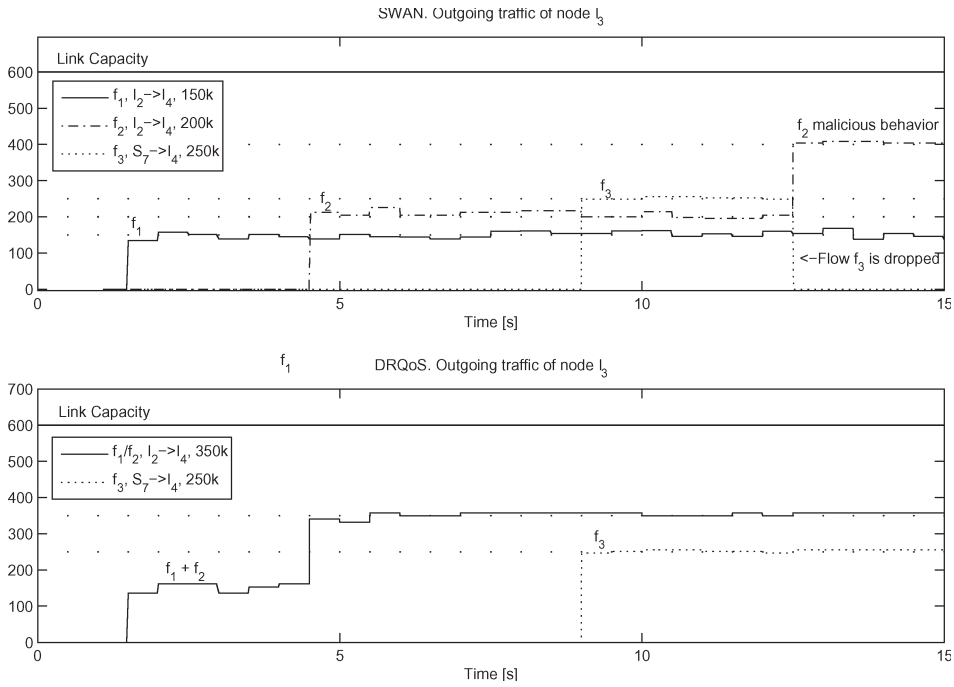

Fig. 7. Flooding attack that causes network congestion.

real-time flow and thus is demoted to a best effort traffic flow with lower priority.

Under DRQoS, however, node I_2 polices the malicious flow f_2 at its negotiated rate, such that congestion does not occur, because the $I_3 - I_4$ link has a capacity of 600 kbps. Thus, none of the traffic flows traversing node I_3 is affected by the flooding attack of flow f_2. This is indicated in the bottom graph of Fig. 7, which shows the traffic on the $I_3 - I_4$ link. Two state table entries are maintained at node I_3, one for the stream (I_2, I_4), which corresponds to the sum of flows f_1 and f_2, and one for the stream (S_7, I_4), which corresponds to flow f_3. Observe that the aggregate rate of stream (S_2, I_4) is limited to 350 kbps, whereas the rate of stream (S_7, I_4) is maintained at 250 kbps. Unlike the case of SWAN, flow f_3 is not affected by the malicious behavior of flow f_2.

B. Overreservation Attacks

To demonstrate DRQoS's ability to address overreservation attacks, we simulate the network topology shown in Fig. 8. In this scenario, three source nodes, S_1, S_2, and S_3, intend to establish traffic flows at the negotiated rates 150, 500, and 300 kbps, respectively. All the sources wish to send their traffic to the D_4 destination node. As in the previous network topology, each wireless link connects two nodes that are 200 m away from each other. The link capacities are all assumed to be 600 kbps. Source S_2 begins transmitting at time $t = 1$ s. Node S_2 negotiates a rate of 500 kbps of network bandwidth and initially acts legitimately by sending at the negotiated rate. However, at $t = 4$ s, node S_2 lowers its transmission rate to 100 kbps, thus performing an overreservation attack. Now suppose that node S_1 initiates a request to establish flow f_1 at time $t = 8$ s, and node S_3 issues a request to establish flow f_3 at time $t = 12.5$ s. If node I_3 does not implement the rate

