A hybrid gateway discovery algorithm for supporting QoS communications in heterogeneous networks

ABSTRACT: In most practical applications, ad hoc mobile devices demand the access to wired hosts located beyond the limits of the mobile ad hoc network, forming a heterogeneous wired-cum-wireless network. This interconnection requires a network device known as gateway. The gateway functions as a bridge between the ad hoc environment and infrastructure-based networks, as Internet. We propose in this paper an algorithm for the discovery and selection of gateways and its integration in a routing protocol for mobile ad hoc networks. Our proposed algorithm is based on an adaptive hybrid strategy where the proactive and reactive methods are combined. Consequently, gateways maintain routing information sending advertisement messages to mobile nodes inside a limited range (proactive area). Those nodes located outside that area must execute a reactive gateway discovery process. The size of the proactive area and the frequency of the control messages are dynamically adapted. The simulation results show that our approach achieves better results in terms of packet loss and delay without increasing the routing overhead.

ARTICLE INFO
Received March 12, 2015
Accepted July 10, 2015

KEYWORDS
Mobile ad hoc networks, heterogeneous networks, gateway discovery, routing protocol, routing protocol evaluation

RESUMEN: En la mayoría de las aplicaciones, los dispositivos móviles ad hoc necesitan tener acceso a dispositivos de red ubicados más allá de los límites de la red móvil ad hoc. Tal como puede ser el caso de servidores ubicados en Internet. Esta interconexión entre redes cableadas e inalámbricas (redes heterogéneas) requiere de un dispositivo de red conocido como Gateway. La principal función de un Gateway es servir de interfaz entre las redes ad hoc y las redes cableadas (por ejemplo, Internet). En este artículo se propone un algoritmo para la búsqueda de Gateways dentro de la red y el establecimiento de rutas hacia estos. El algoritmo desarrollado sigue una estrategia híbrida y adaptativa, es decir, se combinan los métodos de enrutamiento proactivos y reactivos. En consecuencia, los Gateways mantienen la información de las rutas por medio del envío periódico de mensajes de control a los nodos móviles dentro de un rango limitado (área proactiva). Aquellos nodos situados por fuera de dicha zona, deben ejecutar un proceso de búsqueda reactiva de Gateways. El tamaño de la zona proactiva y la frecuencia de los mensajes de control se adaptan dinámicamente. Para su aplicación práctica, el algoritmo desarrollado ha sido integrado dentro un protocolo de enrutamiento para redes móviles ad hoc. Los resultados de las simulaciones muestran que la solución planteada logra mejores resultados en términos del porcentaje de paquetes perdidos y del retardo sin aumentar la sobrecarga debida a los paquetes de control.

1. Introduction

A Mobile Ad-doc Network (MANET) is a type of wireless network with autonomous mobile users that are moving arbitrarily and communicate over multi-hop relays without any centralised administration nor wired infrastructure. MANETs are generally conceived to operate as stand-alone networks, which means that data traffic will be restricted within MANETs. All nodes in MANETs equally participate in the exchange of routing information, by running the same routing protocol. Several routing protocols for MANETs have been proposed, such as AODV (Ad hoc On-Demand Distance Vector protocol) [1], DYMO (Dynamic MANET On-demand Routing protocol) [2] and OLSR (Optimized Link State Routing Protocol) [3]. However, these protocols

* Corresponding author: Wilder Eduardo Castellanos Hernández
e-mail: wc Castellanos@usb bog.edu.co
ISSN 0120-6230
e-ISSN 2422-2844

DOI: 10.17533/udea.redin.n78a11
were originally designed for supporting communications within an autonomous MANET where no intervention of any centralised router exists. Therefore, a mobile ad hoc node cannot obtain the routing information beyond the scope of the MANET. To support communications between mobile ad hoc nodes and host devices in a wired network (e.g. Internet) routing protocols must be modified. The interoperability in heterogeneous networks, i.e., between ad hoc networks and other kinds of networks, such as cellular networks or wired networks, has an increasing attention. Several approaches have been proposed to achieve this interconnection [4, 5]. Basically, routes between wireless ad hoc networks and infrastructure-based networks are established through network devices (named, Gateways) that understand not only the IP stack protocol, but also the MANET protocol suite (see Figure 1). Therefore, a gateway is an interface between MANET and the infrastructure-based networks, such as Internet.

