End-to-End Support for Statistical Quality of Service in Heterogeneous Mobile Ad-hoc Networks

Jamal N. Al-Karaki[a] and Ahmed E. Kamal[a]

[a]Lab for Advanced Networks (LAN)
Dept. of Electrical and Computer Engineering
Iowa State University
Ames, Iowa 50011
U.S.A

In *heterogeneous* Mobile Ad hoc Networks (MANETs), different types of mobile devices with diverse capabilities may co-exist in the same network. The heterogeneity of MANETs makes end-to-end support for Quality of Service (QoS) guarantees more difficult than in other types of networks, not to mention the limited bandwidth and frequent topology changes of these networks. Since QoS routing is the first step toward achieving end-to-end QoS guarantees in heterogeneous MANETs, we propose a QoS routing protocol for heterogeneous MANETs. The proposed protocol, called Virtual Grid Architecture Protocol (VGAP), uses a cross-layer approach in order to provide end-to-end statistical QoS guarantees. VGAP operates on a fixed virtual *rectilinear* architecture (virtual grid), which is obtained using location information obtained from Global Positioning System (GPS). The virtual grid consists of a few, but possibly more powerful, mobile nodes known as *ClusterHeads* (CHs) that are elected periodically. CHs discover multiple QoS routes on the virtual grid using an extended version of the Open Shortest Path First (OSPF) routing protocol and an extended version of WFQ scheduling policy that takes into account the wireless channel state. Moreover, VGAP utilizes a simple power control algorithm at the physical layer that provides efficient energy savings in this heterogeneous setting. Simulation experiments show that VGAP has a good performance in terms of packet delivery ratio, end-to-end packet delay, call blocking probability, and network scalability.

1. *Introduction*

Mobile Adhoc Networks (MANETs) are typically heterogeneous networks with various types of mobile nodes. In military application, different military units ranging from soldiers to tanks can come together, hence forming an ad hoc network. In conference application, different types of mobile devices such as Personal Digital Assistants (PDAs), smart badges, and laptops may exist in the ad hoc network at the same time. Such nodes differ in their power capacities and computational speeds. Thus, mobile nodes will have different packet generation rates, routing responsibilities, network activities, and power draining rates.

Providing end-to-end support for Quality of Service (QoS) guarantees is a central and critical issue in the design of future multimedia heterogeneous MANETs. Quality of service is more difficult to guarantee in MANETs than in other types of networks for the following reasons. First, the absence of a fixed infrastructure coupled with the ability of nodes to move freely cause frequent route breakage and unpredictable topology changes. Second, the limited bandwidth resource is usually shared among adjacent nodes due to the wireless medium. Third, the nodes themselves can be heterogeneous, thus enabling an assortment of different types of links to be part of the same network.

Numerous routing protocols were proposed to solve the routing problem in MANETs, primarily with no QoS guarantees (e.g., [2],[3],[4],[7]). A comprehensive survey on the traditional routing protocols in MANETs can be found in [15]. Among those protocols, position based routing protocols (PBR) reduce control message overhead by utilizing location information of the nodes in the network to find the route. The location information is usually offered by external means such as Global Positioning System (GPS). PBR protocols usually require only accurate neighborhood information and is localized in nature. Because of localization, PBR schemes have good scalability features. PBR protocols have been surveyed in [7]. Quality of Service (QoS) routing is the first step toward achieving end-to-end QoS guarantees in heterogeneous MANETs. Therefore, design of QoS routing protocols is a crucial problem in MANETs. This requires extensive collaboration between the nodes, both to establish the route and to secure the resources necessary to provide the required end-to-end QoS guarantees.

QoS routing in MANETs is comparatively a new field, and has just started to receive attention. Since QoS routing in MANETs is a difficult problem, little work has been done on this topic. Table 1 summarizes the different QoS routing paradigms in MANETs with a reference to an example of that paradigm placed next to it. For the sake of comparison, the table also lists our protocol VGAP, which will be discussed in section 3. The comparison between different routing protocol classes takes into account many factors like the QoS metrics used, the method used to compute

QoS routes, and if the route discovery process returns single or multiple paths. The table is self explanatory.

In order to circumvent the effect of the frequent topology changes in MANETs, topology management and control algorithms are developed. One method of topology control is *topological clustering* (e.g., [8], [11], [12], [13], [14], [18], [19]), which enhances network scalability by creating a hierarchy of layers in the network. Hence, a MANET can be represented by a set of logical clusters with a subset of nodes, inside clusters, acting like virtual base-stations similar to cellular networks. The selected subset of mobile nodes play different roles in the network. In the rest of this paper, we refer to this subset of nodes as ClusterHeads (CHs). The role of CH is a temporary role, which changes dynamically as the network topology or other factors affecting it change. The set of CHs form a wireless virtual backbone. The wireless virtual backbone can be used to perform routing as well as network management and control functions.

Previous schemes for hierarchical routing in MANETs have focused on two main strategies: the *zone based routing* [14,8] and the *dominating set routing* [11,18,19]. The zone based approach divides the network area into zones (clusters) of variable or fixed size. Inside each zone, a special mobile node is usually elected to act as a CH and provide packet forwarding service to other zone members. In the dominating set routing, a mobile node will be either in the dominating set or a neighbor of a node from the set. The dominating set members can be directly or indirectly connected. The hierarchical scheme proposed in this paper belongs to the zone-based category where the zones are fixed and regular in shape resulting in a fixed and regular virtual topology.

Most of the previous studies have assumed symmetric links and assumed that nodes are homogeneous. Moreover, the hierarchy created is not fixed and this raises the need to have complex topology management schemes when topology changes, except [8] and [14]. However, [14] does not use the concept of CHs that is used in [8] and in this paper. Our clustering approach in this paper is different from [8] in many ways. First, we are addressing nodes of heterogeneous capabilities with different transmission ranges. Asymmetric links may exist in a real heterogeneous network due to variation in transmission power of different nodes and noisy channels. Second, our work is more general in the sense that it constructs multi-layer virtual topology, which is capable of supporting different capability levels. Third, we use rectilinear routing instead of diagonal routing and we show analytically that rectilinear routing is better than diagonal routing used in [8]. Fourth, we introduce a simple and low overhead packet forwarding approach based on the concept of source routing. Finally, [8] uses a very expensive route discovery protocol since it has excessive flooding rate and may create loops.

Most of the previously proposed protocols address the issue of routing from a single-layer perspective, that is, at the network layer. Our view of the routing issue is that inter-layer dependencies play a critical role in providing an efficient and comprehensive solution to the routing problem especially when a route has to satisfy certain end-to-end QoS guarantees, and this view is a key design principle in our proposed protocol. This view is also shared by few recent works, where researchers realized the benefits of using cross-layer interdependencies to enhance MANETs performance, e.g., [20]. Our proposed protocol capitalizes on the cross-layer interdependence between the physical, MAC, and network layers to provide end-to-end statistical QoS guarantees in heterogeneous MANETs with minimum overhead.

1.1. *Paper Contributions*

In this paper, we address the important problem of performing routing with end-to-end statistical QoS guarantees in heterogeneous MANETs. From earlier discussion, we observe that QoS routing in MANETs involves two key subproblems: (1) creating a stable and connected routing architecture that is robust against MANETs frequent topology changes and link failures, and (2) performing QoS routing on the resulting virtual backbone. We address each of these two key subproblems separately, while at the same time providing an integrated solution for the QoS routing problem in MANETs.