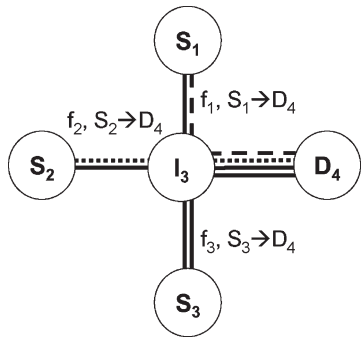

Fig. 8. Topology for overreservation attack.

adjustment mechanism of DRQoS, both flows f_1 and f_3 would be rejected, as indicated in the top graph of Fig. 9.

DRQoS is able to counteract the overreservation attack via rate monitoring and rate adjustment. In the above scenario, when node S_2 lowers its transmission rate to 100 kbps, the rate monitoring mechanism at node I_3 detects the change and lowers the reserved rate for stream (S_2, D_4) to the actual transmission rate of 100 kbps. Thus, the bandwidth that was overreserved by node S_1 is made available for other nodes. As a result, the bandwidth requests of sources S_1 and S_3 are accepted under DRQoS, as shown in the bottom graph of Fig. 9.

VI. Conclusion

We have proposed DRQoS: a QoS signaling protocol for MANETs that is resistant to a class of DoS attacks. The DRQoS protocol employs rate monitoring and traffic policing at the MAC layer to support QoS signaling on top of an arbitrary *ad hoc* routing protocol. The DRQoS protocol requires each node to maintain state information for each aggregate in-out traffic stream traversing an input–output pair, as opposed to

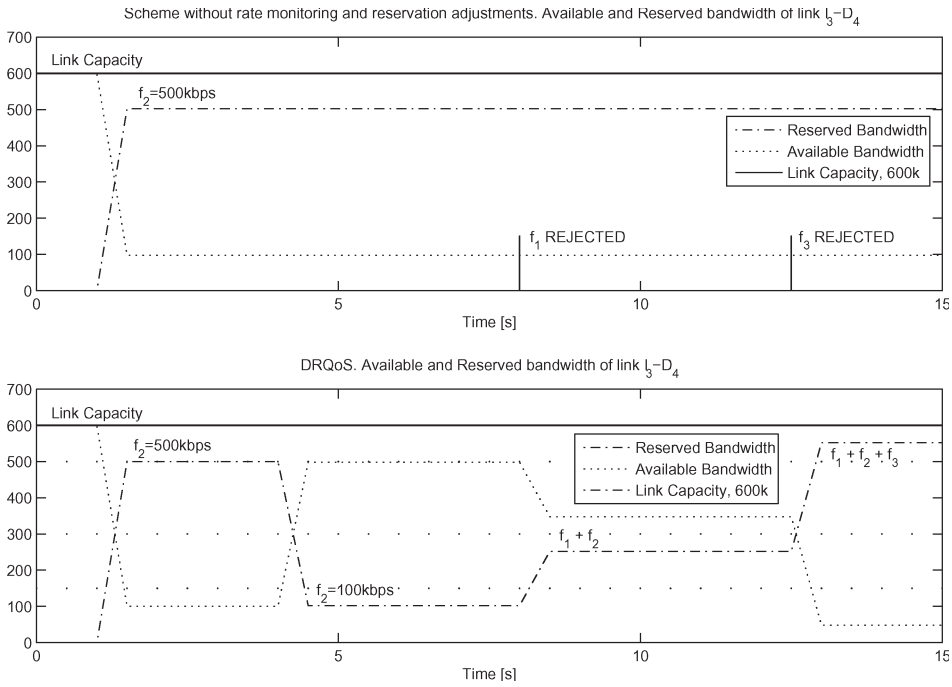

Fig. 9. Overreservation attack.

every flow, thus making the scheme more scalable. Each node performs traffic policing and rate monitoring/adjustment functions on each in-out stream to prevent DoS conditions. The protocol provides resistance to flooding, overreservation, and state table exhaustion while providing QoS to real-time traffic and service differentiation between real-time and best effort traffic. Simulation results from the ns-2 implementation of the proposed DRQoS protocol confirm the ability of DRQoS to provide resistance against flooding and overreservation attacks. The DRQoS protocol could be an important component in an overall architecture to provide security and QoS in MANETs.

