Congestion Avoidance Routing Protocol for QoS-Aware MANETs

Consolee Mbarushimana and Ali Shahrabi
School of Engineering and Computing
Glasgow Caledonian University
Glasgow G4 0BA, U.K.
{consolee.mbarushimana, a.shahrabi}@gcal.ac.uk

Abstract—With the increase of multimedia traffic over the past few years and traffic differentiation introduced by IEEE 802.11e, nodes with delay-sensitive multimedia traffic tend to be busy for long periods, thus exacerbating the congestion problem in Mobile Ad Hoc Networks (MANETs). Although most of the existing routing protocols are based on the shortest path algorithm, some other metrics like load and delay have also been considered in some other research. In this paper, we first expose that the performance of MANETs routing protocols is highly dependent on the type of traffic generated or routed by intermediate nodes. This paper proposes a Type of Service Aware routing protocol (TSA), an enhancement to AODV, which uses both the ToS and traditional hop count as route selection metrics. TSA avoids congestion by distributing the load over a potentially greater area and therefore improving spatial reuse. Our simulation study reveals that TSA considerably improves the throughput and packet delay of both low and high priority traffic under different network operational conditions.

Index Terms--IEEE 802.11e, On-demand, Routing, ToS.

I. INTRODUCTION

MANETs routing protocols have received much attention in the literature because of the complex characteristics of such networks. The main classifications are reactive (AODV, DSR) and proactive (DSDV, OLSR) protocols. In several studies carried out to evaluate their performance [2], [3] the negative points of reactive protocols were found out to be high delay and packet loss due to stale routes, whereas the performance of proactive protocols is very much affected by the high routing load they produce. Shortest path algorithm used in the existing routing protocols does not provide optimal results when the primary route is congested. Congestion aware routing protocols [4], [6], [8], [12], [14] were therefore proposed to that effect due to the fact that besides route failures, network congestion is the other important cause of packet loss in MANETs.

The amount of multimedia traffic is rapidly increasing in today's networks. These types of traffic are delivered in form of UDP flows without the capability of managing congestion. They, as a result, aggressively use up more bandwidth than TCP flows. Additionally, multimedia traffic tends to continuously utilize network resources much longer than TCP traffic. For example, the PSTN is sized for average call duration of 2 minutes, but VoIP connections usually last longer than this. While it is necessary to provide resources to delay-sensitive applications to maintain their performance at an acceptable level, it is also important to protect best effort connection-oriented traffic so that all users can get a reasonable QoS.

Although congestion aware routing has been proposed for MANETs over the last decade [4], [7], [9], the effect of the different traffic types has not been considered in any routing protocol. Distinguishing different types of traffic that each node deals with in QoS-aware MANETs and responding to routing protocol requests correspondingly can potentially increase the performance of routing protocols and the whole network. In this paper, we first investigate the interaction between the ToS and the routing protocols, we then propose a novel Type of Service Aware routing protocol (TSA) a cross-layer enhancement to AODV. TSA uses both ToS and traditional hop count as route selection metrics and avoids congestion by distributing the load over a potentially greater area and therefore improving spatial reuse.

The rest of the paper is organized as follows; Section 2 gives a brief review about related work, QoS-aware MANETs and AODV. In Section 3, we explain our proposed routing protocol in details. The performance evaluation and simulations results are presented in Section 4. Finally, some concluding remarks are given in Section 5.
II. BACKGROUND

A. Related Works

Several load aware routing algorithms have been proposed. Those include the dynamic load aware on-demand routing protocol (DLAR) proposed by Lee and Gerla in [6], and the Load-Balanced Ad Hoc Routing protocol (LBAR) proposed in [5]. Yuan et al. proposed another scheme in [13], which relies on intermediate nodes not to reply to route request messages when their load exceeds a certain threshold. A similar idea was also adopted by Lee and Riley in [7]. The routing scheme proposed in [1] keeps track of the size of traffic in bytes. Load aware routing in ad hoc (LARA) [9] uses a metric called traffic density to represent the degree of contention at the MAC layer.