Our contribution in this paper is twofold. *First*, we realize that most of the routing problems in MANETs are partly due to the unstable topology of these networks. Therefore, having a fixed, stable, and connected routing architecture (i.e., virtual wireless backbone) that is robust against node movement and frequent link failures in MANETs can simplify those routing problems. We investigate providing such a stable topology with low overhead. For this purpose, we devise a novel zoning strategy for heterogeneous MANETs. The zoning process maps the network physical topology onto a virtual two-dimensional grid topology. The virtual grid consists of a few, but possibly more powerful, mobile nodes known as *ClusterHeads* (CHs), which are elected periodically. The proposed fixed virtual topology provides stable routing backbone that is used to perform routing with the required end-to-end QoS guarantees. *Second*, we propose a QoS routing protocol for heterogeneous MANETs. The proposed protocol, called Virtual Grid Architecture Protocol (VGAP) uses cross-layer design methodology in order to provide end-to-end QoS guarantees. VGAP operates on the derived fixed virtual grid, where CHs discover multiple QoS routes using an extended version of the OSPF routing protocol, called Mobile OSPF (M-OSPF). To compute QoS routes, M-OSPF employs an extended version of

Table 1
Comparison of QoS routing algorithms for Ad hoc networks, OD:On-Demand; C,F: Clustered or Flat; Dist:Distributed; BW:Bandwidth, DLY:Delay, S:Single, M: Multiple, P-Global:Partially Global, TDMA: Time Division Multiple Access.

	Core-based [16]	Ticket-Based [17]	Bandwidth-based [29]	Predictive-based [10]	Position-based [30]	VGAP
QoS metric	BW	BW,DLY	BW	BW,DLY	BW	BW,DLY
State maintenance	Local	Local	Global	2-hop info	Local	P-Global
QoS state propagation	BW changes	Periodic	Periodic	OD	OD	BW changes
Routing class	Dist	Dist	Dist	Dist	Dist	Dist
Route computation	OD	OD	OD	OD	OD	OD
Routing architecture	Clustered	Flat	Flat	Flat	C,F	Clustered
Single or Multiple paths	S	M	M	S	M	M
Power issues	No	No	No	No	No	Yes
Scheduling issues	No	No	No	No	No	Yes
Channel access	TDMA	TDMA	TDMA	TDMA	TDMA	802.11,CDMA

WFQ scheduling policy, called Ad hoc WFQ (AWFQ). AWFQ takes into account the varying time characteristics of the wireless channel state into account, yet providing an upper bound on the statistical end-to-end delay guarantees of a certain flow. The possibility of using OSPF as a routing protocol in MANETs has been recently discussed by the IETF MANET group [26] but with no specific details.

The rest of this paper is organized as follows. In Section 2, we introduce the network model, the virtual network topology, and the power control scheme. The proposed extension to the OSPF routing protocol used by our protocol for QoS route discovery and maintenance is presented in Section 3. An extension to the WFQ scheduling discipline in MANETs is introduced in Section 4. Performance evaluation of the proposed scheme and the simulation results are presented in Section 5. Finally, conclusions and directions for future research work are presented in Section 6.

2. *Network Model and Virtual Network Topology*

We consider a heterogeneous network that consists of N mobile nodes of different capabilities that are randomly distributed in the network area. To utilize location information, we assume that each mobile node is equipped with a Global Positioning System (GPS) card[1].

2.1. The Zoning Process

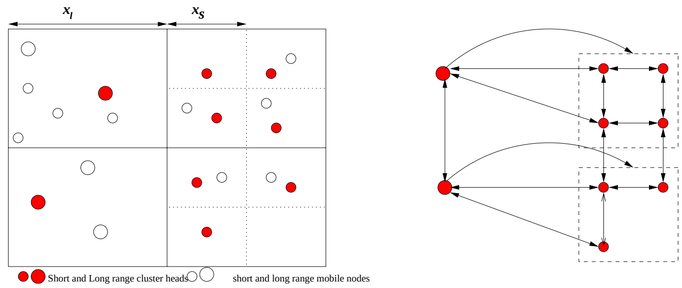
For the sake of simplicity, but without loss of generality, we assume that mobile nodes are of two different transmission ranges: r_s for a Short Range (SR) nodes and r_l for a Long Range (LR) nodes[2]. The transmission range in each case is a function of the energy level at the mobile node. The network area is divided into fixed, disjoint, and regular shape zones (clusters). To create a simple rectilinear virtual topology, we select the zones to be square in shape[3].

A zone may have a mixture of SR and LR nodes. Initially, the network area is divided into large zones where the zone side length x_l is chosen such that two LR mobile nodes in adjacent horizontal/vertical zones and located anywhere in the zone can communicate with each other directly. Therefore, x_l is selected as $x_l = \frac{r_l}{\sqrt{5}}$. If the zone has only short range nodes, the zone is divided into *subzones* where each subzone side length (x_s) is computed as $x_s = \frac{x_l}{2}$. The number of layers in the virtual topology is determined by the number of transmission ranges in the network. In Figure 1(b), the left side zones have both SR and LR nodes while the right side zones have only SR nodes. Therefore, each of the right side zones is divided into four subzones to allow direct communication between any two short range mobile nodes in adjacent horizontal/vertical subzones. As a result, the control overhead is limited to the communication between CHs and in the associated vertical and horizontal directions only.

2.2. *Cluster Head Election*

Since mobile nodes differ in their capabilities, an eligibility criteria is needed to determine which mobile node can become a CH. Note that parameters like remaining battery energy, node speed, and node location should count toward this decision. The CH role is rotated periodically among nodes in each zone. The CH periodicity helps to balance the nodes' load distribution, achieves fairness, and provide fault tolerance against single node failure. We define the time between two successive CH election instants as the CH election period.

[1] A GPS-free approach [5] may also be used if GPS is infeasible for a certain application.
[2] The extension to more than two transmission ranges is possible and follows the same procedure outlined in this section.
[3] Extension to other virtual topologies is possible with the following tradeoffs. If the zone shape is a hexagon instead of square, the routing function will be more complex. If, on the other hand, the zone shape is linear or triangle, the resulting virtual topology will be very simple but less efficient.



(a) Zoning with Physical Topology

(b) Corresponding Virtual Topology

Figure 1. Network Zoning and the virtual topology

Each mobile node decides whether it will become a CH for the current election period based on an eligibility factor (EF). The least mobile, most power capable, most location suited, and most lightly loaded node is the best candidate to act as CH. The eligibility of a node i to serve as a CH at time t can be calculated as follows:

$$EF_i(t) = a_1 e^{-v_i(t)} + a_2(1 - Q_i(t)) + a_3 B_i(t) + a_4(1 - S_i(t))$$

where $v_i(t)$ is the average speed of node i at time t, $B_i(t)$ is the remaining power in node i battery at time t, $Q_i(t)$ is the fraction of time node i remains as CH during a time window T (defined as a window of time sufficient to capture cluster dynamics) ending at time t, $0 \leq Q_i(t) \leq 1$, and found as follows:

$$Q_i(t) = \frac{1}{W} \sum_{m=1}^{W} Y_i(t - m\delta t)$$

where $Y_i(t)$ is a binary indicator which is 1 if and only if node i was a CH at time t, and δt is a small time increment such that $T = W\delta t$. $S_i(t)$ is the Euclidean distance of node i with respect to the center of a zone at time t, and a_1, a_2, a_3, a_4 are scaling factors that reflect the importance of each parameter. Mobile nodes in each zone will exchange the values of the calculated EF's. The node that has the highest value of EF will elect itself as the CH for the current election period. That is, if a zone z has a nonempty set of n nodes, the CH is selected as follows:

$$CH = arg \max_{i \in n}(EF_i)$$

Each mobile node that has elected itself as a CH for the current period will broadcast an advertisement message to the set of the mobile nodes in its zone. If the CH leaves its zone or fails early, an early election of the next CH takes place after a *handoff* procedure in which the new CH inherits the routing cache of the retired CH. It is worth mentioning that the concept of CH election is not a new concept. However, our zoning approach simplifies its operation. Hence, the routing function will also be simplified.

2.3. The Virtual Topology

The set of CHs and the wireless links connecting them is modelled as a directed graph $G = (V, E)$ where V is the set of CH nodes and E is the set of directed wireless links connecting CH nodes. Each nonempty zone corresponds to a vertex in the graph G. Vertical and horizontal adjacent vertices (CHs) are connected by an edge if the distance between them is less than or equal to the corresponding transmission range and the two adjacent zones are not empty. A link is absent from E if either one or the two neighboring zones are empty or a communication in the respected direction is not possible. We refer to the resulting virtual topology as *Virtual Grid Architecture* (VGA). Note that the CH election process selects the most eligible node in every zone to act as a CH. Thus, the resulting set of elected CHs tend to remain in that status for long time relative to other nodes in their zones resulting in stable virtual topology. Consequently, the routes found over VGA exhibit a relatively long lifetime when the zones maintain CHs. This result proves to be helpful when performing end-to-end QoS routing over VGA. Note also that the control overhead is limited to the communication between CHs and in the associated vertical and horizontal directions only.