When a mobile node in an ad hoc network wants to establish a communication with a wired node, it has to discover the available gateways. The gateway discovery process can be executed following a proactive or reactive method. In a proactive scheme, gateways periodically broadcast an advertisement message throughout ad hoc network. When a node receives the advertisement message, it forwards it to other nodes until this message is flooded in the entire network. This scheme provides good connectivity and lower delay. However, high routing message overhead is incurred since gateway advertisements are disseminated even if there is no communication demand from nodes in MANET. On the other hand, in reactive schemes a mobile node broadcasts a gateway discovery message to discover gateways in ad hoc network. Once a gateway receives the discovery message, it sends a unicast advertisement message back to the requestor node. Although this scheme achieves the routing overhead, it may increase route discovery delay since mobile nodes have to send a gateway discovery message every time they need a gateway. Few solutions [6, 7] propose hybrid gateway discovery schemes where the dissemination of gateway advertisements is kept limited to a small range of n-hops (proactive area). Nodes outside this range use a reactive scheme to find their gateways.

Three key elements must be considered in the hybrid gateway discovery scheme. The first element is the selection of best gateway when the requesting node receives multiple advertisement messages from gateways. Most of proposed schemes for gateway selection only use the hop count metric to select a gateway [8, 9]. Therefore, in these schemes, nodes always select the nearest gateway. However, although a route between a mobile node and a certain gateway has fewer hops than other routes, it may have a higher traffic load. The second element to consider is related to the determination of the optimal proactive area. A large proactive area involves an increase in the routing overhead since routes must be maintained in a wider area. On the contrary, if the proactive area is small there will be less maintenance overhead, but more delay during the gateway discovery is experienced. The third issue is the adjustment of the frequency at which advertisement messages are disseminated within the proactive area. In traditional hybrid gateway discovery scheme, the size of the proactive area and the frequency of the advertisement messages are set up statically and these are never updated. This leads to a rigid implementation of the hybrid scheme. On the other hand, as a consequence of the massive demand for multimedia content from mobile devices, the provision of Quality of Service (QoS) has become an important research topic. In order to efficiently support multimedia traffic in heterogeneous networks, in previous works we have proposed the routing protocol AQA-AODV (Adaptive QoS-Aware Ad hoc On-demand Distance Vector) for mobile ad hoc networks. AQA-AODV provides mechanisms in order to provide certain level of quality of service during the discovery and maintenance of routes. However, this protocol is only designed to operate in a purely ad hoc environment. Therefore, it is not suitable for supporting interconnection between MANETs and wired networks, such as Internet.

For this reason, we propose in this paper a new hybrid gateway discovery algorithm, integrated in the AQA-AODV protocol, for supporting the connectivity between MANETs
and infrastructure-based networks. Our gateway discovery algorithm includes available bandwidth as a metric during the gateway selection routine, which is a more suitable solution for providing time-sensitive communications, such as multimedia services. Moreover, we solve the issues of the conventional hybrid scheme, adapting dynamically the size of the proactive area and the frequency of the advertisement messages.

We conducted a performance evaluation of our hybrid gateway discovery using the network simulator NS-2 [10]. The experimental results show that the combination of the adaptive mechanisms of our hybrid scheme and the bandwidth estimation algorithm of AQA-AODV, provide an efficient solution for establishing stable routes to hosts in wired networks. In addition, results reveal that routing overhead and network congestion are reduced, avoiding unnecessary dissemination of control packets and allowing traffic source to adapt its data rate according to the network conditions.

The rest of the paper is organized as follows. Section 2 describes related works about gateway discovery strategies. In Section 3 we briefly review the main characteristics of AQA-AODV protocol. The proposed hybrid gateway discovery scheme is presented in Section 4. Section 5 shows our simulation results and finally, we present our conclusions in Section 6.

