Efficient Virtual-Backbone Routing in Mobile Ad Hoc Networks*

Jamal N. Al-Karaki † Ahmed E. Kamal‡
Dept. of Electrical and Computer Engineering
Iowa State University, Ames, IA 50011
Email: *jkaraki@hu.edu.jo, kamal@iastate.edu*

Abstract

Since the physical topology of Mobile Ad hoc Networks (MANETs) is generally unstable, an appealing approach is the construction of a stable and robust virtual topology or backbone. A virtual backbone can play important roles related to routing and connectivity management. In this paper, the problem of providing such a virtual backbone with low overhead is investigated. In particular, we propose an approach, called Virtual Grid Architecture (VGA), that can be applied to both homogeneous and heterogeneous MANETs. We study the performance tradeoffs between the VGA clustering approach and an optimal clustering based on an Integer Linear Program (ILP) formulation. Many properties of the VGA clustering approach, e.g., VGA size, route length over VGA, and clustering overhead are also studied and quantified. Analytical as well as simulation results show that average route length over VGA and VGA cardinality tend to be close to optimal. The results also show that the overhead of creating and maintaining VGA is greatly reduced, and thus the routing performance is improved significantly. To illustrate, two hierarchical routing techniques that operate on top of VGA are presented and evaluated. Performance evaluation shows that the VGA clustering approach, albeit simple, is able to provide more stable (long lifetime) routes, deliver more packets, and accept more calls.

1 Introduction

Mobile Ad hoc Networks (MANETs) are distinguished from other communication networks by many features such as the absence of a fixed infrastructure, dynamic topology due to continuous node movement, wireless multihop communication, and strict resource limitations (e.g., limited bandwidth and energy). The dynamic nature of MANETs' topology may cause frequent route breakage while the ephemeral node associations may limit the link lifetime or the route lifetime. Therefore, the design of topology control algorithms and routing protocols is a crucial and challenging problem in MANETs [14]. In recent years, many researchers have strived to design routing and topology control[1] algorithms and protocols for MANETs [1]-[14]. Although MANETs have no physical backbone or infrastructure, an alternative for a physical backbone is the construction of a virtual backbone or infrastructure.

One way to create infrastructure in MANETs is to perform node clustering. In general, clustering approaches incur lower overhead during topology state updates and also have fast convergence [1]. Moreover,

*This research was supported in part by grants ECCS-0601570 and CNS-0626822 from the National Science Foundation.
†J. N. Al-Karaki was with Iowa State University. He is now with the Hashemite University, Zarqa, Jordan.
‡Corresponding author; Tel:(515) 294-3580, FAX:(515) 294-1152, email: kamal@iastate.edu.
[1]In this paper, topology control, clustering, and virtual backbone construction are used interchangeably.

clustering can enable bandwidth reuse by reducing interference, and thus can increase system capacity [4]. The Minimum Dominating Set (MDS) problem and the relevant Minimum Connected Dominating Set (MCDS) problem best describe the clustering approach to topology management. However, finding the minimal set of clusters is an NP-Complete problem [36] even when the complete network topology is available. Therefore, most approaches are based on heuristic solutions that provide a sub-optimum solution of the MCDS problem. In fact, it was shown that any distributed algorithm for constructing a connected set of clusters (sometimes called virtual backbone) requires at least $O(nlogn)$ messages, where n is the number of nodes. As such, the tradeoff between simplicity and optimality plays a critical role in the clustering approaches in MANETs. To be more specific, having a virtual architecture that is simple, stable, and scalable, while being close to optimal can be a reasonable compromise.

The concept of virtual backbones is not new. Early studies on this issue appeared in [1][6][7] and [14]. However, the authors of these studies do not attempt to optimize the size of the virtual backbones. Following the early efforts, clustering or topology control in MANETs received a considerable amount of research [18]-[32]. Linked with topology control is the routing problem in MANETs. In recent years, numerous routing protocols were developed (a detailed survey about these routing protocols can be found in [1]). However, most of these studies focused on homogeneous MANETs where nodes are assumed to be of comparable capability. However, MANETs are typically heterogeneous networks with various types of mobile nodes which need to cooperate. As a result, node heterogeneity can affect communication performance and the design of communication protocols in MANETs. Moreover, it has been found that the construction of a stable virtual backbone will help to perform routing more efficiently [19].

In this paper, we realize that most of the routing problems in MANETs are partly due to the unstable topology caused by random deployment and random movement of mobile nodes. Therefore, having a fixed, stable, and connected virtual backbone can simplify those routing problems. The contribution of this paper is two-fold. First, we address the issue of topology control in both homogeneous and heterogeneous MANETs by developing a simple, fixed, and scalable virtual wireless backbone. The fixed backbone, called the Virtual Grid Architecture (VGA), is created through a novel and simple tessellation (zoning) scheme that makes use of the nodes' locations in the network area. The tessellation scheme maps the network physical topology onto a virtual grid (rectilinear) topology. A simple power control scheme, which captures differences in nodes' transmission power, is used in conjunction with the zoning process to ensure network connectivity especially in heterogeneous MANETs. VGA is used as the underlying backbone for simple and efficient routing in MANETs. A comprehensive analysis of the VGA clustering approach is presented in this paper. The performance of VGA is studied both analytically and through simulations. More precisely, the tradeoffs (in terms of computational complexity and solution quality) between VGA clustering approach and an optimal clustering approach are analytically studied based on an Integer Linear Program (ILP) formulation. For large homogeneous MANETs, we derive expressions for VGA clusterhead (CH) cardinality, VGA worst case path length, and VGA average case path length. We also show mathematically why rectilinear routing is better than diagonal routing over VGA. For arbitrary-connected heterogeneous MANETs with a small to medium number of users, we find the optimal number of connected CHs using our developed ILP. We also derive expressions for the communication overhead over VGA. Both analytical and simulation results show that our proposed clustering algorithm VGA, although

simple, is close to optimal and it performs well in terms of average route length, CH cardinality, and the communication overhead. The overhead of computing and refreshing the virtual backbone for the entire network is greatly reduced, and thus the routing performance is improved significantly. To demonstrate, we develop two routing protocols that operate on top of VGA, namely, on-demand based and transitive closure based routing techniques. Both routing protocols affirm the stability of VGA. In other words, the discovered routes over VGA exhibit lifetime that is sufficiently long to enhance both packet delivery ratio and call acceptance rate. The proposed clustering scheme combines centralized and distributed approaches and it works for small and large networks as well as for homogeneous and heterogeneous networks.

The use of virtual wireless backbone routing in MANETs is not new. VGA routing shares similarities with other hierarchical routing, as described earlier, which have existed since the beginning of packet radio network research. Thus, the novelty of this work is the combination of hierarchical and virtual backbone routing into a simple, yet effective framework, and the introduction of a novel method for distributed construction and maintenance of a virtual backbone. We also provide theoretical analysis and simulation results that prove the efficiency of the proposed framework. Moreover, our proposed clustering scheme distinguishes itself from previous schemes in several ways. First, in almost all previous schemes, variable cluster size or dynamic virtual backbone were used. Hence, complex cluster management schemes are needed to update the virtual topology in response to frequent topological changes. On the contrary, our clusters (later called zones) are fixed and the virtual topology updates are rare. This is due to the stability of our virtual topology and its enhanced connectivity capabilities as shown later in this paper. As long as each cluster is occupied by at least one mobile node, there is no need for virtual topology updates. To be specific, if the physical topology is to change due to the removal or reinstatement of a link, the virtual topology can easily be updated and managed accordingly. Second, many previous algorithms may have a mix of large and small size clusters. Hence, some CHs may become overloaded while other CHs are lightly loaded (i.e., load balancing problem). As such, a few nodes' batteries can be depleted quickly so as to potentially break the connectivity of the network. In our approach, all clusters are fixed and have the same size. Thus, the resulting architecture is fixed, simple, and scalable. Finally, most of the previous clustering algorithms consider the case of only homogeneous nodes where nodes are assumed identical, while our clustering approach considers both homogeneous as well as the more realistic heterogeneous MANETs where the capabilities (e.g., transmission power) of mobile nodes may be different for different nodes within the same network.

The rest of the paper is organized as follows. In Section 2, we present the system model, notation, and some definitions. The VGA clustering approach is presented in detail in Section 3. In Section 4, we discuss performance issues of VGA clustering in large homogeneous networks, and derive some useful performance results. We also analyze the communication overhead of VGA clustering and compare it to other algorithms in the literature. We then present an exact solution for the clustering problem (i.e., finding the minimal set of CHs) in general heterogeneous MANETs through an ILP formulation, which is used for comparison purposes. In Section 5, we present two efficient routing techniques over VGA. Simulation results are presented in Section 6. Finally, we conclude the paper with a few remarks in Section 7.

Notation	Meaning
A	The network service area (m^2).
N	The number of mobile nodes in the network area A.
Z	The number of possible zones in VGA architecture.
Δ	Maximum node degree in the network.
T	A window of time sufficient to capture cluster dynamics.
$Y_i(t)$	A binary indicator, which is 1 iff node i is a CH at time t.
$Q_i(t)$	The fraction of time node i remains as a clusterhead during the time window T ending at time t, $0 \leq Q_i(t) \leq 1$.
$B_i(t)$	The remaining energy in node's i battery at time t.
$C(t)$	Clusterheads cardinality at time t (size of virtual backbone).
$v_i(t)$	Average speed of node i at time t (m/s).
$E_i(t)$	The relative distance between the ith mobile node and the center of the zone at time t.

Table 1: System Model Notation

2 System Model

We consider a MANET with N mobile nodes randomly distributed in a two-dimensional network area A. All nodes are equipped with batteries with possibly different levels of energy density. If nodes differ in their transmission power levels, e.g., heterogeneous network, directional links may exist between different nodes. Let the distance between two nodes i and j be d_{ij}. The condition for the existence of a link from node i to node j ($i \rightarrow j$) is given by ($d_{ij} \leq r_i$) where r_i is the transmission radius of node i and is given by the following equation [37]:

$$r_i = \omega(\frac{P_t}{P_r})^{1/\alpha} \tag{1}$$

where ω is a constant that depends on the gain of the antennas, P_t is the transmission power between nodes i and j, P_r is the minimal received power at node j for successful decoding of the message, α characterizes the steepness of the decrease in the signal power and assumes values in the range $2 \leq \alpha \leq 4$. A link is said to be bidirectional if both $i \rightarrow j$ and $j \rightarrow i$ hold, which is the case of homogeneous networks. Mobile nodes are free to move in A, which is usually captured using a certain mobility model. Mobile nodes can find their location coordinates using a location server such as Global Positioning System (GPS) or use a GPS-free approaches based on triangulation or multilateration when the GPS card may not work. When clustering techniques are used to create a hierarchy in the network topology, a set of nodes, namely, virtual backbone or CHs, are selected to perform specific functions in the network operations. The virtual backbone is modeled in this paper as a directed graph $G(t) = (V(t), E(t))$, where $V(t)$ is the set of CH nodes and $E(t)$ is the set of directed wireless links connecting CH nodes at time t. Two vertices (CHs) are directly connected by a link in $G(t)$ if they satisfy transmission range restrictions. A mobile node communicates with its CH using the IEEE 802.11 MAC protocol mechanism, while channel communication management between CHs is controlled by using code division multiple access (CDMA). The assignments of codes to CHs is according to [33], where neighboring CHs use different spreading codes in order to reduce interference and to enhance spatial reuse of channels. For ease of reference, Table 1 summarizes notation used in this paper. We now introduce some useful definitions used in this paper.