A packet forwarding scheme can be easily implemented over VGA. One simple packet forwarding scheme can be implemented as follows. The standard four directions (North(N), South(S), West(W), East(E)) are used for simple packet forwarding over VGA. Those directions can be encoded using a 2-bit representation such that (00-01-10-11) correspond to (N-S-E-W) directions, respectively. Note that opposite directions differ only in the least significant bit, which simplifies route traversal. The route directions are embedded in the packet header by the source node where each hop is represented by the two bits, therefore resulting in very low overhead. Each intermediate CH inspects the route in the packet header and forwards the packet to the four directions accordingly. An additional 2-bit encoding is also used to select a specific subzone in a neighboring zone that has only SR nodes and the direction of communication is from LR node to a specific SR node in that neighboring zone.

2.4. *A Simple Power Control Scheme*

A power control scheme is needed in heterogeneous networks to achieve network connectivity. The scheme determines the power levels required for inter-zone and intra-zone communications. Since we have two transmission ranges, four types of communications between mobile nodes are defined (Short→Short(S-S), Long→Long(L-L), Long→Short(L-S), and Short→Long(S-L)). Note that a LR node can communicate directly with all other nodes. However, the problem is when a SR node wants to communicate with a LR node inside a zone or with a neighboring zone. To solve this problem, we devise a simple power control mechanism as follows. Assume the distance between two mobile nodes is d. The transmitted signal strength is known to decrease with the distance d according to the relation [32],

$$P_r = \frac{P_t}{d^{\alpha}}$$

where P_t, P_r are the transmitted and received signal power, respectively, and α characterizes the steepness of the decrease in the signal power and it assumes values in the range $2 \leq \alpha \leq 4$. For convenience, we define the ratio of the transmitted signal power of a mobile node to that of an LR node, denoted by RTP, as $RTP = \frac{P_t}{P_{max}} = (\frac{r}{r_l})^{\alpha}$, where P_t and P_{max} are the transmitted signal power of the target node and LR node, respectively and $r = r_s$ or r_l for SR nodes and LR nodes, respectively. Figure 2 shows the different cases of intra-zone and inter-zone communication. The distances d_1, d_2, d_3 represent the maximum distance separation between CHs for inter-zone communication, while d_4, d_5 correspond to the maximum distance separation between mobile nodes and CHs for intra-zone communications. From Figure 2, we find the worst case values of RTP (based on a worst case value of α=4) for different types of communication. The results are summarized in Table 2. The values between parentheses are the worst case values of RTP corresponding to each type of communication. They were obtained through simple derivations. For example, if the sender is a LR node, then it uses its maximum power (P_{max}) to communicate with another LR range in a neighboring zone. An LR node can use a reduced power P_{red} to communicate with a LR mobile node inside a zone, which corresponds to a distance of $\sqrt{\frac{2}{5}}r_l$. Assuming that the receiver sensitivity for all mobile nodes is the same, then

$$P_{red} = \frac{4P_{max}}{25} < \frac{P_{max}}{6}$$

This means that transmission power can be reduced by a factor of 6.

Mobile nodes inside a certain zone, which are not acting as CHs, communicate directly using the IEEE 802.11 MAC protocol mechanism. Channel communication management between different zones (CHs) is controlled by using code division multiple access (CDMA). The assignments of codes to zones is according to [22], where neighboring zones use different spreading codes in order to reduce interference and to enhance spatial reuse of channels.

2.5. *Rectilinear Routing versus Diagonal Routing*

In VGA, routing is implemented in the vertical and horizontal directions only. The extension to diagonal routing is possible but will suffer two side effects. First,

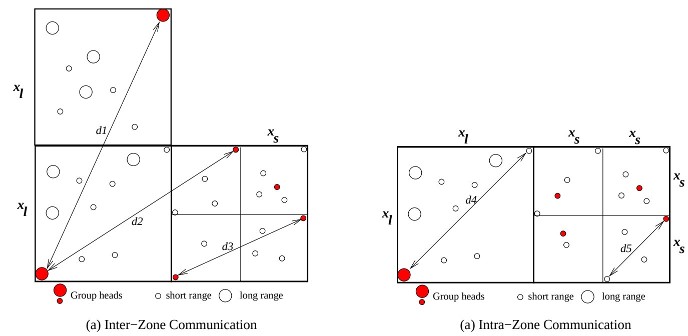


Figure 2. (a) Inter-zone communication (b) Intra-zone communication

it may complicate routing since the number of potential neighbor zones doubles, and second it will reduce the zone side length to a maximum of $\frac{r_l}{2\sqrt{2}}$ per side as opposed to $\frac{r_l}{\sqrt{5}}$ when rectilinear routing is used. In fact, this can adversely affect the route length. In [31], we showed that (1) the probability of rectilinear routing failing while diagonal routing succeeding is low, and (2) the addition of the diagonal routing to VGA will increase the average route length. We refer to VGA with diagonal routing as (D-VGA). First, we find the

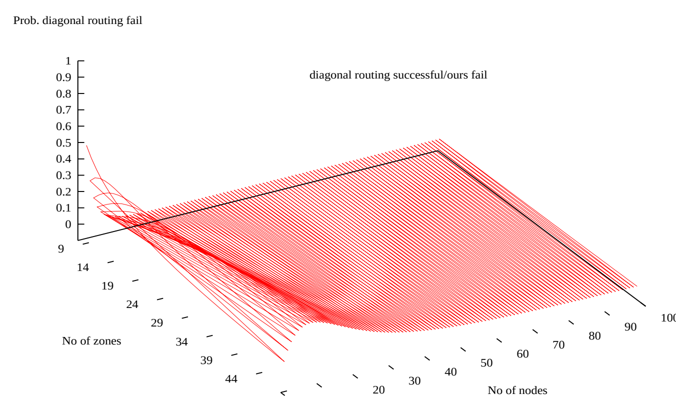


Figure 3. Probability that a rectilinear routing fails

probability that the rectilinear routing from a target zone fails while the diagonal routing succeeds, $P(B)$. Assume that the target zone has at least one node, vertical and horizontal zones are empty, and at least one diagonal zone is not empty. We compute $P(B)$ as:

$$P(B) = \sum_{n=1}^{N} P(B|n)\binom{N}{n}(\frac{|Z|-9}{|Z|-1})^{N-n}(\frac{8}{|Z|-1})^{n} \quad (1)$$

Table 2
Summary of Intra-zone and Inter-zone communication ranges and RTP values

Type	S-S	S-L	L-S	L-L
Intra-zone range(RTP)	$\frac{r_l}{\sqrt{10}}(0.1)$	$0.63r_l(0.39)$	$\sqrt{\frac{2}{5}}r_l(0.39)$	$\sqrt{\frac{2}{5}}r_l(0.17)$
Inter-zone range(RTP)	$\frac{r_l}{2}(0.25)$	$0.81r_l(0.65)$	$\frac{3}{2\sqrt{5}}r_l(0.4)$	$r_l(1.0)$

where $P(B|n) = \frac{\binom{n+3}{3}}{\binom{n+7}{7}}$ and $\binom{N}{n}(\frac{|Z|-9}{|Z|-1})^{N-n}(\frac{8}{|Z|-1})^n$ represents the probability that out of the total N nodes in the network area, $1 \leq n \leq N$ nodes are in the eight zones around the target zone (central zone). We plot $P(B)$ for different numbers of zones and mobile nodes in Figure 3. As shown in the Figure, the probability of a successful diagonal routing given that rectilinear routing fails is very low. Hence, adding diagonal routing may not payoff for the additional complexity that is associated with D-VGA. Next, we find the average route length (i.e., number of hops) for VGA and D-VGA. The average route length using rectilinear routing (VGA) is shorter than the average route length of (D-VGA). This fact is formally stated in the following proposition whose proof has been omitted because of the lack of space.