2. Related works

During the last years, several approaches have been proposed for interconnecting wired and wireless networks. Most of the proposed solutions are based on the IETF Draft “Global Connectivity for IPv6 Mobile Ad-hoc Networks” [11]. This work defines two different schemes for gateway discovery: periodic dissemination of gateway advertisement (GWADV) messages from gateways and reactive broadcast of gateway solicitation (GWSOL) message from nodes. These reactive and proactive schemes are not dependent on any routing solution. However, this approach does not give any metric to select a gateway. Few solutions [8, 9] propose a gateway selection method based on the number of physical hops to gateway. Nevertheless, this metric is not always optimal in the long term. The gateway selection proposed in [6] is based on three parameters: the interface queue size, the total number of neighbours of each node along the route and the hop count. They define a hybrid discovery method where gateways broadcast advertisement messages when they receive a gateway discovery message from a mobile node. The TTL (Time To Live) of the gateway advertisement message is set to a value equal to the distance of the gateway from the requesting node. Another hybrid approach is presented in [12]. They extended the AODV routing protocol with a hybrid scheme for gateway discovery that exploits the opportunistic routing. With this technique, while a mobile node is in transit between two different gateways, it can temporarily store messages before retransmitting. In reference [7], authors apply biologically inspired metaphor to design a hybrid gateway discovery.

Moreover, they propose a distributed gateway discovery based on the prediction model of mobility of nodes, which may require high processing power consumption and nodes have to wait longer period of time to select a path to a gateway. The work presented in [13] proposes a novel technique by which the messages generated by gateways are exclusively forwarded in those areas where links are expected to remain stable.

However, the solutions discussed above do not perform a dynamic adaptation neither to the size of the proactive area nor to the frequency of the advertisement messages. In [4, 14] have studied the problem of the adjustment of the proactive range and the advertisement frequency, respectively. Their findings suggest a dynamic adaptation of these parameters according to networks conditions, such as the mobility of nodes, the position and number of mobile nodes. Several adaptive strategies have been designed. For instance, in [15] proposed a fuzzy logic system to control the frequency of the advertisement message and an adaptation of the TTL value based on the maximal source coverage. In reference [16], authors propose a QoS-based gateway selection mechanism based on three QoS metrics: traffic load of gateway, the path quality from mobile node to the gateway and the hop count to the gateway. Other gateway discovery strategies for heterogeneous networks can be consulted in the surveys [5, 17].

3. Background

A new routing protocol (named AQA-AODV) has been developed as a modification of the well-known AODV protocol. The main modifications affect the route discovery mechanism and the route maintenance strategy of AODV. Specifically, we have implemented an algorithm used for the estimation of the available bandwidth that allows nodes along the path to know their available resources (in terms of bandwidth). In addition, we added a cross-layer mechanism to send information about the available bandwidth to the application layer; hence, source node can easily adapt its transmission rate. In order to implement the above modifications, we have introduced some changes in the format of the packets used during the route discovery procedure of AODV [RREQ, Route Request and RREP, Route Reply]. In particular, we have added new fields to carry information about the bandwidth requirements, transmission rate and a session ID [used to identify each QoS flow]. The new AQA-AODV packets are called QRREQ and QRREP to emphasize the new feature of exchanging QoS parameters.

The most important difference between our proposed protocol and other solutions based on AODV is the adaptive feedback scheme, integrated into the routing protocol, by which the source node can know the available bandwidth and easily adapt its transmission rate according to the state of the route. For this reason, nodes along the path must know their available resources by using some algorithms.

Figure 2 depicts the functional block diagram of AQA-AODV.
The main three elements of AQA-AODV are a bandwidth estimation module, a routing module and a route recovery module. The first module carries out the estimation of the available bandwidth and provides data feedback to the application layer. The bandwidth estimation algorithm uses the HELLO packets, which are employed in AODV to discover neighbours. The traffic source uses the information about the available bandwidth to adjust its bit rate according to the network conditions. On the other hand, the routing module receives the route requests from the application and executes the route discovery routine. The third main module is the route recovery module, which is in charge of re-establishing the connections after a link failure, taking into account the QoS conditions of each of the sessions.