Heterogeneous MANET: A MANET is considered heterogeneous if mobile nodes have different transmission ranges. When transmission ranges are the same, the network is called homogeneous.

Virtual Backbone: The set of connected clusterheads (CHs) that are elected based on certain criteria. The backbone can be fixed or variable depending on the clustering approach.

Route Lifetime: A route or path lifetime is defined as the time until a node or a link in this route fails. To find the lifetime of a route, we must consider all the links in the route because a break in any of the links break the route.

Clusterhead Cardinality: Let $C(t)$ be the CH cardinality of a virtual backbone at time t. The value of $C(t)$ at time t can be defined as:

$$C(t) = \sum_{j=1}^{N} Y_j(t) \tag{2}$$

Network Stability: A MANET is considered stable if the frequent topological changes can be made transparent to the users.

3 The VGA Clustering Approach

In this section, we present the Virtual Grid Architecture (VGA) clustering approach in detail. The main objective of the VGA clustering approach is to create a simple and stable rectilinear virtual topology on which the routing and network management functions can be performed easily and efficiently. Our clustering approach consists of two major steps: network zoning and CH election inside zones.

A. Network Zoning:

The zoning strategy starts by dividing the network area into disjoint, adjacent, fixed size, and regular (symmetric) shaped zones. To create a simple rectilinear virtual topology, we select the zones to be square in shape with possible extension to other virtual topologies like hexagon, line, or triangle with some corresponding tradeoffs (GAF [17] and [9] used similar shapes, but with no analysis). Since zones are fixed, sophisticated algorithms for zone management are not needed. Formally, the network area, defined by the coordinates $(0.0, 0.0) - (MAX_x, MAX_y)$ is divided into fixed-size and equal square zones. Let Z be the set of such zones. Let $z_k \in Z$ be the k^{th} zone, and denote its initial (bottom left) and final (top right) coordinates by $(x_i^{(k)}, y_i^{(k)})$ and $(x_f^{(k)}, y_f^{(k)})$, respectively. Each zone z_k has a unique address (or ZoneID) which can be the ordered pair consisting of the zone's row and column numbers, respectively. Each mobile node is a member of one zone and its zone membership is determined based on its location in the network area. When GPS is used, the node coordinates will be absolute, while it will be relative to other nodes in GPS-free approaches. By knowing its coordinates, a node can calculate its ZoneID easily. We distinguish the zoning process between the two cases of homogeneous and heterogeneous MANETs as follows:

1. *Homogeneous MANETs:* In this case, the zone side length x is chosen such that two mobile nodes in adjacent horizontal/vertical zones and located *anywhere* in their zones, can communicate with each other directly. A necessary and sufficient condition for this communication to happen is to set the zone side length (x) to $\frac{r}{\sqrt{5}}$ (see Figure 1). If communication between diagonally adjacent

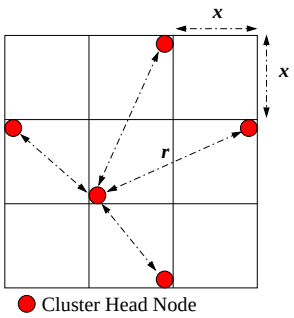

Figure 1: Selection of zone side length.

zones[2] is allowed, the zone size will be smaller since $x = \frac{r}{2\sqrt{2}}$, hence increasing average route length and routing cost as will be shown later. An example of the fixed zoning process applied to homogeneous networks as well as the resulting virtual topology is shown in Figure 2(a).

2. *Heterogeneous MANETs:* For the sake of simplicity, but without loss of generality, let us assume that mobile nodes are of two different transmission ranges: r_s for Short Range (SR) nodes and r_l for Long Range (LR) nodes, where $r_s < r_l$. The transmission range is a function of the transmission power at the mobile node in each case. We want to provide two types of communication:

   - intra-zone communication, which refers to communication between nodes inside a zone, and
   - inter-zone communication, which refers to communication between nodes in neighboring zones.

   Our objective is to provide these two types of communication while using the minimum levels of transmission power. Note that a zone $z \in Z$ may have short, long, or both long and short transmission ranges nodes. Hence, we adopt the following zoning strategy. Initially, the network area is divided into large zones with a zone side length $x_l = \frac{r_l}{\sqrt{5}}$. When a zone has short range nodes only, then it is divided into four[3] *subzones* with the short range, r_s, chosen such that each subzone side length is set to $x_s = \frac{x_l}{2} = \frac{r_l}{2\sqrt{5}}$. Figure 2(b) shows an example of the zoning process in heterogeneous networks and the resulting virtual topology. In Figure 2(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.

   Note that for heterogeneous MANETs, the topology can be changed by varying the nodes' transmitting range. So, further energy can be saved if the virtual network topology used to route/broadcast messages is energy-efficient itself. The goal of topology control is to dynamically change the nodes transmitting range in order to maintain connectivity, while reducing the energy consumed by nodes' transceivers (which is strictly related to the transmitting range) [31]. We now determine the power levels required for inter-zone and intra-zone communications. Since we have two transmission ranges, four types of communication between mobile nodes are defined

[2]Which is referred to as diagonal communication, or diagonal routing in this paper.

[3]In general, if we have m transmission levels, then larger zones will be divided into subzones using a binary strategy, i.e., $x^{(k)} = \frac{x_l}{2^k}, k = 1, 2, 3, \ldots, m$ going from large to small zones. A node joins zone k if its transmission range, r, satisfies $\sqrt{5}x^{(k)} \leq r < \sqrt{5}x^{(k-1)}$.

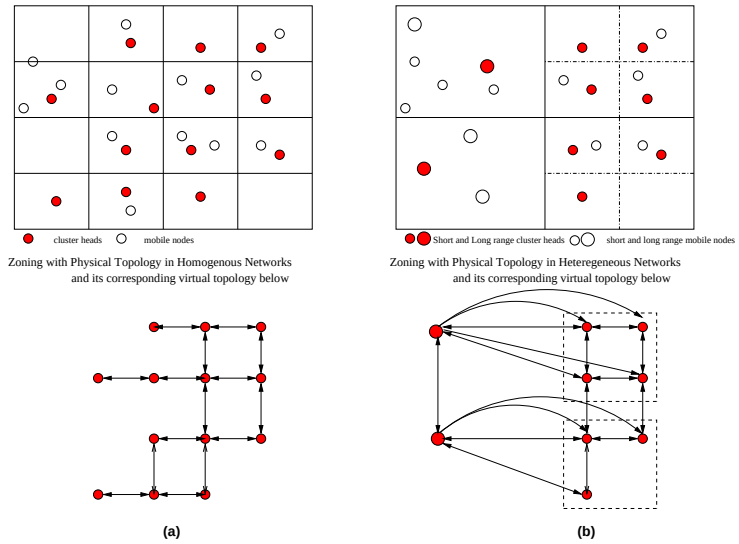

Figure 2: The Fixed zoning process in both (a) Homogeneous networks and (b) Heterogeneous networks.

(Short→Short(S-S), Long→Long(L-L), Long→Short(L-S), and Short→Long(S-L)). We need to ensure that all nodes, LR and SR, will be able to communicate. In general, the transmitted signal strength is known to decrease with distance according to equation (1). Assuming $r_i = d$ and re-arranging equation (1), we have $P_r = \omega \frac{P_t}{d^\alpha}$. To find the relative power consumption, define the ratio of the transmitted signal power of any mobile node to that of LR node (RTP) as follows:

$$RTP = \frac{P_t}{P_{max}} = (\frac{r}{r_l})^\alpha \tag{3}$$

where P_t and P_{max} are the transmitted signal power of a target mobile node and LR node, respectively; and $r = r_s$ or r_l for SR nodes and LR nodes, respectively. Figure 3 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 3, we find the values of RTP for different types of communication. The results are summarized in Table 2 for different ranges of RTP. The values between parentheses are the values of RTP corresponding to each type of communication when α=2. They can be obtained through simple derivations. Note that a LR node can use a reduced power P_{red} to communicate with another 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 the transmission power can be reduced by a factor of 6, hence prolonging the lifetime of the batteries of mobile nodes. It is to be noted that the lifetime extension will be less than six times in this case, since power consumption is a function of other parameters which are independent of the communication distance.

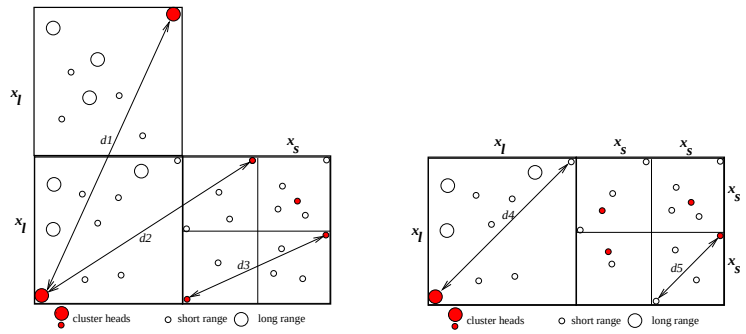

Figure 3: (a) Inter-zone communication (b) Intra-zone communication in VGA.

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)	$\frac{r_l}{2}(0.25)$	$0.81r_l(0.65)$	$\frac{3}{2\sqrt{5}}r_l(0.4)$	$r_l(1.0)$

Since nodes are mobile, it may happen that the existence of LR nodes, SR nodes, or both will vary inside zones according to these movements and therefore zones may consequently get split or merged. Consider the following scenario (see Figure 4). When the only LR node (node A in Figure 4) leaves its zone (zone (0,1)) to a neighboring zone (zone (1,1)) that only has SR nodes, then a split operation takes place in the old zone (four subzones with a SR CH node selected inside each subzone) and a merge process will occur in the new zone of the LR node. In general, split and merge processes take place depending on the situation imposed by node movements. It is important to note that in homogeneous MANETs, the split and the merge processes are not necessary since all zones are fixed and retain the same size regardless of node movement.

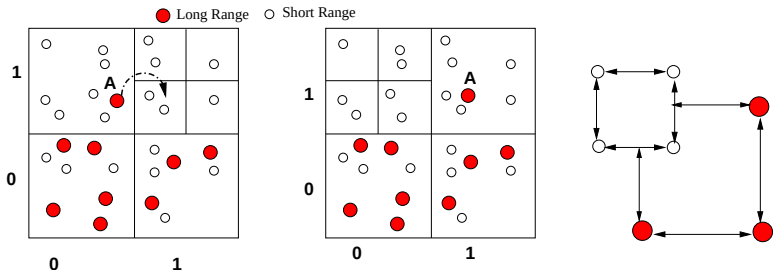

Figure 4: Example of the split and merge processes in heterogeneous MANETs.

B. CH Election:

After the zoning operation is finished, a CH election algorithm is executed in each zone. Nodes in each zone gain network access through the elected CH only. The CH role itself does not change, but CH does change inside each zone. In the following, we discuss three different scenarios for CH change, namely, periodic, adaptive, and on-demand. In all scenarios, the CH election algorithm is performed concurrently in all network zones.