Proposition: If the network area is large and is $L \times L$ square meters, the mobile nodes are uniformly distributed in the network area, and each node has a transmission range[4] of r then the average route (in terms of the number of hops) is given by:

$$\text{VGA} = \frac{\sqrt{5}L}{r} \quad , \quad \text{D-VGA} = \frac{5\sqrt{2}L}{3r}$$

Proof. see [31]. □

This can be intuitively justified by observing that with diagonal routing, zones will be smaller in size, and as such the total network area will contain more zones which lead to longer routes in terms of number of hops. In the following section, we present in details our QoS routing protocol that uses VGA as its underlying routing structure.

3. *Quality of Service Routing using VGA*

In this section, we propose a QoS routing protocol for MANETs. The proposed protocol, called Virtual Grid Architecture Protocol (VGAP), operates on VGA. VGAP employs a routing strategy, which is an extended version of the Open Shortest Path First (OSPF) routing protocol [25] coupled with an extended version of the Weighted Fair Queueing (WFQ) scheduling discipline [23], to provide end-to-end statistical QoS guarantees. The extended version of OSPF is called Mobile OSPF (M-OSPF) while the extended version of WFQ is called Ad hoc WFQ (AWFQ). AWFQ scheduling discipline takes into account the time varying characteristics of the wireless channel as described in section 4. M-OSPF utilizes AWFQ to provide link costs when computing or discovering a QoS route. OSPF is currently used in the Internet. Hence, the use of OSPF in MANETs facilitates the integration between MANETs and the Internet where a MANET can be viewed as one area of the OSPF routing domain as shown in section 3.1.

VGAP is able to compute and discover QoS routes over VGA that satisfy end-to-end bandwidth and delay guarantees. We formally define the QoS route as follows. A QoS route $P(s,d)$ on the virtual rectilinear graph $G = (V, E)$ with a length of n hops and connecting a source-destination pair (s,d) can be represented as a sequence of vertices $v_0, ..., v_k, ..., v_n$ where $0 < k < n$ and $v_0 = s,\ \ v_n = d$. Each link l, $l \in E$, has two associated link metrics, namely, b_l and d_l which correspond to the bandwidth capacity and the packet delay on link l, respectively. Each link l on the path $P(s,d)$ starts at vertex v_{i-1} and ends at vertex v_i for $1 \leq i \leq n$. A source mobile node issues a call to a certain destination with specific QoS requirements in the form $\langle s, d, R_{bw}, D_{max}\rangle$ where R_{bw} is the minimum required bandwidth and D_{max} is the maximum end-to-end delay. A call is admitted if R_{bw} and D_{max} can be satisfied, i.e.,

$$R_{bw} \leq \min_{l \in P(s,d)} b_l \quad ; \; D_{max} \leq \sum_{l \in P(s,d)} d_l$$

We define the path with the minimum number of hops, yet satisfying the required QoS metrics defined above as the *optimal* path.

3.1. *M-OSPF: An Extension to OSPF in MANETs*

OSPF, originally designed for wireline networks, has some limitations when employed in all wireless domains like MANETs. *First*, OSPF has conventionally been designed to work with links that have known bandwidth. This does not carry over to MANETs. MANETs have variations in their available capacities due to the time varying wireless channel characteristics. *Second*, in OSPF, topology update messages are periodically flooded to the network routers[5]. Periodic flooding consumes more resources in a limited bandwidth environment like MANETs. Reducing flooding

[4]Note that r can refer to any of the two transmission ranges r_l, r_s. However, for the sake of comparison between VGA and D-VGA, we set the transmission range to a general r.

[5]We use the terms CH and router interchangeably.

frequency will save bandwidth and as such updates to the QoS routing tables should be performed only when necessary.

In OSPF, a router (CH) has to learn the identity and resource status of all its neighboring routers. Messages, called Link State Advertisements (LSAs), are used for this purpose. Moreover, since MANETs are heterogeneous, some nodes may not be QoS capable (i.e., does not have the resources to support *any* QoS guarantees) and they can only support traditional routes. For a node to advertise its QoS capabilities, we make use of the T bit, currently unused, in the OSPF options field. A QoS capable node sets this bit in its LSAs. The options field is sent in all OSPF Hello packets and all LSAs. Hence, mobile nodes of different capabilities can be mixed within one OSPF area

We propose an extended version of OSPF that works in MANETs, called Mobile OSPF (M-OSPF). Analogous to OSPF terms, the MANET geographical region is modelled as one OSPF area since a MANET normally spans a limited region, which is much smaller than OSPF autonomous system. The set of created zones are modelled as the set of subnets in the OSPF area, and the set of CHs in the zones act as the routers inside the area subnets.

M-OSPF builds its routing tables as follows. First, CHs initialize their routing data structures (tables or topology database as used in OSPF) by using the OSPF Hello protocol to acquire neighbor CHs. In addition to helping acquire neighbors, hello packets also act as *I'am alive* to let CHs know that other CHs are still functional. Each CH sends a message to provide information to its adjacent CHs (which are four or less) about initial link states. These messages, known as Link State Advertisements (LSAa), also include the resource information such as available bandwidth and link processing and queueing delays to all other CHs. Once received, the CH stores this information in the QoS routing table and it is used in the QoS route computation algorithm. M-OSPF employs a threshold based model described later in this section to report link states changes. By comparing already established routing tables to reported link states, failed routers can be detected quickly, and the network topology can be altered appropriately.

M-OSPF uses a *hybrid* criterion in computing end-to-end QoS routes over VGA. When a new request is received by a router (CH), it computes a QoS route locally using the stored QoS routing tables. If the QoS route computation process fails, the router initiates a route discovery process on demand. A user (application layer) sends its request to the network layer for a QoS route with the required metrics in the form $\langle s, d, D_{max}, R_{bw} \rangle$. From the topological database (routing tables) that was generated from LSAs, each router calculates a QoS shortest-path tree, with itself as root which contains all routes that can satisfy R_{bw} first and then finds the routes that satisfy D_{max}. The shortest-path tree, in turn, yields a QoS routing table. The QoS routing table is inspected for an available QoS route. If a QoS route is available, the request is admitted and the route is used. If more than one QoS route is found, the optimal QoS route, if any, is used and other routes are used as backup routes. If no routes can be computed, an on-demand route discovery process is generated. If the route reply also returns no QoS route, the request is rejected. The reason behind using this hybrid QoS routing criteria is to maximize the number of accepted calls (requests).

When a QoS route is broken as a result of a zone becoming empty or the CH fails for any reason, the route maintenance process must be started. A route may also fail if the source node or destination node leaves the original zone and moves to another zone. Note that the former event is rare in dense MANETs. However, mobility of source and destination nodes cannot be controlled. Hence, we need to discover the new location of a source or destination nodes after they leave their zones. For this purpose, a simple on-demand query cycle that consists of location request (LREQ)/location reply (LREP) messages can be used, which is handled similar to RREQ/RREP used in on-demand ad hoc routing protocols, but messages broadcast is limited to only at most four neighbor CHs.

Since periodic message advertisement in OSPF is a costly process, we propose a simple and efficient scheme that can be incorporated in M-OSPF to limit link state updates and to avoid false and unnecessary updates. Our scheme is a threshold-based triggering scheme described as follows. The scheme is characterized by two constant threshold values (Upper threshold (U_{th}) and Lower threshold (L_{th})). Based on the two threshold values, three link classes C_1, C_2, C_3 can be defined. Updates are triggered when a class boundary is crossed. At some particular CH node i, an update is triggered when class boundary is crossed, i.e., if the actual value of the link bandwidth crosses L_{th} or U_{th}. An oscillation around a class boundary may happen each time the available bandwidth value crosses a class boundary hence the available bandwidth fluctuates around that class boundary. In order to stop such behavior, the scheme is augmented with a *hysteresis* mechanism, which requires that change in bandwidth be significant (larger than a certain value) in order to advertise a link state change message. In addition, only a fixed value for the bandwidth (lower bound of the class) is advertised when the link state remains in that class, which may correspond to the IEEE 802.11b data rates of 11 Mbps, 5.5 Mbps, and 2 Mbps.