3.1. Route Discovery in AQA-AODV

When a source node requires a route to a destination node with specific bandwidth requirements, it broadcasts a RREQ packet with the QoS extension (QRREQ) to its neighbour nodes. This packet includes two new fields additionally to the standard fields of AODV: reqBW and Session ID. The field reqBW corresponds to the bandwidth requested by the source node and the Session ID field is used, together with the source address, to uniquely identify each traffic flow. Once a node receives a QRREQ packet, a reverse route entry is created with the Session ID, and the QRREQ packet is rebroadcasts. This process continues until the QRREQ packet reaches the destination node. In AODV, a RREP packet can be created by the destination node or an intermediate node with a “fresh enough” route to the destination. However, only the destination will be able to send the route reply packet (QRREP) in AQA-AODV. This will ensure that all nodes in the selected route satisfy the bandwidth constraints. When the destination node receives a QRREP packet, if it is a new request, a reverse route entry for the new session will be created. Before sending the QRREP to the source, local available bandwidth is checked. Finally, the QRREP will be transmitted to the source with a modified header that includes the field min-bandwidth, which stored the minimum value between the bandwidth requested by the source (reqBW) and the local available bandwidth of the destination node. Once an intermediate node receives the QRREP packet, it compares its local available bandwidth with the bandwidth indicated in the QRREP. If its local available bandwidth is lower, it replaces the value stored in the min-bandwidth field of QRREP, with the value of its local available bandwidth. Otherwise, the node forwards the QRREP. This procedure will ensure that source node knows the minimum bandwidth along the path, which will be the maximum rate that it may transmit. Once source node receives the QRREP packet, it adjusts its transmission rate according to the value of the field min-bandwidth in QRREP and then the transmission of the data packets is started. A more detailed information of AQA-AODV can be consulted in the references [18, 19].

4. Proposed algorithm for gateway discovery in AQA-AODV

We describe in this section the gateway selection mechanism implemented as an extension of the AQA-AODV routing protocol to improve interconnection performance between MANET and the Internet. We have adapted the ideas suggested in the IETF Draft [11] for implementing a hybrid gateway discovery in order to minimize the disadvantages of the proactive and reactive strategies. In contrast with the proactive method, overhead and network congestion are reduced since the GWADV messages of the hybrid approach are propagated only in a limited number of hops away from the gateway (advertisement zone) and not along the entire network. On the other hand, hybrid method shows a better performance than the reactive approach, in terms of latency.

In our hybrid discovery strategy, mobile nodes within a limited range (a certain number of hops away from the gateway) perform a proactive discovery while mobile nodes located outside this range use reactive gateway discovery (see Figure 3).
4.1. Gateway discovery process

In our approach, gateways periodically broadcasts a GWADV (Gateway Advertisement) message with a certain TTL value. The TTL parameter defines the range in which proactive discovery method is used. The mobile nodes that receive the GWADV message update their routing table and then rebroadcast the message. When a mobile node in a MANET needs to establish a route to a fixed node (e.g., located in Internet), firstly it has to find a gateway. Therefore, mobile node looks for a default route (i.e., a route to a gateway) in its routing table. If mobile node finds a default route, it uses this route to transmit packets to the gateway. This fact indicates that mobile node is inside the proactive range. However, if the mobile node is not within the proactive area, it could not find a route to a gateway in its routing table. Consequently, it starts a gateway discovery process sending a Gateway Request message (GWREQ). If other mobile nodes receive this message, they rebroadcast it until the GWREQ message is received by a gateway (see Figure 4(a)).

Once a gateway receives a GWREQ message, it responds sending back a unicast reply message, named GWREP (Gateway Reply), to the requesting node. We propose to add one new field (named BW) in the GWREP message header in addition to the conventional fields. We use the BW field to store the available bandwidth of the nodes along the route from a gateway to the requesting node. In our approach, once an intermediate node between gateway and the requesting node receives a GWREP message, it compares its local available bandwidth with the bandwidth indicated in the GWREP. If its local available bandwidth is lower, it replaces the value stored in the BW field of the GWREP message, with the value of its available bandwidth. Otherwise, the node forwards the GWREP message [see Figure 4(b)]. Consequently, this procedure will ensure that the requesting node knows the available bandwidth of the route to each gateway.

GWREQ and GWREP messages were implemented in our routing protocol by simply adding an I flag to existing route request (QRREQ) and route reply (QRREP) packets in order to make the implementation easier.