Periodic CH election: The approach used here is not new, but its operation is distinguished from others [3]. In each zone, nodes take turn in acting as CH for a certain period. The most eligible mobile node will become a CH for the current period based on an eligibility factor (EF). The objective of EF is to select a node to act as CH for the longest time possible in order to stabilize the virtual backbone. Each node i will calculate its eligibility factor, EF_i, at time t as:

$$EF_i(t) = a_1 e^{-v_i(t)} + a_2(1 - Q_i(t)) + a_3 B_i(t) + a_4(1 - E_i(t)) \tag{4}$$

where $v_i(t)$ is the velocity of node i at time t, $B_i(t)$ is the remaining energy at ith node battery at time t, $Q_i(t)$ is the fraction of time node i served as CH in a large window of time that reflects the energy level difference between the previous period and the current period of that node, $E_i(t)$ is the distance from node i to the center of its zone relative to the maximum distance within the zone at time t, and a_1, a_2, a_3, a_4 are scaling factors. The node that has the highest value of EF will elect itself as the CH in its zone for the current period. That is, if a zone $z \in Z$ has a nonempty set of m nodes, the CH is selected as follows:

$$CH = arg \max_{i \in m}(EF_i) \tag{5}$$

A simple algorithm that performs CH election in zone z is shown in Figure 5. The algorithm is distributed at each node and it is executed in each zone independently. In the algorithm, each node in zone z calculates its EF and broadcasts it to other nodes in the same zone. A simple comparison of EF value at each node determines which node has the maximum EF, i.e., the node that is the most eligible to become a CH. This node will declare itself as the CH and broadcast the zoneID, and its own node ID to other members, using a CH claim message. To save on transmissions, collisions, and also save time, each node, i, determines the transmission time τ_i of its EF_i value, where $\tau_i = \alpha / EF_i$, and α is a constant, and is chosen to determine how long the nodes wait. A node with a high EF value will therefore transmit earlier, hence nodes with lower EF values will not transmit. Nodes' transmissions with the same EF value will collide, and in this case the collision is resolved using random backoff times. Note that the algorithm requires an acknowledgment of successful transmission of the EF value. This can be provided easily by nodes which receive the successful transmission. Notice that even if multiple acknowledgments are transmitted and collide, they can be interpreted as an acknowledgment.

In heterogeneous settings, the broadcast message also includes a field for the node type (SR or LR) that is used for the split and merge processes as was indicated earlier. Following every new election, the new CH inherits the routing cache of the retired CH. However, in order to minimize packet delays incurred due to CH role exchange, we allow the retiring CH to transmit any remaining queued packets if it is not out of service. Otherwise, the new CH is responsible for routing these packets. A CH will act as such for a certain period of time. After the end of this period, another node in the zone will be elected as the CH.

It is worth mentioning that the election period can be variable or fixed. Fixing this period results in simpler operation but may not be optimal. Varying the period in an adaptive manner can result in

```
Algorithm: ClusterElection (z, i)
Input: zone z ∈ Z, m=set of nodes in z, EF_i=eligibility factor for node i
Output: A CH for current period in zone z

initialization: CH-elected=false;
if (Zone z not empty) then
    Compute τ_i = α/EF_i;
    Wait for τ_i and listen to medium;
    //If a larger EF value is received, another node is CH
    if (During τ_i, EF_j ≥ EF_i was received) then
        Interrupt wait;
        Listen for CH claim message(z, j);
        CH = j;
        CH-elected = true;
    else
        //If no larger EF value is received, try to transmit EF_i
        while (CH-elected == false) do
            Transmit EF_i;
            if (Successful) then
                Transmit CH claim message(z, i);
                CH = i;
                CH-elected = true;
            else
                //Collision has occured with a node with a similar EF value
                Wait for random time and listen to medium;
                if (During wait, EF_j ≥ EF_i was received) then
                    Interrupt wait;
                    Listen for CH claim message (z, j);
                    CH = j;
                    CH-elected = true;
                end
            end
        end
    end
end
```

Figure 5: Cluster head election algorithm in a certain zone z

better behavior but may require more computation, overhead, and power consumption. The tradeoff between these two options is studied and the results are discussed in Section 6.

Dynamic CH election: Under this scenario, a CH in each zone is initially selected based on equation (4). Unlike periodic CH election, the eligibility factor is only recalculated in a certain zone when needed. A new CH is initiated when the current CH is about to leave its zone or its battery energy falls below a certain threshold. If all nodes have the same energy level, then the CH role will be rotated among nodes based on their IDs in ascending order. A higher ID will takeover when its descendant is out of service due to battery energy depletion.

On-demand CH election: In this strategy, a CH node may ask to be replaced on demand due to many reasons (e.g., overloaded or simply needs to switch off). This strategy can be mixed with the previous two strategies for exceptional events.

It is worth mentioning that although the concept of CH election is not a new concept, our zoning approach simplifies its operation and reduces the overhead of the whole process.

4 Properties of VGA Clustering Approach

In this section, we study the properties of the VGA clustering approach. In particular, we focus on different and important aspects of VGA, e.g., VGA size, route length over VGA, VGA clustering overhead, connectivity, and stability. We start by studying the performance tradeoffs between VGA clustering approach and an optimal clustering approach for both homogeneous and heterogeneous MANETs.

4.1 The Homogeneous MANET Case

In this subsection, we compare VGA clustering to optimal clustering in homogeneous MANETs. The optimal network coverage and the least number of fixed zones are achieved when the network area is divided into fixed and equal *hexagons* [41], such that at the center of each hexagon there is a node acting as a clusterhead. This is very restrictive and can only be guaranteed with a very large number of nodes, which are uniformly distributed over the network area. The likelihood of this happening in MANETs is small, but will be used here for reference. We present our analytical results along three aspects: (1) Justify the use of the rectilinear virtual topology and explain why diagonal routing (henceforth referred to as Diagonal VGA (D-VGA)) is not adopted in VGA, (2) Find closed form expressions for CH cardinality, (3) Find the average and worst case route length (in terms of number of hops) over VGA. Figure 6 shows the zoning strategies and the resulting virtual topologies for VGA, D-VGA, and optimal zoning, respectively.

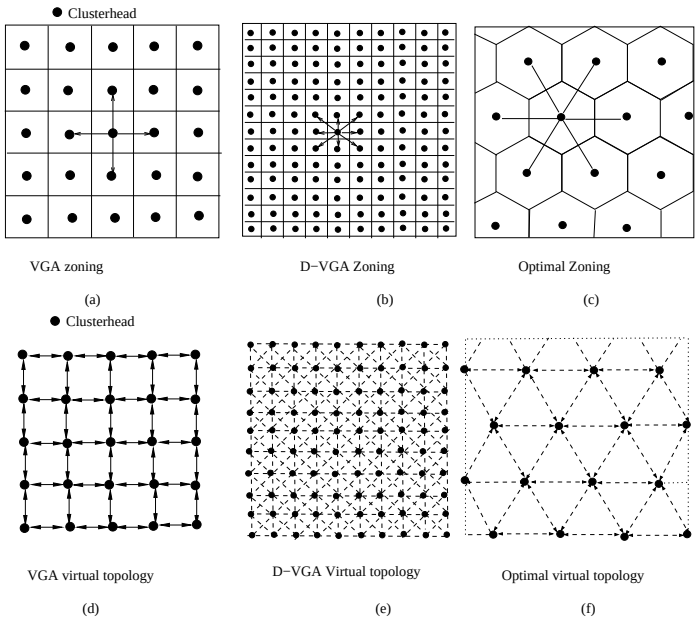

Figure 6: The zoning process with corresponding virtual topologies for VGA, D-VGA, and optimal cases .

- **Rectilinear Routing versus Diagonal Routing:** In VGA, a CH communicates only with its vertical and horizontal neighbors directly. In D-VGA, however, a CH communicates with all of its eight neighbors. The number of zones in D-VGA is 60% more than that of VGA per network area as the zone size is smaller in D-VGA given the same communication range, r. To illustrate, let the network area $A = L \times L$. In VGA, the zone side length is $x_v = \frac{r}{\sqrt{5}}$, while the zone side length of

D-VGA, $x_d = \frac{r}{\sqrt{8}}$. Hence, the number of zones in VGA is $\frac{L}{x_v} * \frac{L}{x_v} = \frac{5L^2}{r^2}$, while it is $\frac{L}{x_d} * \frac{L}{x_d} = \frac{8L^2}{r^2}$ in D-VGA. This represents an increase in the number of zones in D-VGA by a factor of 1.6. Consequently, the average number of routing hops is expected to increase in D-VGA as will be shown later in this section. As such, D-VGA has higher latency, more total power consumption, and more information processing per node than VGA.

Next, we consider the probability of diagonal routing succeeding given that rectilinear routing has failed. If this probability is low, then the selection of rectilinear routing over VGA is justified. Consider a target zone, and its eight neighboring zones under VGA as shown in Figure 7(a). Let B be the event where the target zone has at least one node, the neighboring vertical and horizontal zones are empty, and at least one diagonal zone is not empty. Also, assume that the total number of nodes in the network area is N, and that there are $|Z|$ zones in the network area. Let $P(B)$ be the probability of the occurrence of event B (i.e., the rectilinear routing from the central zone fails while the diagonal routing succeeds). We express $P(B)$ as:

$$P(B) = \sum_{n=1}^{N} Pr(B|n) \binom{N}{n} (\frac{|Z|-9}{|Z|-1})^{N-n} (\frac{8}{|Z|-1})^{n} \tag{6}$$

where $Pr(B|n)$ is the probability of the event B given that n nodes are uniformly distributed in the eight zones surrounding the target zone and is given by $\frac{\binom{n+3}{3}}{\binom{n+7}{7}}$. The expression $\binom{N}{n}(\frac{|Z|-9}{|Z|-1})^{N-n}(\frac{8}{|Z|-1})^{n}$ represents the probability that out of the total N nodes in the network area, 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 7(b).

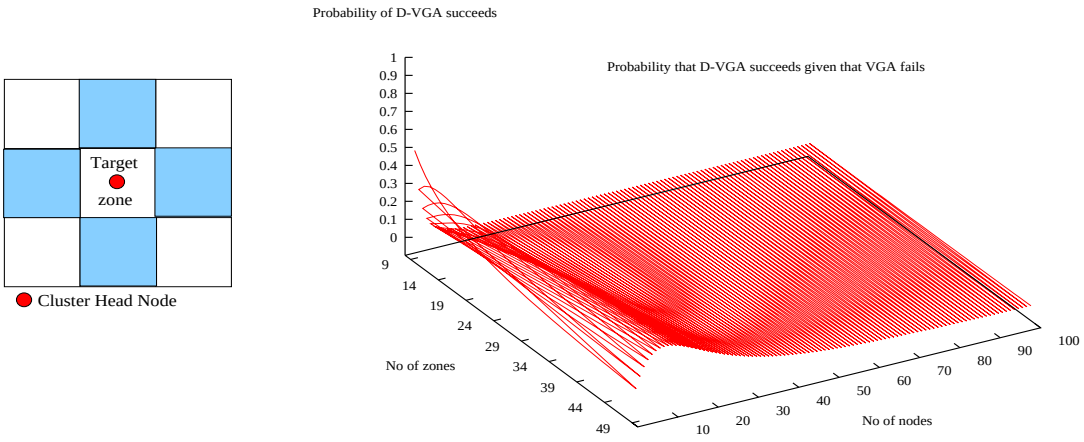

Figure 7: (a) Diagonal routing capability option (b) Probability that D-VGA succeeds given that VGA fails.

As shown in the figure, the probability of successful diagonal routing and rectilinear routing failure is very low. For example, when the number of zones $|Z|$=40 and the number of nodes N=80, the value of $P(B)$ is 0.0015749. Other values of $|Z|$ show the same trend. Hence, adding diagonal

routing does not significantly improve the probability of finding a route in the case rectilinear routing fails. This justifies our exclusion of diagonal routing from VGA. We used the assumption of random node distribution for the purpose of analytically comparing between rectilinear routing and diagonal routing. This was done for the sake of mathematical tractability, and we did not assume any specific mobility model.