3.2. *Route Maintenance in VGAP*

One main advantage of the proposed protocol is that the routes found exhibit a relatively longer lifetime when the zones maintain CHs, i.e., do not become empty. This is due to the stability feature of the proposed VGA. To see this, we note that VGA is updated only when a zone becomes empty or becomes occupied after being empty. The former results in link deletion and the latter results in link addition to the virtual grid structure. For dense networks, these events become very rare leading to a very stable structure. This is because a vertex in VGA does not change as long as there is at least one node in the respected zone. However, when a link is deleted (i.e., link fails) for a certain route, the route maintenance process in VGAP will be initiated. A link can also fail if a CH becomes congested or fails for other reasons (e.g., runs out of power).

VGAP can recover from this situation as follows. When a link breaks or gets heavily congested, an error message is sent back to the sender, which causes it to use one of the backup routes found earlier as an alternate route. If no backup route is available, the source node looks for a new route. Note that this puts the burden on the source node to always find an alternate solution. However, trying local re-routing (detour) first before the global source re-routing may help alleviate the problem and provide quick recovery from the error, hence improving the network performance. We refer to the re-routing scheme we use as ($local-then-global$). Depending on the situation, the upstream CH of the failed link may find an alternate route on demand by contacting its neighbor possible four CHs, where it could happen that one of them may have a route to the destination sharing joint links with original route. If a neighbor CH has a knowledge of a route to the destination with the required QoS, the current CH informs the source node about the original route failure and the rerouted route, if found, such that the source will decide whether to continue transmitting packets with this locally rerouted path or resort to another global solution, i.e., send packets on a different new route where backup routes can be used. The above discussion results in the reduced transmission delay of packets directed to a broken or congested link and the increased packet delivery ratio of VGAP protocol as shown in the simulation results.

4. *A Modified WFQ Scheduling Policy for MANETS*

In this section, we propose an extension of the Weighted Fair Queueing (WFQ) [23] for MANETs, called Ad hoc WFQ (AWFQ) that is used in VGAP to provide end-to-end statistical bounds on the required delay guarantees. If nodes adopt a WFQ-like service discipline and the source traffic is constrained by a leaky bucket, an upper bound on the end-to-end delay and bandwidth guarantees can be provided [23][6].

[6]Reference [23] also provides a detailed description of the operation of WFQ.

Formally, given a leaky bucket $\langle \sigma, b \rangle$ that constrains the source node traffic where σ denotes the rate at which tokens are accumulated and b is the depth of the token bucket (in bytes), a required bandwidth r, link propagation delay is τ_i; then the end-to-end delay bound on a path P with m hops is given by [23],

$$D_{WFQ} = \frac{b}{r} + \frac{(m-1)L_{max}}{r} + \sum_{i=1}^{m} \frac{L_{max}}{R} + \sum_{i=1}^{m} \tau_i$$

However, when fair queueing algorithms are used over wireless networks, the delay bound may not hold. This is due to the bursty and location dependent channel errors of a wireless link. In fact, most of the scheduling algorithms that work well in wireline networks do not carry over their desirable properties to wireless environment; they rather lose those properties. Hence, error compensation schemes were devised for infrastructered wireless packet networks based on the lead/lag model (see [28] for a useful survey). The proposed AWFQ also uses an error compensation scheme. However, AWFQ differs from all previous related work in the following ways. First, previous compensation models work for wireless networks with infrastructure where base stations always maintain the scheduler. Base stations use a reliable wireline network to communicate with each other. In MANETs, all hops are wireless and therefore packets are prone to bad transmission media on all hops along the path to the destination node. Second, all previous compensation schemes work incorrectly unless all flows have the same packet size. In our scheme, lead and lag of all flows are defined in terms of virtual time which is flexible enough to deal with packet size differences. Third, the amount of compensation in other schemes is not bounded. If the amount of compensation for lagging flows is large, error-free flows will encounter service degradation that may affect the QoS guarantees of these flows. In our extended version of WFQ, we limit the amount of compensation to a maximum value C_{max}, that is also dependent on the amount of additional delays encountered by the lagging flows. For ease of reference, Table 3 summarizes all notations used in this section.

4.1. *The Proposed Extension: Ad hoc WFQ (AWFQ)*

Under AWFQ, if the k^{th} packet of flow i, denoted by $p_{(k,i)}$, cannot be transmitted to some other CH node due to channel error, its transmission will be deferred and other flows that have error-free channels will be allowed to transmit. As this packet has a deadline to meet, and because it encountered additional delay at the current node, this additional amount of delay should be compensated for at downstream nodes (CHs) in order for the packet not to miss its deadline. AWFQ does this by modifying (decreasing) the packet virtual finish at the next node in a manner that increases the probability of meeting the end-to-end delay guaran-

Table 3
AWFQ Model parameters

Notation	Definition
$p_{(k,i)}$	k^{th} packet of flow i
ψ_i	Service share of flow i
θ	Amount of increment in service share at next node
$W_i(t_1, t_2)$	Amount of traffic served in (t_1,t_2) of flow i
$W_i^c(t)$	Waiting time counter value of flow i at time t
$C(p_{(k,i)})$	Amount of service time reduction of packet $p_{(k,i)}$ at the next node
C_{max}	The maximum amount of compensation time given to packet $p_{(k,i)}$ at the next node
$A_{(k,i)}, A_{(k,i)}^{(2)}$	Arrival time of k^{th} packet of flow i at nodes 1 and 2, respectively.
$S_{(k,i)}, F_{(k,i)}$	Service start and finish times of k^{th} packet of flow i, respectively.
$S_{(k,i)}^v, F_{(k,i)}^v$	Virtual service start and finish times of k^{th} packet of flow i at node 1, respectively.
$S_{k,i}^{v2}, F_{k,i}^{v2}$	Virtual service start and finish times of k^{th} packet of flow i at node 2, respectively.
V_{new}	New virtual finish time of a donor packet at node 2.
$V(x)$	Virtual time corresponding to real time x
R	Transmission link rate (bps)
$B(t)$	Set of all backlogged (nonempty) flows at time t
$D(t)$	Set of all flows perceiving channel errors and get delayed

tees. The amount of compensation (the modified virtual service time) is computed at the node where the delay occurred, and is carried in the packet header to the next node. We assign each flow i a waiting time counter, $W_i^c(t)$, which stores the amount of additional time delay a flow i packets experience due to channel error. We define θ as the service share increment of the lagging packet applied at the next node and ψ_i as the service share of flow i.

We now show how the virtual finish time of the packet that incurs additional delay is modified at the next node in accordance with the amount of the additional delay encountered. Without loss of generality, we carry the analysis at any two consecutive nodes along the route (current node and next node) and with the aid of the ideal Generalized Processor Sharing scheduler service curve used by WFQ shown in Figure 4.