The restriction to aggregate in-out streams makes DRQoS scalable and resistant to state table exhaustion attacks. However, if the number of flows traversing a DRQoS node is small, it may be advantageous for the node to perform per-flow traffic management to provide more fine-grained security and QoS. As the number of flows increases, flow state could be aggregated dynamically to conserve memory. More generally, a set of flow aggregates could be managed by a DRQoS node to provide different granularities of security and QoS in a dynamic fashion as memory and computational resources allow.

ACKNOWLEDGMENT

The authors would like to thank Prof. K. Gaj for helpful discussions and V. Papadimitriou for assistance in developing the simulation code for DRQoS.

REFERENCES

[1] S.-B. Lee, G.-S. Ahn, X. Zhang, and A. Campbell, "INSIGNIA: An IP based quality of service framework for mobile *ad hoc* networks," *J. Parallel Distrib. Comput.*, vol. 60, no. 4, pp. 374–406, Apr. 2000.

[2] G.-S. Ahn, A. T. Campbell, A. Veres, and L. H. Sun, "Supporting service differentiation for real-time and best-effort traffic in stateless wireless *ad hoc* networks (SWAN)," *IEEE Trans. Mobile Comput.*, vol. 1, no. 3, pp. 192–207, Jul.–Sep. 2002.

[3] M. Hejmo, B. L. Mark, C. Zouridaki, and R. K. Thomas, "Denial-of-service resistant QoS signaling for mobile *ad hoc* networks," in *Proc. ACM Workshop SASN*, Washington, DC, Oct. 2004, pp. 23–28.

[4] Y. Ge, T. Kunz, and L. Lamont, "Quality-of-service routing in *ad hoc* networks using OLSR," in *Proc. Hawaii Int. Conf. Syst. Sci.*, 2003, pp. 300–308.

[5] Q. Xue and A. Ganz, "*Ad hoc* QoS on-demand routing (AQOR) in mobile *ad hoc* networks," *J. Parallel Distrib. Comput.*, vol. 63, no. 2, pp. 154–165, Feb. 2003.

[6] L. Chen and W. B. Heinzelman, "QoS-aware routing based on bandwidth estimation for mobile *ad hoc* networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 3, pp. 561–572, Mar. 2005.

[7] P. Sinha, R. Sivakumar, and V. Bharghavan, "CEDAR: A core-extraction distributed *ad hoc* routing algorithm," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 8, pp. 1454–1465, Aug. 1999.

[8] S. Chen and K. Nahrstedt, "Distributed quality-of-service routing in *ad hoc* networks," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 8, pp. 1488–1505, Aug. 1999.

[9] C. R. Lin, "On-demand QoS routing in multihop mobile networks," in *Proc. IEEE INFOCOM*, 2001, pp. 1735–1744.

[10] C. Zhu and M. Corson, "On-demand QoS routing in multihop mobile networks," in *Proc. IEEE INFOCOM*, 2002, pp. 958–967.

[11] C. Zouridaki, M. Hejmo, B. L. Mark, R. K. Thomas, and K. Gaj, "Analysis of attacks and defense mechanisms for QoS signaling protocols in MANETs," in *Proc. WIS Workshop*, Miami, FL, May 2005, pp. 61–70.

[12] J. Mirkovic and P. Reiher, "A taxonomy of DDoS attacks and defense mechanisms," *ACM Comput. Commun. Rev.*, vol. 34, no. 2, pp. 39–53, Apr. 2004.

[13] Y. Xiao, "IEEE 802.11e: QoS provisioning at the MAC layer," *IEEE Wireless Commun.*, vol. 11, no. 3, pp. 72–79, Jun. 2004.

[14] T. You, C. Yeh, and H. Hassanein, "DRCE: A high throughput QoS MAC protocol for wireless *ad hoc* networks," in *Proc. IEEE ISCC*, 2005, pp. 671–676.