- **Clusterheads Cardinality:** We define *the cardinality* of the virtual topology, VGA, as the number of elected CHs. The CH cardinality, $C(t)$, of the virtual topology at time t can be expressed as:

$$C(t) = \sum_{j=1}^{N} Y_j(t)$$

where $Y_j(t)$ is 1 if and only if node j is a CH at time t. We find C [4] for VGA, D-VGA, and for optimal clustering as follows. It is known that the area, R, of a hexagon (which is used in optimal clustering) with side length x is given by $R = \frac{3\sqrt{3}}{2}x^2$. Two CHs located at the centers of two neighboring hexagons will be able to communicate when $x = \frac{r}{\sqrt{3}}$; hence the area of the hexagon can be written as $R = \frac{\sqrt{3}}{2}r^2$. For VGA and D-VGA, the zone sizes are given by, $R = x^2 = (\frac{r}{\sqrt{5}})^2 = \frac{r^2}{5}$ and $R = x^2 = (\frac{r}{2\sqrt{2}})^2 = \frac{r^2}{8}$, respectively. Assuming a large network area of size $L \times L$ square meters, the number of clusters C, (which is also the expected CH cardinality) can be found easily as:

$$\text{VGA: } C = \frac{5L^2}{r^2}\ ,\ \text{D-VGA: } C = \frac{8L^2}{r^2},\ \text{Optimal: } C = \frac{2L^2}{\sqrt{3}r^2}$$

Table 3 summarizes the results and also shows the ratio between the number of clusters in VGA and D-VGA and the optimal number of clusters. We note from the table that VGA has a smaller CH cardinality than D-VGA. Also, for the case of hexagons, if nodes are allowed to be anywhere inside the hexagon, then the CH cardinality ratio of VGA to optimal zoning would be even smaller.

Table 3: Comparison of clusterhead cardinality for VGA, D-VGA, and optimal clustering.

Case	Optimal	VGA	D-VGA
Number of clusters(S)	$\frac{2L^2}{\sqrt{3}r^2}$	$\frac{5L^2}{r^2}$	$\frac{8L^2}{r^2}$
CH cardinality Ratio	1.0	4.33	6.93

- **The Worst-case and Average-case Path Length:** we now show that the average path length of VGA is shorter than the average path length of D-VGA. Also, we show that although VGA is simple, the average path length over VGA is closer to the optimal average path length than D-VGA. On the other hand, the worst case path length of VGA is higher than that of D-VGA. We state these facts formally in the following proposition:

Proposition 1: If the network area is large and equal to $L \times L$ unit area, mobile nodes are uniformly distributed in the network area, and each node has a transmission range of r meters, then the worst

[4] From this point onward, we drop the time index for convenience, e.g., we use C instead of $C(t)$.

case path length (PL) of VGA, D-VGA, and Optimal routing (in terms of the number of hops) is given by:

$$PL^{max}_{VGA} = \frac{2L\sqrt{5}}{r} - 1 \ , \ PL^{max}_{D-VGA} = \frac{L\sqrt{8}}{r} \ , \ PL^{max}_{Opt} = \frac{L}{r}(\frac{1+\sqrt{3}}{\sqrt{3}}),$$

while the average case path length is given by:

$$PL^{avg}_{VGA} = \frac{\sqrt{5}L}{r} \ , \ PL^{avg}_{D-VGA} = \frac{5\sqrt{2}L}{3r} \ , \ PL^{avg}_{Opt} = \frac{(3\sqrt{3}+11)L}{18r}.$$

Proof. See the Appendix. □

The increase in D-VGA average path length can be intuitively justified by observing that in D-VGA, zones will be smaller in size. As such, the same network area in D-VGA will contain more zones than VGA, which leads to longer routes in terms of the number of hops. The results of our simulations and a simulation study carried in [9] are in agreement with our derived equations where the simulation results show that the average path length of VGA is smaller than D-VGA.

4.2 The Heterogeneous MANET Case

In this section, we consider the general problem of finding the minimum set of connected CHs in arbitrary connected graphs. A minimal set of CHs will reduce communication overhead and simplify the connectivity management. However, finding the minimum connected virtual backbone is similar to finding the Minimum Connected Dominating Set (MCDS) of a given set of nodes in the network, which is an NP-hard problem [36]. In this section, we provide an optimal solution for MCDS problem. The problem will be defined first for homogeneous setting and then generalized to a heterogeneous setting:

Problem MCDS: Given a set of mobile nodes (ψ) distributed in the network plane, select the smallest set of CHs (M) such that:

1. For each node $i, i \in \psi$, there exists at least one CH node $j \in M$, such that $d_{ij} \leq r$.
2. The undirected graph $G = (M, E)$ induced by M (i.e. $\forall u, v \in M$, define an edge $(u, v) \in E$ if $d_{uv} \leq r$) is connected.

We formulate the problem MCDS as an Integer Linear Program (ILP). The proposed ILP works for both homogeneous as well as heterogeneous MANETs of small to medium size. First, we will outline the ILP for homogeneous MANETs and then show how the ILP can be easily extended to handle the heterogeneous case. Using the notation in Table 4, the MCDS problem can be solved by solving the ILP shown in Table 5. The objective of the ILP is to minimize the total number of CHs in the network, i.e., finding M. The set M is not known beforehand, but is determined by the ILP, otherwise, the ILP would be redundant. In the ILP, constraints (8) and (9) are for non-CHs to CH connections where constraint (8) guarantees that node i uses mobile node j as its CH, and constraint (9) guarantees that node i has a one hop connection to at least one CH. The two constraints in (10) and (11) guarantee that if x_{ij}=1, then nodes i and j are both CHs, i.e., $(x_{ij} = 1) \iff (y_{ii}{=}1) \wedge (y_{jj}{=}1)$. The constraints (12)-(17) are used for virtual backbone

Table 4: The ILP variables in heterogeneous MANETs.

Notation	Meaning
N	The number of mobile nodes in A.
M	The smallest set of CHs.
c_{ij}	Adjacency indicator, which is 1 if and only if there is an edge between i and j in G; otherwise, c_{ij} is 0. Note that we assume that the network is symmetric, and hence $c_{ij} = c_{ji}$. The values of c_{ij} are an input to the ILP.
y_{ij}	A binary indicator, which is 1 if and only if mobile node i uses j as its clusterhead. Note that y_{jj} is always 1 if and only if j is a clusterhead.
x_{ij}	A binary indicator, which is 1 if and only if both mobile nodes i and j are CHs.
r_{ij}^h	A binary indicator, which is 1 if a cluster head i can reach clusterhead j in h hops.

connections. All i and j in the constraints are indices over all nodes. Note that constraints (13) - (16) implement the following equation

$$r_{ij}^h = \bigvee_k (r_{ik}^{h-1} \wedge r_{kj}^1), \quad \forall i, j, h \; (i \neq j, k, k \neq j), 2 \leq h \leq N-1 \tag{7}$$

where constraints (13) and (14) implement the conjunction ($\wedge$) operation, while constraints (15) and (16) implement the disjunction ($\vee$) operation. These constraints, together with the minimization of the objective function, guarantee that all CHs form the MCDS. In particular, these constraints guarantee that a clusterhead i is reachable by any other clusterhead j through at most h hops, $1 \leq h < N$, where constraints (12) and (7) (which is implemented using equations (13)–(16) as indicated above) recursively find a route between two CHs i and j using other CHs as intermediate nodes, while constraint (17) makes sure that only intermediate CHs that are connected are used to find a route between CHs i and j.

The solution of the above ILP guarantees that for any arbitrary connected graph, the generated set of CHs will be minimal, connected, and provide 100% network coverage. Note that the above formulation can be extended to handle heterogeneous MANETs by not requiring the matrix of c_{ij} to be symmetric. This means that the link between two nodes may exist in one direction only, i.e., unidirectional graph. We use the results obtained from solving the ILP[5] as a baseline against which our proposed clustering algorithm (VGA) or any other related heuristic algorithm, can be compared.

4.3 Control Overhead Analysis of VGA Clustering

In this subsection, we investigate the control overhead of the VGA clustering approach and compare it to some related work in the literature. Our objective is to express the control overhead as a function of the number of mobile nodes N, the maximum node degree Δ, and the CH cardinality (C). We are also interested in finding the message overhead and the time complexity of VGA clustering approach. Note

[5]Since this problem is NP-hard, the ILP can only solve small to medium sized problems in a reasonable time. A number of approximate algorithms and heuristics for the MCDS problem have been introduced in the literature, and can be used to solve large instances of the problem.

Table 5: The ILP formulation for finding MCDS.

Objective function: $Minimize \quad \sum_i y_{ii}$ (Number of CHs in G).

Subject to:

$$\sum_j y_{ij} \geq 1\ ,\ \forall i \tag{8}$$

$$y_{ij} - y_{jj}c_{ij} \leq 0 \quad \forall i,j\ (i \neq j) \tag{9}$$

$$x_{ij} \leq \frac{y_{ii} + y_{jj}}{2} \quad \forall i,j\ (i \neq j) \tag{10}$$

$$x_{ij} \geq (y_{ii} + y_{jj} - 1) \quad \forall i,j\ (i \neq j) \tag{11}$$

$$r_{ij}^1 = x_{ij}c_{ij} \quad \forall i,j\ (i \neq j) \tag{12}$$

$$R_{ikj}^{h-1} \geq r_{ik}^{h-1} + r_{kj}^1 - 1 \quad \forall i,j,h\ (i \neq j,k,k \neq j), 2 \leq h \leq N-1 \tag{13}$$

$$R_{ikj}^{h-1} \leq \frac{r_{ik}^{h-1} + r_{kj}^1}{2} \quad \forall i,j,h\ (i \neq j,k,k \neq j), 2 \leq h \leq N-1 \tag{14}$$

$$r_{ij}^h \leq \sum_{k \neq j} \frac{R_{ijk}^{h-1}}{N-1} \quad \forall i,j, 2 \leq h \leq N-1 \tag{15}$$

$$r_{ij}^h \geq \sum_{k \neq j} R_{ijk}^{h-1} \quad \forall i,j, 2 \leq h \leq N-1 \tag{16}$$

$$\sum_{h=1}^{N-1} r_{ij}^h \geq x_{ij}\ ; \quad \forall i,j\ (i \neq j) \tag{17}$$

that for both cases of homogeneous and heterogeneous MANETs, one salient feature of our zoning process is that the resulting virtual grid limits the control overhead to the communication between CHs in the vertical and horizontal directions only. For clarity purposes, the VGA control overhead is divided into three parts: (1) Cluster formation and CH election (OH_1), (2) Cluster update or maintenance due to link addition/deletion (OH_2), and (3) Location management information due to node movements between zones (OH_3). Thus, the total communication overhead of VGA, $OH = \sum_{i=1}^{3} OH_i$.

1. Cluster formation and CHs election (OH_1); two rounds of communication are needed to elect a clusterhead (broadcast the eligibility factors and the ID of the elected clusterhead node). Since all nodes will be involved in this phase, the cluster formation overhead, OH_1, is $O(N)$.