MAC layer channel contention information, number of packets in the interface queue, and the traditional hop count are the three metric used for route selection by CSLAR (contention sensitive load aware routing protocol) [8]. A similar protocol, Contention and Queue-aware Routing (CQR) was proposed in [4]. A congestion adaptive routing (CRP) in which a route is adaptively changeable based on the congestion status of the network is proposed in [12]. The Delay-Oriented Shortest Path Routing protocol (DOSPR) [10] and Load-Aware On-demand Routing (LAOR) [11] are based on path delay. Other approaches using metrics indirectly related to the load have been proposed. Like the Load Balance Routing using Packet Success Rate proposed in [15], and Lifetime-aware Leisure Degree Adaptive Routing protocol (L-LDAR) [14].

B. QoS-Aware MANETs

There has been a tremendous increase in multimedia applications over the past few years. This type of applications requires QoS guarantees in terms of delay, bandwidth, packet loss and jitter. With the prospects of future MANETs commercial applications, it is desirable to support these services in MANETs as well.

In EDCA, four access categories (ACs) are defined by using different arbitration interframe space (AIFS), initial and maximum contention windows (CWmin and CWmax respectively). The AIFSi for a given ACi is given by

\[ AIFS_i = AIFSN_i \times \delta + SIFS \]  

where \( i \) is the AC number, \( \delta \) is the time interval of a slot and AIFSNi is constant dependent on the AC. With small values of AIFSN, high priority classes start decreasing their backoff counter earlier than low priority classes. High priority ACs are also given smaller values of \( CW_{\text{min}} \) and \( CW_{\text{max}} \), which results into shorter backoff intervals. The backoff interval (BI) is randomly chosen in the range \([0, CW_i]\), where \( CW_i = 2^{k-1}CW_{\text{min}} \). In real life, multimedia traffic like voice (AC3) and video (AC2) are assigned higher priority over best effort (AC1) TCP based applications (e-mail, FTP).

C. Ad Hoc On-demand Distance Vector Routing Protocol (AODV)

AODV minimizes the number of broadcasts by creating routes on-demand. Figure 1 illustrates a simple route discovery in AODV. The node S seeking a route to a destination D broadcasts a RREQ (route request) message to neighbour nodes. In the simple case scenario where nodes N1, N2, and N3 have a route to the destination, they reply with a RREP (route reply) message containing the number of hops (hop count hc) to the destination. AODV, being a distance vector protocol that uses the hop count as the metric, will choose the path through node N2 as it is the shortest.

III. ToS AWARE ROUTING

Congestion or adaptive routing has been investigated in several studies as we explained in section 2. The approaches in all the cited studies converge in evaluating or assessing the level of activity in intermediate nodes by measuring either the load or the delay. Based on the gathered information, the optimal path is established trying to avoid the already or likely to become congested nodes. However, none of the research reported has evaluated the effect the type of service of the traffic carried by intermediate nodes has on the performance of routing protocols.

In the route discovery process of most of MANETs routing protocols, the nodes with a route to the destination advertise themselves as candidate to route traffic to the destination, without considering the status of their queues. This might result into long delays or packet drops for newly arriving traffic, failing to be transmitted ahead of the already queuing traffic. This effect is very pronounced in mixed traffic QoS-aware MANETs, where different priority traffic conversations coexist. In such networks, low priority traffic has to
queue for long periods if it is routed through a node engaged in a high priority traffic conversation. This problem might also occur for new high priority connections, if their traffic encounters a node already busy with same priority traffic on its route, as it will be transmitted in a FIFO fashion. An important point to be taken into consideration is that in today's networks, delay-sensitive traffic conversations tend to last longer than best effort traffic. Downloading a webpage or e-mail lasts just a few seconds, while voice calls and video conferences can last several minutes.