[15] H. Wu, X. Wang, Y. Liu, Q. Zhang, and Z.-L. Zhang, "SoftMAC: Layer 2.5 MAC for VoIP support in multi-hop wireless networks," in *Proc. IEEE Conf. SECON*, Santa Clara, CA, Sep. 2005, pp. 441–451.

[16] J. S. Turner, "New directions in communications (or which way to the information age)," *IEEE Commun. Mag.*, vol. 24, no. 10, pp. 8–15, Oct. 1986.

[17] M. Kazantzidis and M. Gerla, "End-to-end versus explicit feedback measurement in 802.11 networks," in *Proc. IEEE ISCC*, Jul. 2002, pp. 429–434.

[18] S. H. Shah, K. Chen, and K. Nahrstedt, "Dynamic bandwidth management for single-hop *ad hoc* wireless networks," in *Proc. IEEE Int. Conf. PerCom*, Dallas, TX, Mar. 2003, pp. 195–203.
[19] K. Fall and K. Varadhan. (2005, May). *The ns Manual*. [Online]. Available: http://www.isi.edu/ nsnam/ns/ns-documentation.html

Marek Hejmo received the B.S. degree in electrical engineering and the M.S. degree in computer engineering from AGH University of Science and Technology, Cracow, Poland, in 1999 and 2000, respectively. He is currently working toward the Ph.D. degree in information technology at George Mason University, Fairfax, VA.

His research involves security and quality-of-service aspects of mobile *ad hoc* networks. Other research interests include mobile and wireless communication, *ad hoc* networking, performance analysis, and analytical modeling.

Brian L. Mark (S'91–M'95) received the B.A.Sc. degree in computer engineering with an option in mathematics from the University of Waterloo, Waterloo, ON, Canada, in 1991 and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 1995.

He was a Research Staff Member at the C&C Research Laboratories, NEC USA, Princeton, from 1995 to 1999. In 1999, he was on part-time leave from NEC as a Visiting Researcher at Ecole Nationale Supérieure des Télécommunications, Paris, France. In 2000, he joined the Department of Electrical and Computer Engineering, George Mason University, Fairfax, VA, where he is currently an Associate Professor. His main research interests are in the design, modeling, and analysis of communication systems, communication networks, and computer systems.

Prof. Mark was a corecipient of the Best Conference Paper Award for the IEEE Infocom'97. He was also a recipient of the National Science Foundation CAREER Award in 2002.

Charikleia Zouridaki (S'01) received the B.S. degree in physics from Aristotle University of Thessalonica, Thessalonica, Greece, in 2000 and the M.S. degree in computer engineering from George Mason University, Fairfax, VA, in 2002. She is currently working toward the Ph.D. degree in information technology at George Mason University.

Her research interests include network security, systems security, and communication networks. Her research focuses on security of wireless networks.

Ms. Zouridaki is a student member of the IEEE Women in Engineering. She is also a member of Phi Beta Delta: an honor society for international scholars.

Roshan K. Thomas received the B.Sc. degree from the University of Lagos, Lagos, Nigeria, in 1985, the M.S. degree from the University of Houston, Houston, TX in 1987, both in computer science, and the Ph.D. degree in information technology with a specialization in computer security from George Mason University, Fairfax, VA, in May 1994.

He is currently a Senior Principal Scientist at SPARTA, Inc., Centreville, VA, and prior to that worked as a Senior Scientist at McAfee Research Laboratories. He has over ten years of experience as a Researcher at the Principal Investigator level in various aspects of computer security including access control models, network security, secure distributed database management, and multilevel-secure object-oriented distributed computing. He is currently a co-PI on a National Science Foundation sponsored project called SEQUOIA, which is investigating the integration of security-aware quality-of-service mechanisms into *ad hoc* wireless routing protocols.

Dr. Thomas served as the Cofounder of the First IEEE International Workshop on Pervasive Computing and Communication Security (PerSec 2004) and served as the PC Cochair for the second workshop (PerSec 2005).