2. Cluster update or maintenance due to link addition/deletion (OH_2): When a link state changes in VGA (addition/deletion), it will be detected by at most four neighboring CHs. Therefore, the average number of packet transmissions needed for this update is computed as follows: (each CH node directly affected) $\times$ (3 neighbors/node) $\times$ (2 message/neighbor) $\times$ (1 hop/message) $\times$ VGA cardinality (C) $= 6\ C$ transmissions. If each clusterhead maintains a global view of VGA, then the number of packet transmissions drops to (4 level-2 neighbor nodes directly affected) $\times$ (1 messages/neighbor) $\times$ (1 hop/message) $= 4$ packet transmissions, i.e., it takes on the order of $O(1)$.

3. Location management information due to node movements between zones (OH_3): when a node moves

to another zone, its location information needs to be updated. The location management overhead is O(N/Z) since we assumed uniform node distribution.

The time needed to finish the VGA clustering scheme is dependent on both the number of zones (Z) and the maximum node degree Δ in each zone. For a large network, the number of zones is comparable to the maximum node degree; hence the time complexity of running VGA is given by $O(max(\Delta, Z))$.

A set of related approximations are summarized in Table 6 in addition to our proposed architecture (VGA). It is clear from the table that VGA outperforms most other solutions.

Table 6: Performance comparison between different algorithms. Here *opt* is the size of the optimal MCDS; Δ is the maximum node degree; C is the size of the generated connected dominating set; Z is the number of zones in VGA; V,E: the number of nodes and number of edges in the virtual network graph, respectively.

	Cardinality	Message Overhead	Time Complexity	Msg Length	Hops
[6]-I	$\leq (2ln\Delta + 3)opt$	$O(VC + E + VlogV)$	$O(V + C\Delta)$	$O(\Delta)$	2-hop
[6]-II	$\leq (2ln\Delta + 2)opt$	$O(VC)$	$O(C(\Delta + C)$	$O(\Delta)$	2-hop
[11]	$--$	$O(V\Delta)$	$O(\Delta^2)$	$O(\Delta)$	2-hop
[10]	$\leq 8opt + 1$	$O(VlogV)$	$O(V\Delta)$	$O(1)$	1-hop
[12]-I,II	$\leq 8opt + 1$	$O(V)$	$O(V\Delta)$	$O(1)$	1-hop
VGA	$\leq 4.33opt$	$O(V)$	$O(max(\Delta, Z))$	$O(1)$	1-hop

5 Routing over VGA

Although VGA shares some similarity with grids in wireline networks (e.g., Manhattan Street Network) in terms of structure, the routing strategies developed for such networks, e.g., deflection routing, hot-potato routing, and vertical routing, are not directly suitable for applications in MANETs. This is because these protocols assume that all links are always available, while in MANETs links can be added/deleted over the lifetime of the network. We believe that VGA can be used to perform routing in MANETs efficiently due to its stable structure. In this section, we investigate two methods to route packets over VGA. The first method is based on the well-known *on-demand* routing techniques [16] but with a major difference: packet broadcasting is constrained to at most four neighbors. That is, our version of on-demand routing utilizes the notion of a virtual backbone to efficiently direct the route querying control traffic in a specific order. The second method is a novel approach and is based on the concept of *transitive closure* from discrete mathematics.

5.1 On-Demand Routing Approach

A simple on-demand packet forwarding scheme (OD-PFS) can be easily implemented over VGA. The OD-PFS scheme can be implemented as follows. The standard four directions (North(N), South(S), West(W), East(E)) are used for simple packet forwarding in the resulting virtual grid, VGA. Those directions can be encoded using a 2-bit representation in the packet header such that (00-11-10-01) correspond to (N-S-E-W)

directions, respectively. Note that opposite directions have complementary codes, which simplifies route traversal over VGA. OD-PFS builds routes using a route request (RREQ)/route reply (RREP) query cycle similar to the one used in the Ad hoc on-demand distance vector routing protocol (AODV) in MANETs [16]. This cycle starts by having the source send a RREQ packet which propagates until the path to the destination is found upon which a RREP packet is propagated back to the source. To illustrate the operation of OD-PFS, consider the following two examples covering the cases when the intermediate nodes do not know a path to the destination and when some of them do, respectively. For the first case, consider a certain source node (x) requesting communication to a destination node (y) as shown in Figure 8.

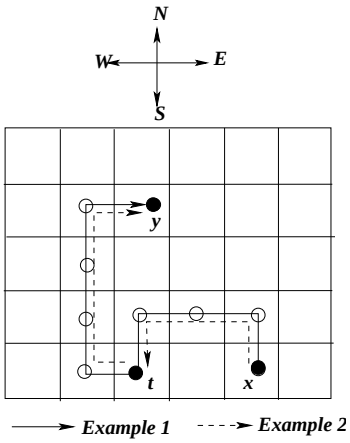

Figure 8: On-Demand Packet Forwarding/Routing over VGA: An example.

Node x initiates a route discovery by sending the RREQ packet. Assume that the complete route traversed by the RREQ when it learns a path to y (i.e, RREQ reaches y) is (N-W-W-S-W-N-N-N-E) from left to right. Therefore, the RREP will follow the path (W-S-S-S-E-N-E-E-S) to x which is simply obtained by reversing the direction in the path and taking the 1's complement of the direction of each hop. For the second case, assume the RREQ issued by x reaches an intermediate node, t, after four hops and RREQ contains the partial path, P_1 = (N-W-W-S). Also assume that node t has a path to y given as P_2 = (W-N-N-N-E). Then t will concatenate both paths P_1 and P_2 as one path $P = P_1||P_2$, where || is the concatenation operation. Hence, the resultant path to the destination is P= (N-W-W-S-W-N-N-N-E) and node t will use RREP to send P back to the source node using the following path (N-E-E-S). In the case of heterogeneous MANETs, an additional 2-bit field per hop is also used to select a specific subzone in a neighboring zone that has only SR nodes and the direction of communication is from an LR node to a specific SR node in that neighboring zone.

5.2 Transitive Closure Routing Approach

In this subsection, the concept of *transitive closure* from discrete mathematics [38] will be used to discover paths over VGA. Although the reader will notice a similarity between the routing approach described here and Dijkstra's algorithm in matrix operations, our approach automates the calculations and effectively uses matrix multiplication. In our approach, each CH maintains an up-to-date image of the adjacency matrix ($A = [a_{ij}]$ $i, j = 1, ..., n$). The adjacency matrix reflects the link status between the set of neighboring CHs such that a value of 1 in the entry (i,j) indicates that there is a link going from CH i to CH j; and

0 if no link exists. The set of CHs exchange the adjacency matrix whenever it is updated due to link addition/deletion.

Now, we illustrate how the concept of transitive closure can be used to find a path over VGA represented by the graph $G = (V, E)$. Let R be the relation on the set V with n elements, where n is the number of CHs in G. The relation R defines the one way connectivity (or reachability) between CHs and it can be mathematically defined as follows,

$$R = \{(x, y) | x, y \in V, \ \ y \ \text{ is reachable from } \ x \ \text{ in one hop}\}$$

where (x,y) is the ordered pairs of CHs in V. The relation R is typically asymmetric on the set V in heterogeneous setting where unidirectional links may exist in general, i.e., if $(x,y) \in R$, then it is not necessary that $(y, x) \in R$. Moreover, if there is a path in R from x to y, then there is such a path with length not exceeding n. Let T_R be the boolean matrix corresponding to the relation R. Then the boolean matrix of the transitive closure is defined as, $T_R^* = T_R \vee T_R^2 \vee T_R^3 \vee \ldots T_R^{n-1}$. To compute the matrix T_R^*, successive boolean powers of T_R (or A), up to the $n - 1^{st}$ power are computed. Since all entries in T_R are either 0 or 1, and since each row or column in T_R contains at most four nonzero elements, then each matrix multiplication operations can be done using at most $4n^2$ addition operations. This process can even be performed in n^2 bit vector operation; hence, these products can be computed using n^3 bit operations. Note that if all links are symmetric, then the number of computations can be halved. However, we consider the general case of asymmetric links. The search for the existence of a path between two mobile nodes (s,d) can be determined as follows. Suppose we compute T_R^i. If T_R^i of (s,d) is equal to 1, then a route of length i exists between (s,d). If T_R^i is not 1, then we calculate T_R^{i+1} and check for a route again. In the worst case, the transitive closure is used to determine if a path of length $n - 1$ between (s,d) exists or not. This is done by inspecting $T_R, T_R^2, T_R^3, \ldots, T_R^{n-1}$. For a square network area, it can be shown that the maximum path length is $(n + \sqrt{n})/2$, which can be easily found by inspecting the path between any two nodes in opposite corners.

This process checks for path existence between the source-destination pair. The actual path can be constructed by inspecting the binary multiplication results in a backward manner, i.e. if the binary multiplication of T_R was performed i times and results in a value of 1, the actual path of i hops can be constructed by backtracking through $T_R^{i-1}, T_R^{i-2}, \ldots, T_R^1$. The discovered path is then embedded in packet headers and is utilized to deliver packets using the packet forwarding scheme described earlier.

An important feature of a routing protocol is its ability to adapt or react to topological changes. When a zone becomes empty in VGA due to node movements, the routing protocol should be able to route packets around the empty zone. In our routing approach, the responsibility of detecting whether the next CH has failed or not falls on the direct upstream CH in the complete route list that is embedded in the packet header. Figure 9 illustrates how this is done. The figure shows the current active path from zone 1 to zone 8 between a certain source-destination pair. Assume that a CH in zone 5 leaves its zone and moves to a different zone, which will result in service disruption to current path. One solution is that CH in zone 2 will look for an alternate route (local). In Figure 9(d), current CH in zone 2 was able to find an alternate path locally and continues the connection without service disruption.

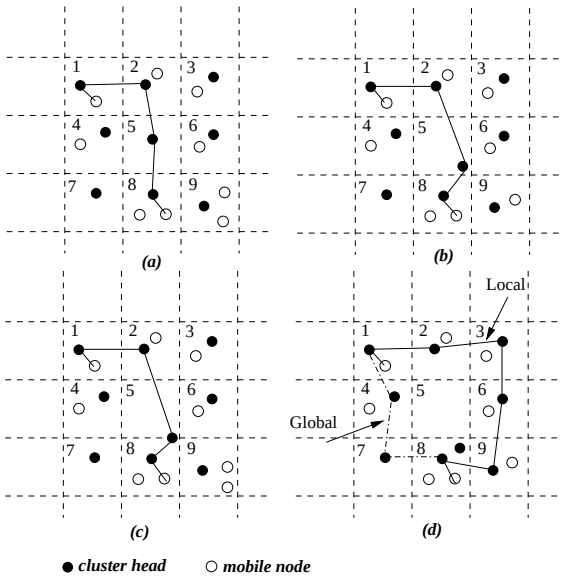

Figure 9: The illustration of local path restoration in VGA.