This paper suggests a simple yet effective approach to alleviate this problem and at the same time achieving load balancing and spatial reuse. We propose a modification to AODV, which we call Type of Service Aware (TSA) routing protocol. TSA is a cross-layer approach that uses the ToS of the traffic in the nodes queues, and couples it with the traditional hop count in the route selection process. The fundamental nature of TSA is in the route discovery process. A source node broadcasts a RREQ like in AODV. On receiving the RREQ, the intermediate node with a route to the destination therefore wishing to generate a RREP, first checks its availability by considering the AC of the traffic in its queues. Using the same colour scheme as in [12], the node is classified in one of the four categories; green, yellow, orange or red, depending on whether the traffic belongs to background, best effort, video or voice AC respectively. The different congestion levels are shown in Table 1. Using this classification, a node with no traffic or with delay-insensitive traffic is considered more flexible to receive more traffic than a node with delay-sensitive traffic, which is likely to be busy for a prolonged time.

After determining its availability, the intermediate node generates a RREP packet. In this approach, we propose to modify the information contained in the RREP message so that it reflects the availability status of the node. We propose to add to the actual number of hops to the destination, an additional number proportional to the availability level. The resulting hop count \(hc\) in the RREP is therefore given by

\[
hc = \text{actual} \_ hc + i \times n
\]

where \(i\) is the congestion level and \(n\) is a constant proportional to the average hop count of the network. On receiving the RREPs, the destination will choose the route with the smallest number of hops. In case the node buffers traffic of different classes, the class with highest congestion level is considered. The congestion level of every node is updated every time there is change in traffic type, and the congestion level is periodically propagated to neighbours in the HELLO messages.

Using the example network of Figure 1, consider that the intermediate nodes are labelled green to red as shown on Figure 2 and as it is a small network \(n=1\). The intermediate nodes will therefore reply with RREPs with the modified hop count values shown in the figure. Applying the shortest path algorithm, TSA will therefore choose to use the next hop as node \(N_1\), unlike AODV that would have chosen the heavily congested node \(N_2\).

TSA is a cross layer solution that works in conjunction with the MAC layer. The MAC layer is responsible for updating the IP layer whenever there is a change in the traffic ToS. In our implementation, the MAC layer generates an interrupt when there is a change in its availability level. This interrupt is directly fed to the network layer, which in turn interprets it and adjusts the hop count to be included in the RREPs.

### IV. PERFORMANCE EVALUATION

#### A. Simulation Setup and Parameters

Our simulations were conducted using OPNET Modeller 11.5. We simulated a network consisting of 50 mobile nodes moving in a 1000x1000 m area and a nominal transmission range of 250m. The MAC layer protocol used is EDCA with the default settings. The network traffic consisted of long-lived FTP file transfers.
and the delay-sensitive traffic was simulated by establishing G711 CBR voice connections. As the protocol (TSA) proposed in this paper is a congestion avoidance routing protocol, we evaluated its performance compared to AODV’s under different load levels. The two protocols were also evaluated under different mobility levels. Their performance is assessed by the throughput, the packet end-to-end delay and the routing load.

B. Simulation Results

1) Effect of Voice Connection Time

As we mentioned earlier, one of the reason behind our proposal is that multimedia traffic conversations tend to last longer, which makes the nodes unavailable to service other connections. We therefore evaluate the effect of voice connection time on the performance of the two routing protocols. This time was varied as a percentage of the total simulation time.

TCP traffic opportunities to be transmitted when voice transfer is taking place are reduced; the longer the voice connections last, the poor the TCP throughput becomes. As seen on Figure 3(a), long-lived voice connections have a very negative impact on TCP, whose throughput drops below 25% when there is constant voice traffic transfer. Using TSA as the routing protocol, the route selected tries to bypass those nodes with voice traffic. TSA is therefore able to achieve a remarkable improvement of 20% in TCP throughput in case of long lasting voice connections. As for the voice traffic, TSA is also able to deliver a large amount compared to AODV (Figure 3(b)). New connections voice traffic is routed through nodes with no traffic or with best effort traffic; therefore, it is able to get through easily and the number of dropped packets is reduced.