It is important to note that, the TTL value as well as the sending frequency of the GWADV messages may have a high impact in the overhead and the overall network performance [4, 20]. Therefore, these parameters should be dynamically adapted depending on the networks conditions. With the aim of adapting the TTL of the GWADV message, we propose an adaptive approach based on the distribution of mobile nodes. In our approach, each gateway keeps a mobile node list to record the information of every active mobile node that has an active route to Internet established through it. The information includes mobile node IP address, numbers of hops to the gateway and lifetime. This information is easily obtained from the IP header of data packets. Gateway can periodically query the mobile node list in order to calculate the average distance (in number of hops) of the mobile nodes registered in that list. This average distance will be the TTL value of the next GWADV message to be broadcast by the gateway. The average number of hops
is an approximate indicator of how the mobile nodes are distributed around a gateway. If the average number of hops decreases means that only few mobile nodes are closed to gateway, then the TTL value should be decreased in order to reduce the overhead by avoiding the dissemination of unnecessary GWADV messages. Otherwise, the TTL value should be increased to cover more mobile nodes avoiding the propagation of excessive GWREQ messages.

In addition, the interval between two consecutive gateway advertisements (known as advertisement interval) should be carefully selected. A short advertisement interval allows mobile nodes to maintain an updated route to gateways. However, it can saturate the network with GWADV messages unnecessarily, increasing the traffic overhead both and the power consumption. On the other hand, a high advertisement interval could lead to the storage of out-of-date routing information in the mobile nodes. This fact implies that mobile nodes have to start a reactive process, sending a GWREQ message, when they need to forward packets to the gateway. Therefore, an autonomous and dynamic algorithm to adapt the value of the advertisement interval is strongly recommended. In our approach, we implemented the algorithm presented in [21], where the advertisement interval varies according to the network stability. If the mobile nodes are near the gateway, routes can be assumed as more stable and therefore, the advertisement interval may be increased. If the nodes remain just a short period of time under the coverage of the gateway, most routes probably will be broken in the short term and GWADV message should be sent more frequently (advertisement interval must be decreased). The minimum value that can be assigned to the advertisement interval is 3 seconds, in accordance with the recommendations given in the standard [22]. And the maximum value that advertisement interval can reach is 30 seconds according to the outcomes of the study [13]. Between these two limits, the value of the advertisement interval is set according to a linear function, which is proportional to the network stability estimation performed by the gateway. The estimation of the network stability factor is based on the variation of the number of nodes that are under the coverage of the gateway during two consecutive advertisement intervals. For more details about the estimation of the stability parameter, the reference [21] can be consulted.

4.2. Route establishment

Once a mobile node selects a route to a gateway, it can initiate a route discovery process towards the fixed node (i.e. Internet host) using the conventional route discovery packets of AQA-AODV. Specifically, mobile node sends a unicast QRREQ packet to a fixed node, informing its available bandwidth. This bandwidth value is the minimum value between the bandwidth informed by route during the gateway discovery process and its own estimation of the available bandwidth. When the fixed node receives the QRREQ packet, it configures its route to the gateway, adjusts its bit rate according to the available bandwidth and sends data packets to the requesting mobile node.

5. Performance evaluation

In this section, we present the evaluation of our proposed solution and we compare it with other similar approaches under the same network conditions. For this evaluation, we have conducted extensive simulations using the network simulator NS-2 [10]. Our approach (hereinafter cited as AQA-Hybrid) was compared with the well-known reactive approach (also known as AODV+) proposed in [8] and the recent strategy [QGWS, QoS-based Gateway Selection] developed in [16].

The metrics used for the performance analysis of the proposed algorithms were: i) the packet loss rate, which corresponds to the ratio between the number of lost packets and the total number of packets sent by the MANET nodes; ii) the average end-to-end delay, which is defined as the time elapsed since the emission of the data packet from the source until it reaches its final destination and; iii) the normalized routing overhead, that represents the ratio between the number of control packets and the number of received data packets.