6 Simulation Results

In this section we present the simulation results of the proposed protocols in both homogeneous and heterogeneous networks. We divide the experiments into two suites, corresponding to the VGA clustering approach and the routing techniques over VGA. All experiments were performed using the NS 2.26 simulator [39], except for finding the optimal selection of CHs, which is evaluated using our ILP formulation. Each ILP problem instance was solved using the software tool CPLEX [40]. The simulation settings are as follows. A MANET with mobile nodes randomly placed within a fixed-size network area of size 1000×1000 m^2 and with a number of mobile nodes ranging between 10 and 300 were simulated. In homogeneous networks, transmission range was set to 150 meters. In heterogeneous MANETs, the long range transmission distance is set to 250 meters, while short range is set to 125 meters so that SR node zone side length is half that of LR node zone. Links have a transmission rate that is equiprobably selected from a pre-defined set of transmission rates (1, 2, 5.5, and 11 Mbps). The values of weighting factors in the CH election algorithm were set to $a_i = 0.25,\ i = 1 \ldots 4$. New session requests arrive according to a Poisson arrival process with mean arrival rate of 0.1 arrivals/second. The length of each session is exponentially distributed with the mean value being a simulation parameter. When a new session is generated, the routing protocol is initiated. Connections are established between randomly selected pairs of source/destination nodes. For the communication pattern, we used bursty traffic with ON-OFF periods. The ON and OFF periods of each flow are exponentially distributed with mean values being simulation parameters. We vary the session duration, the ON period length of each session, the packet generation rate during the ON period, and the mean OFF period length to obtain different levels of the offered load. Unless otherwise specified, the default offered load was set to 0.8. Each simulation experiment runs for a duration of 5000 seconds. No data collection was performed for the first 100 seconds of the simulation to avoid transient period and to ensure that the initial route discovery process is stable. The packet sizes are fixed at 512 bytes. Node mobility is captured by a more realistic mobility model, called Real Mobility Model (RMM), that

was originally proposed in [34], and recently appeared in [35]. In brief, node velocities and directions of movement in RMM are taken from probability distributions that mimic real user mobility behavior. The model is shown to be stationary and is able to produce mobility traces that closely resemble real mobility traces. Simulation experiments were conducted in both cases when nodes have infinite and limited energy, respectively. In the on-demand CH election scenario, a CH node will initiate a new or early CH election when its battery capacity drops to one third of its initial capacity. For simplicity, we assume that for each packet, the amount of energy consumed from the battery is 1.5mW for receiving, 3.0mW for transmitting, and 1.0mW for listening. All parameters for the physical layer were set to the default values of the NS-2 simulator.

6.1 Results of VGA Clustering Approach

In this section, we study the performance tradeoffs of the VGA clustering approach. In particular, we compare the performance of VGA clustering with other clustering algorithms proposed in the literature (summarized in Table 6). We also compare VGA clustering to the optimal clustering algorithm presented earlier in the paper. The metrics of comparison are:

(*i*) **CHs cardinality:** As mentioned earlier, the size of the virtual topology needs to be as small as possible. The bigger the size of the virtual topology, the higher the message overhead and the longer the routes. First, we compare the size (cardinality) of the virtual topology computed by different algorithms against VGA. We denote these algorithms by A_1, A_2, A_3, which correspond to algorithms in [15], [11], and [10], respectively. We randomly distribute the mobile nodes in the fixed network area such that they form a connected graph. In the case of optimal zoning, we divided the area into a number of hexagons and found the number of these hexagons which represents the upper limit on optimal CH cardinality. Figure 10 compares VGA to other algorithms in terms of the number of CHs. As Figure 10 shows, VGA is able to find a set of CHs that has a smaller cardinality than that obtained by other algorithms. This reaffirms the results of Table 6, namely that the cardinality obtained by VGA is closer to the optimal than other algorithms.

To compare the performance of VGA with the optimal sized MCDS obtained by solving the ILP for small MANETs, we generate connected graphs of 5, 10, 15, 20, 25, and 30 nodes. Each node uses a fixed transmission range of 150 meters. We used the ILP to solve for MANETs of these graphs, where the ILP was able to find the MCDS for these graphs in an average computation time equal to 0.32, 13.28, 27.42, 840.63, 1011.5, and 2016.42 seconds, respectively. We compare the cardinality of VGA against the optimal number obtained by solving the ILP. Figure 11 shows this comparison. In almost all cases, VGA is able to obtain cardinalities that are within 50% of the optimal solution. The reason is that VGA uses a fixed virtual topology that is made from zones which approximate the optimal hexagon zones. Hence, fewer CHs are needed to cover the network area. Note that the ILP, being computationally expensive, can only solve for networks with a limited number of nodes. Hence the comparison with VGA is also limited to networks with small numbers of nodes.

(*ii*) **Average route length:** Figure 12 shows the average route length of both VGA and D-VGA as well as the optimal average route length when the number of mobile nodes is varied between 10 and 300. The average path length was found through simulations. The average path length of VGA is closer

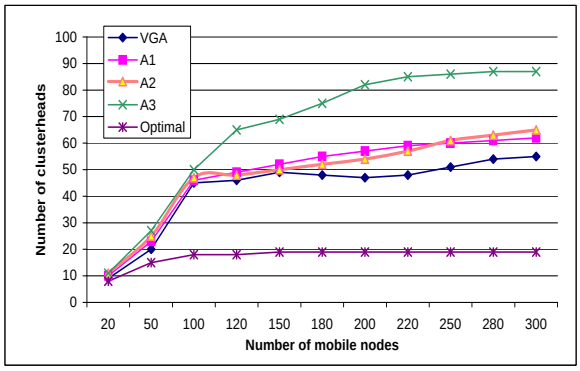

Figure 10: CHs cardinality; homogeneous MANETs.

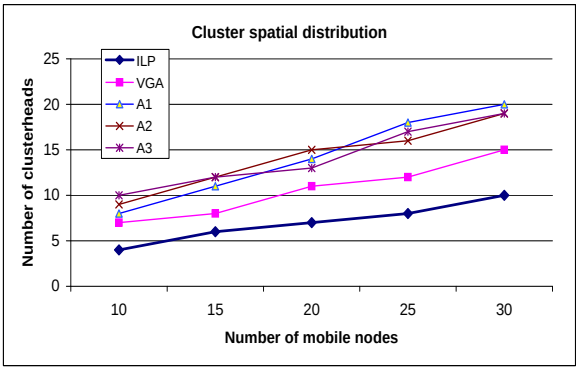

Figure 11: CHs cardinality; heterogeneous MANETs.

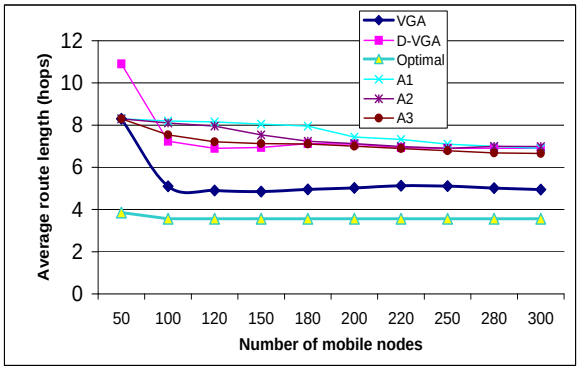

Figure 12: Average path length; VGA and D-VGA.

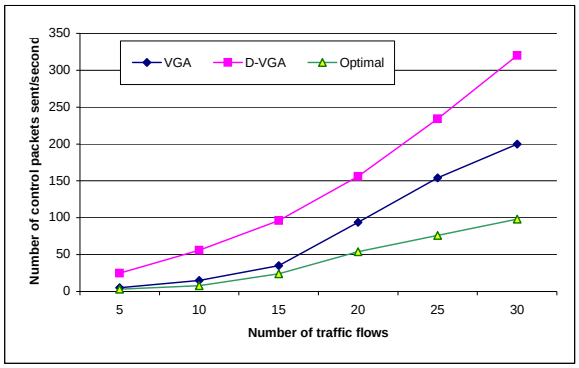

Figure 13: Communication Overhead in VGA and D-VGA.

to the optimal path length than D-VGA (the optimal path length was calculated using the closed form expression). The reason, as mentioned earlier, is that the cluster (zone) size in VGA is larger than the cluster size in D-VGA. Hence, on the average, fewer zones are encountered between any node pair in VGA routing. These results match the results obtained earlier in the form of closed form expressions in section 4.1.

(*iii*) **Control overhead:** The control overhead is defined as the total number of control messages sent every second in the network, which are exchanged between CHs. As might be expected, larger cardinalities mean higher communication overhead. Figure 13 shows the total number of control messages sent in VGA and D-VGA. As shown in the figure, VGA uses a smaller number of control packets than D-VGA in a constant manner as the amount of traffic in the network increases. This is due to the fact that in VGA, control packets are forwarded in four directions rather than in all possible eight directions used in D-VGA.

6.2 Routing over VGA

The two routing techniques proposed in Section 5 were simulated using the NS 2.26 simulator [39]. The simulation setup is as described in the beginning of this section. We evaluate the performance of the proposed routing techniques based on the following metrics:

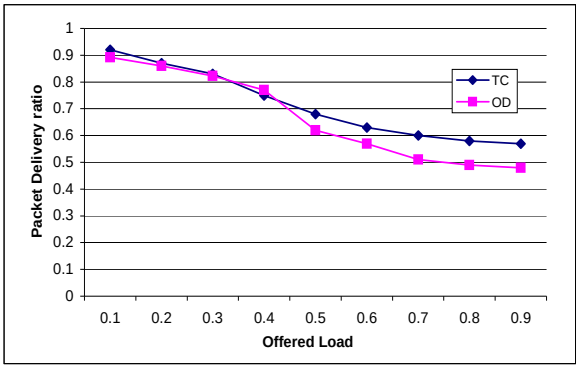

Figure 14: Packet Delivery Ratio vs. offered load.

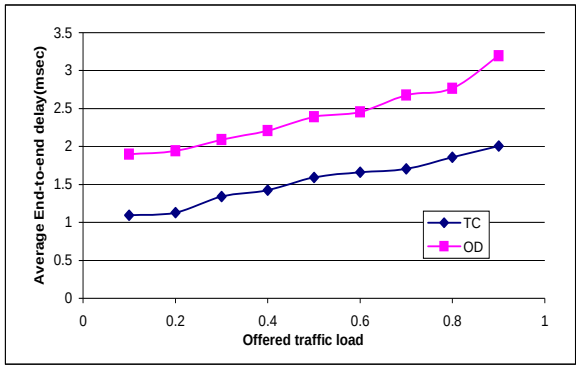

Figure 15: Network end-to-end delay vs. offered load.

(a) 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 generated by the application layer sources of the accepted calls. This ratio directly affects the maximum throughput that the network can support. Figure 14 shows the percentage of the packets delivered using both on-demand routing (OD) and transitive closure (TC) over VGA. Both schemes are able to achieve high packet delivery ratios and they maintain acceptable levels even when the offered load in the network increases. The stability of VGA allows both OD and TC based routing techniques to achieve high packet delivery. However, TC-based routing outperforms OD-based routing at higher offered load. This is because TC-based routing is able to locate paths quicker than OD-based based on local information and computation.

Recall that the optimal path is defined as the path with the minimum number of hops since it uses the fewest resources. We measured the percentage use of optimal paths used in TC-based routing. It is found that TC-based routing is able to use around 82% of the optimal paths when serving various calls. This is directly reflected on the high packet delivery ratio of VGA.

(b) Average end-to-end packet delay: The average end-to-end packet delay versus the offered traffic load for both routing techniques is shown in Figure 15. Note that a link becomes unreliable when it is broken and/or saturated with heavy traffic. 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 due to mobility. Therefore, packets that need to be re-routed will be queued and therefore encounter longer delays. Both routing techniques employ the efficient re-routing solution (*local-then-global*) discussed previously in order to reduce the average end-to-end packet delay. In Figure 15, it can be observed that TC-based routing improves the average end-to-end delay by as much as 20% over OD-based routing.

(c) Route Breakage: We measured the percentage of broken routes over VGA when the OD-based routing is used and compared the results to the well-known routing protocols, AODV and DSR. The mobility model used is RMM. Figure 16 shows this comparison. It is obvious from the figure that OD-based routing running on top of VGA is able to maintain more than 87% of the discovered routes, while the traditional AODV was able to maintain less than 65% of the routes while performing better than DSR.