TSA is able to constantly deliver both TCP and voice traffic with delays smaller than AODV’s. AODV and TSA TCP segment delays are comparable for short time connections, but for longer connections, the improvement in TSA segment delay becomes remarkable as seen on Figure 4(a), as busy nodes are avoided. For any voice packets that bypass red nodes, they are automatically transferred ahead any existing best effort traffic as they have higher priority, hence the lower voice delay achievable by TSA (Figure 4(b)).

This paper also evaluates the normalized routing load, which is the ratio between routing traffic generated to the successfully received data traffic. We observe a sharp increase in the routing load normalized on TCP when the voice connections last longer. As TSA is able to deliver more TCP traffic than AODV, its routing load is smaller as shown on Figure 5(a). Voice throughput increases with increase in connection time, hence the decrease in the routing load normalized on voice as seen on Figure 5(b).

2) Effect of Number of Traffic Sources

The number of traffic sources has a significant impact on the performance of on-demand routing protocols. The increase in number of VoIP nodes negatively affects TCP performance for the two routing protocols as seen on Figure 6(a), because more transmission opportunities are given to the nodes with delay-sensitive voice traffic. The voice throughput on the other hand (Figure 6(b)) is increased when the number of voice connections increases. However, the increase is not linear due to self-contention between voice connections themselves. Similarly, the increase in number of TCP
connections makes the contention harder and hence affecting the network throughput. TCP throughput increases with the number of TCP sources (Figure 7(a)), but again due to the contention between the TCP connections, the increase is not a linear function of the connections number. This increase in TCP traffic has a negative impact on voice throughput as seen on Figure 7(b). With TSA, new connections use only the least congested nodes and therefore the load is distributed across the network. TSA constantly achieves higher throughputs than AODV. The packet delay increases when both voice and TCP connections are increased due to increased medium contention, but is again smaller in TSA networks as seen on Figure 8 and 9.

3) Effect of Nodes Mobility

In MANETs, nodes mobility plays a high role in determining the performance of routing protocols. The nodes mobility results into frequent topology changes and route failures; the performance of the routing protocol depends on how quickly it can recover from the failures. In congested networks, it will take longer for the routing protocol to repair the broken route or establish a new route to the destination.

Different mobility models were obtained by varying the pause time from 0 to 100 seconds. The throughput is shown on Figure 10. In low mobility scenarios, there are fewer route failures, fewer packets are lost, hence the increase in the network throughput. The two protocols behave similarly, the difference in favour of TSA being due to its use of congestion avoidance. Similarly, the packet end-to-end delay is reduced for the two protocols in low mobility (Figure 11), as less time is spent in discovering and repairing routes.

V. CONCLUSION

During the route discovery process of routing protocols in MANETs, nodes advertise themselves as capable of reaching the destination irrespective of the status of the traffic already in their queues. The new arriving traffic might face long waits or get dropped failing to get transmitted ahead existing high priority traffic. This problem was exposed in this paper. As the main concern in QoS-aware MANETs is to achieve service guarantees for delay-sensitive applications; such incidents are therefore common and almost inevitable.

In this paper, we therefore propose a new Type of Service Aware (TSA) routing protocol, which avoids such nodes in the route discovery process. TSA is a cross-layer congestion-avoidance routing protocol in which the routes through nodes engaged with delay-sensitive traffic for extended periods are only selected as the last resort, even when they are shorter. Avoiding busy nodes alleviates congestion, results in less packets drop and in a short end-to-end delay. In addition, TSA distributes the load on a large area, thus increasing the spatial reuse. Our intensive simulation study has
confirmed the advantages of TSA over AODV. We intend to explore the application of the idea introduced in TSA to other routing protocols in our future study.

REFERENCES