5.1. Simulation environment

The simulated scenario consists of 2 gateways, 2 fixed routers, 2 wired hosts and 30 mobile nodes randomly distributed in an area of 1000x300 m. The mobile nodes move according to the random waypoint model with a variable speed from 2 to 10 m/s. The gateways are located at the x- and y-coordinates (100,100) and (900, 100). The wireless channel capacity for each mobile node is 2Mb/s, using the IEEE 802.11b DCF MAC layer and a transmission range of 250 m. The radio propagation model is Two Ray Ground and the queue type is Drop Tail with maximum length of 50. Two scenarios with different traffic load were tested. The first scenario has 10 traffic sources and the second one has 15 sources communicating with nodes in the wired network. All traffic flows are Constant Bit Rate (CBR) streams over UDP with a packet size of 512 bytes. The packets transmission is established between a random source-destination pair. The simulation time was configured to 300 seconds and each data point shown in figures is the average of 10 simulations with different random seed. In Table 1 some simulation parameters are described and a graphical representation of the simulated scenario is shown in Figure 5.
### 5.2. Results

The results presented in Figure 6 show the percentage of packet loss obtained for the three solutions. We can observe that the differences between our proposal and other approaches are very significant. Comparing Figure 6(a) and 6(b), it can be seen that with an increase in the number of sources, the performance of the AODV+ and QGWS approaches decreased. Whereas AQA-Hybrid maintained its percentage of lost packets below 11%, about 2 and 3 times lower than QGWS and AODV+, respectively. The high packet loss of the AODV+ and QGWS approaches could be caused by network congestion in the mobile nodes. This congestion is generated when intermediate nodes between a mobile node and a gateway, cannot efficiently retransmit the packets because they do not have enough available bandwidth. Connectivity losses could be an additional explanation for the high packet losses. However, because of the high density of nodes as well as the fact that there are two gateways in the network, it is more probable that there are always routes to gateways. Consequently, packet losses due to connectivity losses are occasional. The dropped packet rate increases as the speed increase because the link lifetimes decrease. Therefore, routes break easier and the transmission of the data packets through broken routes will increment the losses.

As expected, the AODV+ proposal presents a high delay, which is significantly reduced using the hybrid approaches. Moreover, the delay associated to AQA-Hybrid is evidently lower than delay obtained by the other analysed approaches. Figure 7 shows that the average end-to-end delay of AQA-Hybrid is always below 0.12 s in both scenarios. There is a similar trend in relation to the packet loss metric since as the number of sources increases, the difference in the performance of the three approaches also increases.

Even though the traffic load increases, AQA-Hybrid remains a similar performance due to its adaptive scheme since the traffic sources can adapt its data rate to the available bandwidth avoiding network congestion. Consequently, the delay associated to the detection and recovery of the link failures is also reduced. The values of delay and packet loss obtained with AODV+ and QGWS allow us to infer that these proposals are not suitable for supporting multimedia traffic.
6. Conclusions

A new hybrid gateway discovery strategy has been proposed in this paper for improving the interoperability of MANETs and wired networks. Furthermore, we conducted a performance evaluation of our approach together with other similar and recent approaches. Our hybrid gateway discovery algorithm has been integrated in the QoS-aware routing protocol named AQA-AODV. AQA-AODV is a routing protocol suitable for purely ad hoc environments that includes a mechanism for estimation of the available bandwidth and a feedback scheme to provide information to the application. With this information, traffic source may adapt its data rate to network conditions.

With the hybrid discovery procedure added to AQA-AODV, it is not only possible to establish routes according to the QoS requirements of the requesting node in a MANET, but it is also possible to establish communications with a host in a wired network through gateways. Compared with other hybrid approaches, our hybrid gateway discovery incorporates the available bandwidth as a metric during the gateway selection. Moreover, our solution dynamically adapts the size of the proactive area and the frequency of the advertisement messages in order to make a more efficient dissemination of control packets.

The simulation results show that the most substantial improvement of the proposed solution is the ability to reduce the network congestion (i.e., packet loss and delay) without increasing routing overhead.

As a future work, we are considering the evaluation of our routing protocol in wired-cum-wireless networks where video streaming services are provided. Rate adaptive techniques will be evaluated since a video source could exploit the feedback information from AQA-AODV to tune a parameter on the source coding in order to adapt the bit rate to the available bandwidth.

7. References