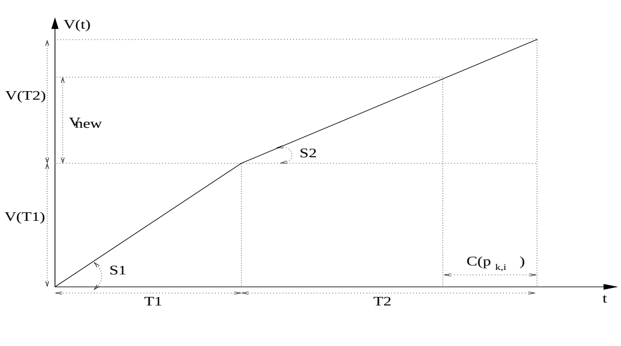


Figure 4. GPS service curve at two consecutive nodes in a route

In the service curve, the following parameters are used. We call the current and next node nodes 1 and 2, respectively. T_1 and T_2 are the GPS packet service times at nodes 1 and 2 and are equal to $\frac{L(p_{(k,i)})\sum_{j\in B_1}\psi_j}{\psi_i}$ and $\frac{L(p_{(k,i)})\sum_{j\in B_2}\psi_j}{\psi_i}$, respectively, where B_1 and B_2 are the set of backlogged flows at nodes 1 and 2 respectively. $V(T_1)$ and $V(T_2)$ are the packet virtual service times at nodes 1 and 2, and we assume that $V(T_1)$ is equal to $\frac{L(p_{(k,i)})}{\psi_i}$, S_1 and S_2 are the service curve slopes at nodes 1 and 2 and they are equal to $\frac{1}{\sum_{j\in B_1}\psi_j}$ and $\frac{1}{\sum_{j\in B_2}\psi_j}$, respectively. V_{new} is the new virtual finish time of the packet at node 2. $C(p_{(k,i)})$ is the amount of compensation (reduction in service time) given to a certain packet at the next node and we set $C(p_{(k,i)}) = \min(W_i^c(t), C_{max}, \frac{L(p_{(k,i)})\sum_{j\in B_2}\psi_j}{\psi_i})$, where C_{max} is the maximum compensation given to a packet. Note that $C(p_{(k,i)})$ is upper-bounded by the amount of additional delay encountered by flow i ($W_i^c(t)$) and also by the real service time under GPS. We seek to find the new virtual time (V_{new}) of the delayed packet. From Figure 4, we have $\frac{T_2}{T_2 - C(p_{(k,i)})} = \frac{V(T_2)}{V_{new}}$. Substituting the values of T_2 and $V(T_2)$ from above, and noticing that ψ_i is the packet service share at the next node, we have

$$V_{new} = \frac{L(p_{(k,i)})}{\psi_i} - \frac{C(p_{(k,i)})}{\sum_{j\in B_2}\psi_j} \quad (2)$$

Hence, the new virtual finish time of the delayed packet at the next node is given by

$$F_{k,i}^{v2} = max\{F_{k-1,i}^{v2}, V(A_{k,i}^2)\} + V_{new}$$

Decreasing the virtual finish time of the packet will assign the packet a precedence in the scheduling at the new node. The above reduction in service time corresponds to an increment in the service share at next node by a value of θ, which is computed as follows. Since $V_{new} = \frac{L(p_{(k,i)})}{\psi_i(1+\theta)}$, the value of θ can be easily found as:

$$\theta = \frac{1}{\frac{L(p_{(k,i)}) \sum_{j \in B_2} \psi_j}{C(p_{(k,i)}) \psi_i} - 1}$$

4.2. *End-to-End Statistical Delay Guarantees using AWFQ*

In this subsection, we derive the probability that a packet, experiencing a channel error, is still meeting its end-to-end delay requirement when AWFQ is used. For simplicity, we model the wireless channel using a two state Markov chain where a wireless channel will be either in Good(G) or Bad(B) states. Assume for simplicity and for the sake of analysis that a packet needs one time unit for transmission. Let $1/\lambda_g$ and $1/\lambda_b$ be the average time the channel is in good and bad states, respectively. Then, the transition probability matrix of the Markov model becomes:

$$= \begin{bmatrix} GG & GB \\ BG & BB \end{bmatrix} = \begin{bmatrix} 1-\lambda_g & \lambda_g \\ \lambda_b & 1-\lambda_b \end{bmatrix}$$

and the steady state probabilities of being in Good/Bad states, π_G, π_B, are given by

$$\pi_G = \frac{\lambda_b}{\lambda_g + \lambda_b} \quad ; \quad \pi_B = \frac{\lambda_g}{\lambda_g + \lambda_b}$$

A packet with erroneous channel continues to wait until its channel state becomes good or the packet misses its deadline and in this case it is dropped. As such, the time elapsed before the successful transmission of an undropped packet follows a geometric distribution at each node with a probability generating function found as follows. At a certain node, a packet at the head of the queue will not wait if its channel is in good state (with probability π_G). If it finds the channel in a bad state (with probability π_B), it will wait for a geometrically distributed length of time with probability λ_b. For a path of length m hops and assuming independence between hops, the probability generating function of the total additional waiting time is:

$$G_{D,m}(z) = (\pi_G + \frac{\pi_B z \lambda_b}{1-(1-\lambda_b)z})^m$$

From $G_{D,m}(z)$, we can find the probability mass function d_n, which is the probability that the total additional delay is equal to n time units and it is found as follows:

$$d_n = \frac{1}{n!} \frac{d^n(G_{D,m}(z))}{dz^n}|_{z=0} \tag{3}$$

By finding d_n, we can find all desired performance metrics. To find d_n, let $X = (G_{D,m}(z))^{(1/m)}$ and $Y = G_{D,m}(z)$ and let $q = 1-\lambda_b$, $p = \pi_B \lambda_b$, $a = p/q$. Then X can be written as:

$$X = \pi_G + (-a + \frac{a}{1-qz}) = (\pi_G + \frac{\pi_B z \lambda_b}{1-(1-\lambda_b)z})$$

The n^{th} derivative of X is easily found as:

$$X^{(n)} = (-1)^n n! \frac{a(-q)^n}{(1-qz)^{n+1}}$$

while the n^{th} derivative of Y is found to be:

$$Y^{(n)}|_{z=0} = \frac{m\sum_{i=1}^{n} X^{(n+1-i)} Y^{(i)} - \sum_{i=2}^{n} Y^{(n+1-i)} X^{(i)}}{X^{(1)}}|_{z=0} \tag{4}$$

Substituting equation (4) in equation (3), we can recursively find the expression for d_n. Using the resulting equation, the probability of having an additional delay time in AWFQ, D_{AWFQ}, which is less than a certain value $\mathbb{W}$ can be found as follows:

$$P(D_{AWFQ} \le \mathbb{W}) = \sum_{i=0}^{\mathbb{W}} d_i$$

Therefore, the deferred packet still meets its deadline if $(D_{AWFQ} \le D_{max} - D_{WFQ} + \mathbb{C})$, where $\mathbb{C}$ is the total amount of applied compensation time to even $\mathbb{W}$ and is bounded by C_{max} at each hop. Hence,

$$Pr(meeting\ \ D_{max}) = Pr(D_{AWFQ} \le D_{max} - D_{WFQ} + \mathbb{C}) \tag{5}$$

5. *Performance Evaluation*

The VGAP protocol was simulated using NS 2.26 simulator [33]. A heterogenous MANET in an area of size (2000m×2000m) with a variable number of mobile nodes was simulated. The long range transmission distance is set to 250 meters. Mobile nodes were initially placed randomly within the fixed-size network area. Mobile nodes roam around with speed uniformly generated between 0∼20 m/s, and uses a realistic probabilistic random walk mobility model [6]. The number of mobile nodes is set to 200 unless otherwise stated. Links have a maximum data rate of 11Mbps. Calls arrive to the system according to poisson process with mean inter-arrival time that was varied to control the traffic load into the network. The call holding time is exponentially distributed with a mean of 60 seconds. Unless otherwise specified, the default offered traffic load was set to 0.8. Each call is specified as a randomly chosen source-destination (*s*-*d*) pair. The packet sizes are exponentially distributed with mean 512 bytes. Every node has an initial energy level (*full*) at the beginning of a simulation. For every transmission and reception of packets, the energy level is decremented by a specified value, which represents the energy usage for transmitting and receiving. The initial energy of each node is 8000 joules so that the energy level does not reach zero in the simulation period. Power requirements for receiving is set to 1.0 W and for transmitting to 3.0 W. These two values conform with commercial specifications. Two types of calls were generated, Audio and Video. The requested bandwidth of each call is uniformly distributed between 16∼64 Kbps Audio and 0.5∼2 Mbps Video. Audio calls constitute 40% of the total generated calls and the rest are video calls. For the channel prediction, we used a two state Markov chain which models the wireless channel as being either in Good(G) or Bad(B) states. In the simulations, we set $\lambda_g = 0.1$ and $\lambda_b = 0.9$ per average packet transmission time. For the link state update model, we set $((L_{th})$=0.3 and $(U_{th}))$=0.7 of the link rate. In another setting, we also vary those two values to study the behavior of the model and its effect on the performance of VGAP. We evaluate the performance of the proposed protocol based on the following metrics: (1) Packet delivery ratio (2) End-to-end packet delay (3) Call acceptance ratio (4) Path optimality (5) Scalability (6) Normalized power consumption.