(d) 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 the network for the length of the

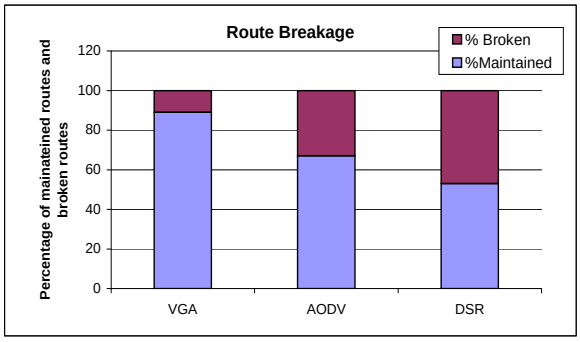

Figure 16: Percentage of broken routes due to mobility.

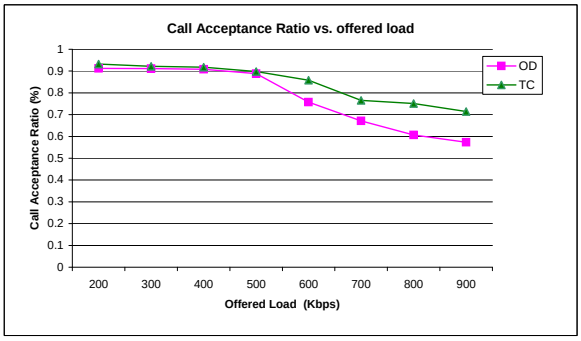

Figure 17: Call acceptance rate in the two routing techniques.

simulation period. Calls are rejected if routes to the required destinations cannot be found. Figure 17 shows the call acceptance ratio as a function of the offered traffic load for both routing criteria. As the figure shows, both the OD- and TC-based routing techniques are able to achieve more than 90% call success ratio at light loads. However, TC-based routing technique shows better performance than OD-based routing as the offered traffic load increases. This is because the distributed OD-based routing technique took longer to find a path compared to the TC-based approach.

(e) Routing Overhead: This is defined as the number of routing control packets per second that was sent and measured over the entire simulation experiment. Each transmission of a routing packet over each hop counts as one transmission. Routing overhead is an important metric for comparing routing protocols, as it measures the scalability of a protocol, the degree at which it will function in congested or low-bandwidth environments, and its efficiency in terms of consuming battery power. Protocols that send large numbers of routing packets can also increase the probability of packet collisions and may delay data packets in the network queues. Figure 18 shows the normalized routing overhead of the two routing techniques for variable offered traffic loads compared to the dynamic source routing (DSR) protocol and using the RMM mobility model. Normalized control overhead is defined as the ratio of the number of control packets generated to the number of packets that arrive at receivers. It can be seen that both routing techniques running over VGA generate lower normalized control overhead than DSR. Moreover, TC-based routing requires lower control overhead than OD-based routing, especially as the offered load increases. The control overhead in TC-based routing includes link state updates overhead besides other control packets.

(f) Cluster head election strategies: We also studied the effect of having fixed or variable periodic CH election as well as having on-demand CH election on the performance of the two routing techniques running on top of VGA. We measured the packet delivery ratio as a function of the offered load. We tried different fixed values for the CH election period and the reported results are for a fixed period of 2 seconds. For a variable period CH election, we developed a simple formula to set the CH election period in each zone such that the length of this period is dependent on three parameters at the beginning of CH election process: the number of nodes in the zone (n), the elected CH node available energy (e), and the speed of the elected CH (v). CH election is executed in each zone whenever the CH election timer expires. Note that due to node mobility, membership of each zone changes and this affects the CH election algorithm.

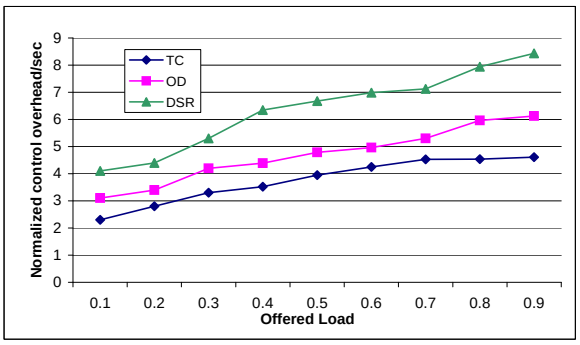

Figure 18: Network control overhead versus offered load using the TC and OD techniques.

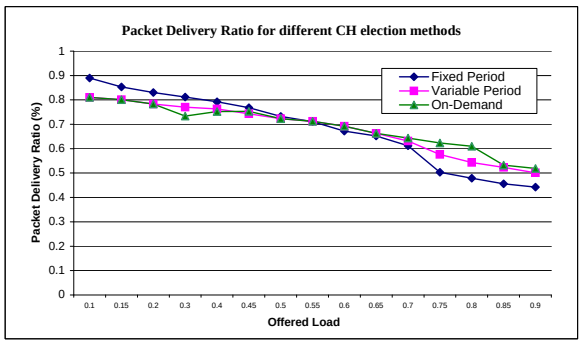

Figure 19: Packet delivery ratio versus offered load under different CH election strategies.

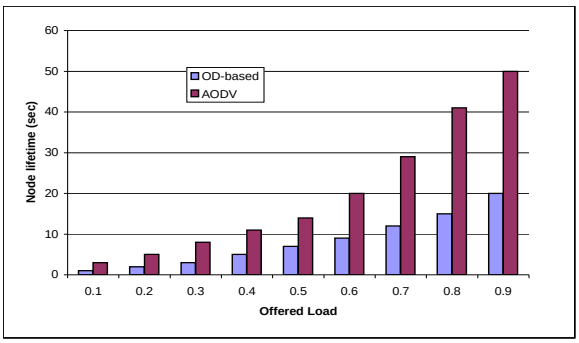

Figure 20: Effect of power control scheme of VGA on node lifetime.

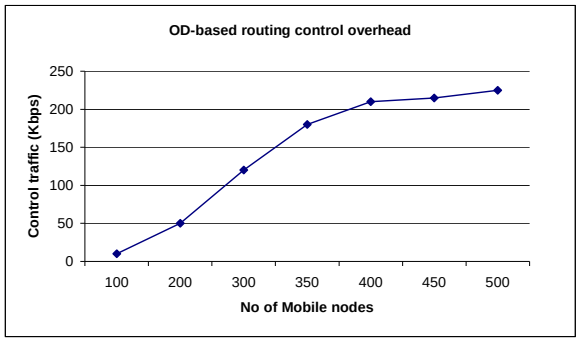

Figure 21: OD-based routing scalability as node density increases.

One simple equation is therefore: $CH_{period} = \frac{e}{v*n}$. Naturally, the CH period in the network decreases when there are more nodes in the zone or when the current CH moves quickly. As for the on-demand CH election strategy, a node can demand a new CH election if its energy level drops below 33% of its initial energy capacity. For this purpose, we assigned each mobile node a fixed amount of initial energy capacity of 8000J. Figure 19 shows the difference between the three CH election strategies when the CH election period is fixed to 2 seconds and when the CH election period is variable and is computed in each zone using the previous equation. For this experiment, each mobile node velocity v is initially selected per the PMM model such that it falls between (0∼5 m/s). We fixed the number of nodes in the network to 100. As can be seen in the figure, there is a tradeoff between the three strategies for different traffic loads. For low to moderate traffic loads, the fixed-period strategy resulted in the higher packet delivery ratios. However, on-demand strategy performs better at higher offered loads which indicates that CH nodes were suffering and hence needed to be replaced more often. This resulted in more load balancing among mobile nodes. Based on these observations, a hybrid approach which is load dependent may be used.

(g) Normalized energy consumption: The normalized energy consumption is defined as the total energy consumed in delivering the packets of accepted calls. The difference between the minimum and the maximum node lifetime under OD-based routing technique compared to plain AODV was measured. The difference reflects the time needed before the first sensor node or the network need to be replaced, i.e., it is directly related to the lifetime of both node and network, respectively. Figure 20 shows this comparison. OD-based routing technique that works on top of VGA is able to extend node lifetime more than AODV.

This is due to the use of the power control scheme of VGA. The bars shown in the figure represent the difference between minimum and maximum node life time for each offered load. Note that VGA is able to balance the energy consumption among mobile nodes which results in extending their lifetimes.

(h) 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. We use the same simulation environment presented previously except for the number of nodes and the grid (network area) size. We increase the network size to 3000m×2000m, and vary the number of nodes between 100 to 500. We generated 30 CBR sessions for this experiment. Each node moves with a speed randomly generated between 2∼20 m/s using the PMM model. Figure 21 shows the scalability of VGA clustering approach. As the the number of nodes increases, the normalized control overhead of OD-routing over VGA scales almost linearly with the increment in the number of mobile nodes after n=400. Furthermore, as the number of nodes increases, the overhead tends to decrease due to the stability of VGA, which is enhanced by the increase in node density in the network. When the node density is low meaning that we have sparse node distribution, the number of available paths was limited and hence the control overhead is also limited.

7 Conclusions

In this paper, we presented a novel clustering scheme that results in a fixed rectilinear virtual topology, called the Virtual Grid Architecture (VGA). The VGA clustering approach forms a fixed size clusters and selects clusterhead (CH) nodes using different eligibility scenarios. We conducted a performance study of the proposed clustering scheme. In particular, we focused on the tradeoffs between the proposed simple clustering scheme and an optimal clustering in MANETs. VGA clustering approach can be applied in both homogeneous and heterogeneous MANETs. For homogeneous MANETs and under the assumption of random distribution of a large number of nodes, we compared the performance of VGA clustering approach with the optimal clustering approach. Our results show that, although simple, the performance of the proposed VGA clustering is closer to the optimal than other existing algorithms, and it performs well in terms of CH cardinality, path length, and communication overhead. For heterogeneous MANETs, we provide an exact solution for the minimal clustering problem in arbitrary connected graphs by formulating the problem as an Integer Linear Program (ILP). We derived expressions for the communication overhead, and we showed that the proposed clustering scheme incurs lower communication overhead than related algorithms in the literature. Since topology stability in MANETs is a very important requirement, we show that our VGA clustering approach produces stable routes that exhibit long lifetimes. Overall, analytical as well as simulation results show that VGA clustering, although simple, is very good and it is better than other clustering schemes introduced in the literature, while being not too far from optimal.