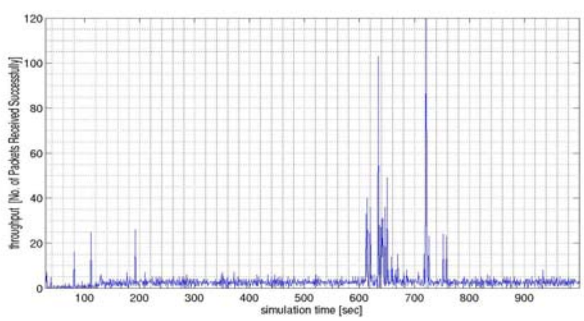


Figure 5. Network throughput versus simulation time

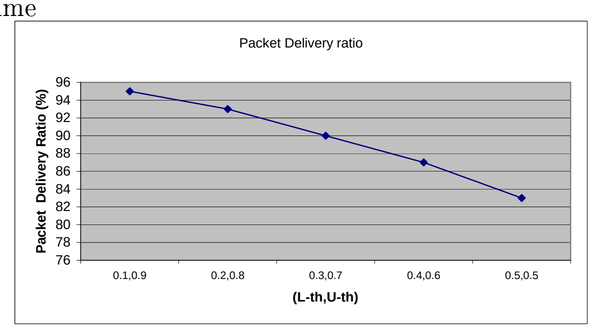


Figure 6. Packet Delivery ratio versus the link state update thresholds

Packet Delivery Ratio: We define the packet delivery ratio as the ratio between the number of packets received by the destination and the number of packets

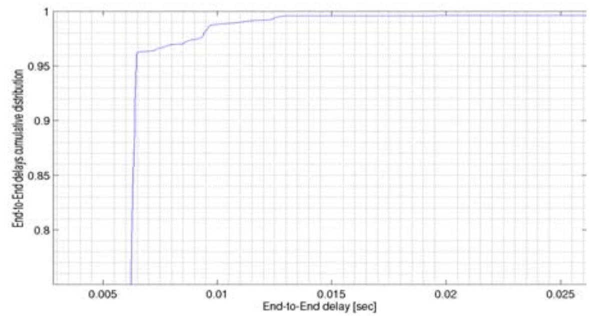


Figure 7. Cumulative distribution of end-to-end packet delay

generated by the application layer sources of the accepted calls. Packet delivery ratio is important as it describes the loss rate that will be seen by the transport protocols, which in turn affects the maximum throughput that the network can support. Figure 5 shows the evolution of throughput with simulation time from which the network status can be inspected at different time periods. The figure shows the number of packets delivered successfully every second during the simulation period. The expected successful delivery rate of the sent packets (average throughout) is 5 packets/sec. From the throughput curve, we can see that network was able to meet the expectations at most of the time. From the figure, we can also know or predict the packet loss in the network. Stability of VGAP architecture (the virtual grid) allows it to achieve high throughout. Figure 6 shows the packet delivery percentage using M-OSPF in VGAP and employing the link state advertisement model that uses two threshold levels. The thresholds (L_{th}, U_{th}) are changed from (0.1,0.9) to (0.5,0.5), respectively. VGAP is able to achieve high packet delivery ratio and maintain acceptable levels when the link updates increase (increase in control traffic).

End-to-end Delay: The end-to-end packet delay is also studied. When a link is unreliable, the node fails to forward packets, causing packet drops or longer delays. At low traffic load, nodes rarely experience congestion but often experience broken links. Therefore, packets that need to be re-routed will be queued and therefore encounter longer delays. The packet end-to-end delay was set to 10 msec. VGAP was able to satisfy the delay requirements for most of these packets as shown in Figure 7. The figure shows the cumulative probability distribution of the end-to-end delay of the generated packets in the network. We notice that most packets were delivered successfully within there required delay bounds.

Call Acceptance Ratio: Call acceptance ratio is defined as the number of successfully accepted route requests divided by the total number of requests generated in

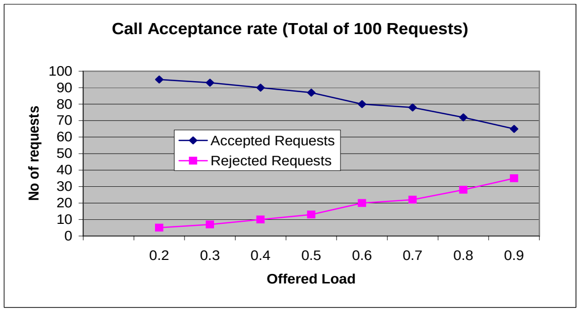


Figure 8. Call acceptance rate in VGAP.

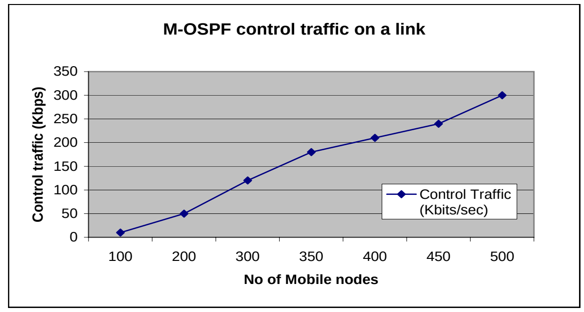


Figure 10. Control traffic of M-OSPF.

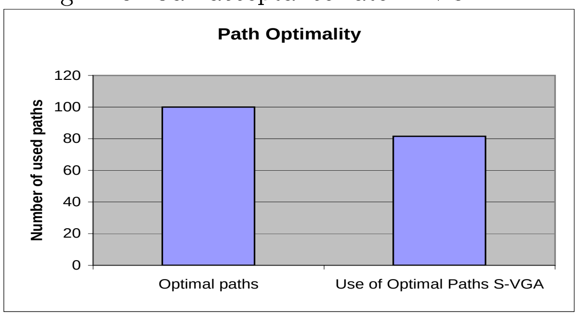


Figure 9. Use of optimal paths in VGAP.

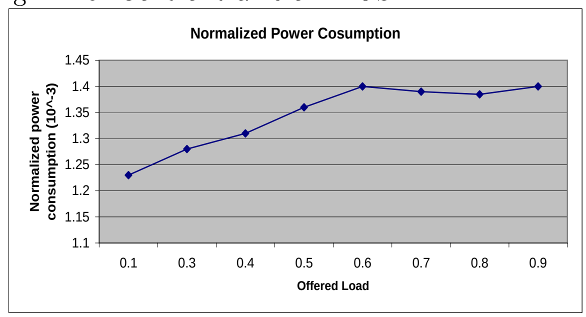


Figure 11. Normalized power consumption.

the network for the length of the simulation period. Figure 8 shows the call acceptance ratio versus the network offered load. Since VGAP finds routes that exhibit longer lifetime due to its stable architecture, it is expected that the call acceptance ratio will be high. This is clear in Figure 8 where we show the number of the accepted and rejected calls versus the offered load for every instance of the offered load. As the offered load increases, the call acceptance rate drops with a reasonable rate. For higher levels of offered loads, VGAP still accepts more than 65% of the calls.

Path Optimality: The path with the minimum number of hops which still satisfies the required QoS metrics is an optimal path since it uses less resources. The number of optimal paths used in the simulated networks as a percentage of the set of all paths used for routing is shown in Figure 9. VGAP which employs M-OSPF was able to use 80% of the optimal paths when serving requests. This efficient use of the resources also explains why VGAP was able to achieve high packet delivery ratio.

Scalability: Scalability is an important factor that refers to the adaptability of the protocol to larger networks, in terms of area and number of nodes. To show the scalability of VGAP, we vary the number of nodes from 100 to 500 and measure the control traffic generated by M-OSPF. Figure 10 shows that the control traffic increases almost linearly with the increase in the number of the mobile nodes. This is because the resulting virtual topology uses less control traffic and the control messages generated by VGAP is triggered when a change to this fixed virtual topology occurs. Note that as long as there is a single node in any zone already in the virtual topology, there is no change to virtual topology.

Normalized Power Consumption: This is defined as the total consumed energy divided by the number of delivered packets. We measure the power consumption because it is a precious resource in mobile communications. Figure 11 shows the normalized power consumption in VGAP. The power consumption increases gradually with increasing the load, but maintains comparable levels. This is a result of the power control scheme and the zoning process used of VGAP.

6. *Conclusion*

We have presented VGAP, a QoS routing scheme for heterogeneous mobile ad hoc networks. VGAP is simple, scalable, and reliable. The proposed protocol utilizes the interaction of the bottom three layers of MANETs to enhance the network performance and support of quality of service requests. The protocol overlays a regular rectilinear virtual topology on the physical topology using a simple zoning approach that takes into consideration the transmission range difference in a heterogeneous MANETs. For simplicity, we assumed two transmission ranges (Long range and short range) in this paper. However, the proposed scheme works for more than two transmission ranges with simple extensions. VGAP runs an extended version of the OSPF (M-OSPF) on the virtual grid to compute QoS routes. In order to provide QoS guarantees, M-OSPF employs an extended version of WFQ scheduling policy (called AWFQ) that takes wireless channel status into account when estimating link costs in terms of link bandwidth and link delay. Our simulation results show improved performance in terms of bandwidth and end-to-end delay guarantees and more network resilience to link failure and topology variations. The protocol also employs a power control algorithm at the physical layer to provide communication between nodes and maintain network connectivity in this heterogeneous setting.

REFERENCES

1. H. Xiaoyan, X. Kaixin, M. Gerla,"Scalable routing protocols for mobile ad hoc networks", IEEE Network, Volume: 16 Issue: 4 , July-Aug. 2002 Page(s): 11-21.
2. D.B. Johnson and D.A. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks," Mobile Computing, vol. 353, T. Imielinski and H. Korth, eds., Kluwer Academic, Boston, 1996, Page(s): 153-181.
3. M.R. Pearlman, and Z.J.Haas,"Determining the Optimal Configuration for the Zone Routing Protocol," IEEE Journal on Selected Areas in Comm., Volume 17 No. 8 , Aug. 1999, pp. 1395-1414.
4. Y.-B. Ko and N. H. Vaidya,"Location-Aided Routing (LAR) for Mobile Ad Hoc Networks", MOBICOM'98, October 1998, Dallas.
5. S. Capkun, M. Hamdi, J. Hubaux,"GPS-free positioning in mobile ad-hoc networks", Proceedings of the 34th Annual Hawaii International Conference on System Sciences, 2001 Page(s): 3481-3490.
6. T. Camp, J. Boleng, V. Davies,"A Survey of Mobility Models for Ad Hoc Network Research", Wireless Communications & Mobile Computing (WCMC): Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications, 2002. Page(s) 332-347.
7. I. Stojmenovic, "Position-based routing in ad hoc networks", IEEE Communications Magazine, July 2002, pp 128-134.
8. W. Liao, Y. Tseng, J.P. Sheu,"GRID: A fully location-aware routing protocols for mobile ad hoc networks", Telecommunication Systems, volume 18:1-3,pp 37-60, 2001.
9. S. Basagni, I. Chlamtac, V. Syrotiuk, "Dynamic source routing for ad hoc networks using the global positioning system", IEEE Wireless Communications and Networking Conference, 1999 (1), Page(s): 301-305.
10. J. Chen; J. Wang; S. Deng; Y. Tang;"QoS routing with mobility prediction in MANET", IEEE PACRIM. 2001, Vol. 2 , 2001 Page(s): 357-360.
11. J. Wu and H. LI,"A Dominating-Set-Based Routing Scheme in Ad Hoc Wireless Networks", Telecommunication Systems 18:13, 1336, 2001.
12. D. J. Baker, A. Ephremides, and J. A. Flynn, "The design and simulation of a mobile radio network with distributed control," IEEE J. Select. Areas Commun., vol. SAC-2, pp. 226237, Jan. 1984.
13. M.S. Corson and A. Ephremides, "A distributed routing algorithm for mobile wireless networks", Wireless Networks 1 (1995) 61-81.
14. M. Joa-Ng and I.-T. Lu, "A peer-to-peer zone-based two-level link state routing for mobile ad hoc networks", IEEE Journal on Selected Areas in Communications 17(8) (1999) 1415-1425.
15. E.M. Royer and C.-K. Toh,"A review of current routing protocols for ad hoc mobile wireless networks", IEEE Personal Communications (April 1999) 46-55.
16. P. Sinha, R. Sivakumar, V. Bharghavan, "CEDAR: a core-extraction distributed ad hoc routing algorithm". In IEEE Infocom99, New York, March 1999.
17. S. Chen and K. Nahrstedt, "Distributed Quality-of-Service Routing in Ad Hoc Networks," IEEE JSAC, vol. 17, no. 8, Aug. 1999, pp. 1488-1505.
18. K. Xu, X. Hong, and M. Gerla.,"An Ad Hoc Network with Mobile Backbones", In Proceedings of IEEE International Conference on Communications (ICC 2002), New York, NY, April 2002.
19. J. Wu, and H. Li.,"A Dominating-Set-Based Routing Scheme in Ad Hoc Wireless Networks", Wireless Networks in the Telecommunication Systems, 3:63-84, 2001.
20. D. Ayyagari, A. Michail, A. Ephremides,"A unified approach to scheduling, access control and routing for ad-hoc wireless networks", VTC 2000-Spring Tokyo. 2000 , Volume: 1 , 2000 Page(s): 380-384.
21. J. Broch, D. Maltz, and D. Johnston,"Supporting Heirarchy and Heterogenous Interfaces in Multi-Hop Wireless Ad Hoc Networks", IEEE proceedings of the workshop on Mobile Computing, I-SPAN, June 1999.
22. Y. Wu et.al., "Spreading code assignment in an ad hoc DS-CDMA wireless network", ICC 2002. Volume: 5 , 2002 Page(s): 3066 -3070 vol.5
23. A. Parekh, R. Gallager,"A Generalized Processor Sharing Approach to flow control in Integrated Services Networks - The Multiple Node Case," IEEE/ACM Transactions on Networking, Vol 2 No.1, pp 137-150, April 1994.
24. Q. Ma and P. Steenkiste,"Quality-of-service routing for traffic with performance guarantees", IFIP International Workshop on Quality of Service, 1997.
25. J. Moy. "OSPF Version 2". INTERNET-RFC 2328, April 1998.
26. F. Baker, "OSPF v3 as a MANET routing protocol" presentation, http://www.ietf.org/proceedings/
27. S. Lu, V. Bharghavan, and R. Srikant,"Fair queueing in wireless packet networks", In Proceedings of the ACM SIGCOMM Conference (SIGCOMM-97), pages 6376, New York, Sept. 1418 1997. ACM Press.
28. S. Bucheli, J.R. Moorman, J.W. Lockwood and S. M. Kang, "Compensation modeling for QoS support on a wireless network", ; Global Telecommunications Conference, 2000. GLOBECOM '00. IEEE , Volume: 1 , 2000 Page(s): 198-202 vol. 1.
29. C. Zhu, M.S. Corson,"QoS routing for mobile ad hoc networks", INFOCOM 2002. Volume: 2 , 2002 Page(s): 958-967.
30. S.H. Shah, K. Nahrstedt,"Predictive location-based QoS routing in mobile ad hoc networks", IEEE ICC 2002, Volume: 2, Page(s): 1022-1027.
31. J. N. Al-Karaki, A. E. Kamal,"Virtual topology based routing in infrastructreless wireless networks", Technical Report, Iowa State University, 2003.
32. D. Goodman, "Wireless Personal Communications Systems". Reading, MA: Addison-Wesley, 1997.
33. The NS-2 simulator, http://www.isi.edu/nsnam/ns