Appendix
Proof of Proposition 1

Proposition: If the network area is large and is equal to $L \times L$ square meters, mobile nodes are uniformly distributed in the network area, and each node has a transmission range of r meters, then the worst case path length (PL) of VGA, D-VGA, and Optimal routing (in terms of the number of hops) is

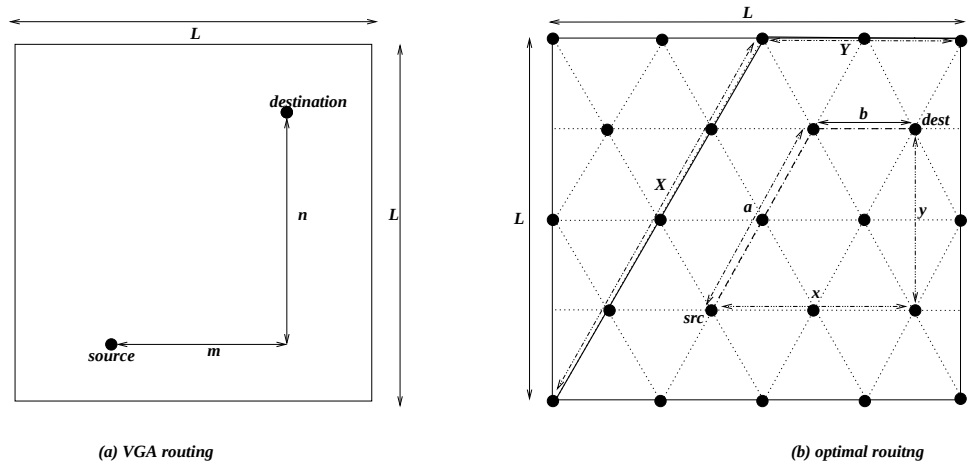

Figure 22: Figure used in the proof of Proposition 1

given by:

$$PL_{VGA}^{max} = \frac{2L\sqrt{5}}{r} - 1 \ , \ PL_{D-VGA}^{max} = \frac{L\sqrt{8}}{r} \ , \ PL_{Opt}^{max} = \frac{L}{r}(\frac{1+\sqrt{3}}{\sqrt{3}}),$$

while the average case path length is given by:

$$PL_{VGA}^{avg} = \frac{\sqrt{5}L}{r} \ , \ PL_{D-VGA}^{avg} = \frac{5\sqrt{2}L}{3r} \ , \ PL_{Opt}^{avg} = \frac{(3\sqrt{3}+11)L}{18r}.$$

Proof. Divide the network area into $Z = M \times M$ smaller zones. Let the horizontal length of an arbitrary route be m rows (zones) and vertical length be n columns (zones). With VGA routing, the zone side length is $\frac{r}{\sqrt{5}}$, then

$$M_{VGA} = \frac{L\sqrt{5}}{r},$$

and the worst case number of hops (H_{VGA}) is = $2M$-1 = $\frac{2L\sqrt{5}}{r} - 1$, which is $\approx \frac{2L\sqrt{5}}{r}$ for large L. Adding diagonal routing will decrement the zone side length to $\frac{r}{2\sqrt{2}}$, hence for D-VGA,

$$M_{D-VGA} = \frac{L\sqrt{8}}{r},$$

and the worst case number of hops (H_{D-VGA}) is = $max(m,n)$=$\frac{L\sqrt{8}}{r}$ for large L. Therefore, in the worst case, the rectilinear routing increases the number of hops by a factor of $\frac{\sqrt{20}}{\sqrt{8}} = 1.6$. For the optimal case, when the zones are hexagons, the worst-case path length can also be calculated as follows. The virtual topology connectivity for the optimal case is shown in Figure 22(b), where each clusterhead can communicate with 6 possible neighbors. The worst case number of hops (H_{opt}) in this case can easily be found by noting that the angles of movement are $0^o, 60^o, 120^o, 180^o, 220^o, and \ \ 300^o$. The worst-case path length (in terms of the number of hops) is $X + Y$ where X and Y are shown in Figure 22(b) and they can be easily obtained as $X = \frac{2L}{\sqrt{3}}$ and $Y = L(1 - \frac{1}{\sqrt{3}})$, hence the worst-case path length is

$$H_{opt} = \frac{L}{r}(\frac{1+\sqrt{3}}{\sqrt{3}})$$

Now, we will find the average hop length of a route (i.e., route length) between a source-destination (s, d) pair and for both diagonal and rectilinear routing, and compare it to the optimal route length. Let x

and y be two random variables representing the length of a route in the horizontal and vertical directions, respectively (see Figure 22(a)). Let $f_x(x)$ and $f_y(y)$ be the probability distribution functions of x and y, respectively. Further, assume that x and y are uniformly distributed in the network area. Hence, $f_x(x)$= $f_y(y)$= $\frac{1}{L}$. Then,

$$\text{VGA routing} : \ m = \frac{x}{r/\sqrt{5}}, \ n = \frac{y}{r/\sqrt{5}} \quad \text{D-VGA routing} : \ m = \frac{x}{r/\sqrt{8}}, \ n = \frac{y}{r/\sqrt{8}}$$

and the number of hops for both cases is:

$$\text{D-VGA}=max(m,n) \quad \text{VGA routing}=m+n-1$$

assuming $m+n >> 1$, and also assuming that the horizontal and vertical granularity is very large, then we can assume m and n are continuous. Therefore, the average path length, APL for the VGA routing is:

$$APL_{VGA} = \int_0^L m f_x(x)dx + \int_0^L n f_y(y)dy = \frac{\sqrt{5}L}{r} = \int_0^L \frac{x}{r/\sqrt{5}}\frac{1}{L}dx + \int_0^L \frac{y}{r/\sqrt{5}}\frac{1}{L}dy = \frac{\sqrt{5}L}{r}$$

For D-VGA routing, we have

$$APL_{DVGA} = \int_0^L (\int_{x=0}^{y} n f_x(x)dx + \int_{x=y}^{L} m f_x(x)dx) f_y(y)dy$$

$$= \int_0^L (\int_{x=0}^{y} \frac{y}{r/\sqrt{8}}\frac{1}{L}dx + \int_{x=y}^{L} \frac{x}{r/\sqrt{8}}\frac{1}{L}dx)\frac{1}{L}dy = \frac{5\sqrt{2}L}{3r}$$

What is left is to find the average path length of the optimal case. Let x and y represent the length of a route in the horizontal and vertical directions, respectively (see Figure 22(b)). It can be easily seen that the number of hops between an arbitrary source-destination pair is the sum of the two segments $a+b$, where a and b can easily be found as $a = \frac{2y}{\sqrt{3}r}$. Note that the value of b is dependent on the relation between x and y such that:

$$b = \begin{cases} (x - \frac{y}{\sqrt{3}})/r & , \ if \ \ x \le \frac{y}{\sqrt{3}} \\ (\frac{y}{\sqrt{3}} - x)/r & , \ if \ \ x \ge \frac{y}{\sqrt{3}} \end{cases}$$

Hence the average path length of the optimal case, APL_{opt}, can be found by finding:

$$APL_{opt} = \int_0^L \{\frac{2y}{\sqrt{3}r} + \int_{x=0}^{\frac{y}{\sqrt{3}}} (\frac{y}{\sqrt{3}} - x)\frac{1}{r} f_x(x)dx + \int_{x=\frac{y}{\sqrt{3}}}^{L} (x - \frac{y}{\sqrt{3}})\frac{1}{r}\frac{1}{L} f_x(x)\} f_y(y)dy$$

Recall that $f_x(x)$= $f_y(y)$= $\frac{1}{L}$, then

$$APL_{opt} = \int_0^L \{\frac{2y}{\sqrt{3}r} + \int_{x=0}^{\frac{y}{\sqrt{3}}} (\frac{y}{\sqrt{3}} - x)\frac{1}{r}\frac{1}{L}dx + \int_{x=\frac{y}{\sqrt{3}}}^{L} (x - \frac{y}{\sqrt{3}})\frac{1}{r}\frac{1}{L}dx\}\frac{1}{L}dy = \frac{(3\sqrt{3}+11)L}{18r}$$

□

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] J.N. Al-Karaki, A.E. Kamal, "On the Optimal Clustering in Mobile Ad hoc Networks", Proceedings of IEEE Consumer Communications and Networking, Las Vegas, Nevada USA / January 5-8, 2004.

[3] L. Bao and J.J. Garcia-Luna-Aceves, "Topology management in ad hoc networks", Proceedings of MobiHoc (2003).

[4] V. Kawadia and P. R. Kumar, "Power control and clustering in ad hoc networks", proceedings of INFOCOM 2003, Volume: 1, Page(s): 459-469, April 2003.

[5] I. Stojmenovic, "Position-based routing in ad hoc networks", IEEE Communications Magazine, July 2002, Page(s): 128-134, 2002.

[6] B. Das and V. Bharghavan, "Routing in ad hoc networks using minimum connected dominating sets", IEEE ICC' 97, Montreal, Canada, June 1997.

[7] A. Farago, I. Chlamtac, S. Basagni, "Virtual path topology optimization using random graphs", Proceedings IEEE INFOCOM '99, New York-USA.

[8] J. N. Al-Karaki, A. E. Kamal, "End-to-End Support for Statistical Quality of Service in Heterogeneous Mobile Ad hoc Networks", Computer Communications. Vol. 28, No. 18, Nov. 2005, pp. 2119-2132.

[9] W. Liao, Y. Tseng, J.P. Sheu,"GRID: A fully location-aware routing protocols for mobile ad hoc networks", Telecommunication Systems, Vol. 18, No. 3, Page(s): 37-60, 2001.

[10] K.M. Alzoubi, P.-J. Wan and O. Frieder, "New distributed algorithm for connected dominating set in wireless ad hoc networks", Proceedings of IEEE HICSS, 2002.

[11] J. Wu and H. Li, "On calculating connected dominating set for efficient routing in ad hoc wireless networks", Proc. of the 3rd International Workshop on Discrete Algorithms and Methods for MOBILE Computing and Communications, 1999, Seattle, WA USA, Page(s): 7-14.

[12] X. Cheng and D. Du,"Virtual backbone-based routing in multihop ad hoc wireless networks", Technical Report, Department of Computer Science and Engineering, University of Minnesota, Minneapolis, 2002.

[13] J. Wu, F. Dai, M. Gao, and I. Stojmenovic "On calculating power-aware connected dominating set for efficient routing in ad hoc wireless networks", Journal of Communications and Networks , Vol. 5, No. 2, March 2002, Page(s): 169-178.

[14] P. Sinha, R. Sivakumar, and V. Bharghavan,"CEDAR: a core-extraction distributed ad hoc routing algorithm", Proceedings of INFOCOM 1999.

[15] M.R. Pearlman, and Z.J.Haas,"Determining the optimal configuration for the zone routing protocol", IEEE Journal on Selected Areas in Communications, Vol.: 17 No. 8 , Aug. 1999 Page(s): 1395–1414.

[16] C. E. Perkins and E. M. Royer, "Ad hoc on demand distance vector routing", proceedings of 2nd IEEE Workshop. Mobile Computing Sys and Apps., Page(s): 90-100, Feb. 1999.

[17] Y. Xu, J. S. Heidemann, and D. Estrin. "Geography-informed energy conservation for ad hoc routing", In Mobile Computing and Networking, pages 70-84, 2001.

[18] X. Du, D. Wu, "Adaptive cell relay routing protocol for mobile ad hoc networks", IEEE Transactions on Vehicular Technology, Volume 55, Issue 1, Jan. 2006 Page(s):278-285.

[19] Z. Tao, G. Wu, "An analytical study on routing overhead of two-level cluster-based routing protocols for mobile ad hoc networks", 25th IEEE International Performance, Computing, and Communications Conference (IPCCC), 10-12 April 2006 Page(s):8-12.

[20] A. Koyama, Y. Honma, J. Arai, L. Barolli, "An enhanced zone-based routing protocol for mobile ad-hoc networks based on route reliability", International Conference on Advanced Information Networking and Applications (AINA), Volume 1, 18-20 April 2006 Page(s):6-11.

[21] J. Wu, F. Dai, "Virtual Backbone Construction in MANETs Using Adjustable Transmission Ranges", IEEE Transactions on Mobile Computing, Volume 5, Issue 9, Sept. 2006 Page(s):1188-1200.

[22] J. Wu, L.Wei Lou, F. Dai, "Extended multipoint relays to determine connected dominating sets in MANETs", IEEE Transactions on Computers, Volume 55, Issue 3, March 2006 Page(s):334-347.

[23] K. Robinson, D. Turgut, M. Chatterjee, "An Entropy-based Clustering in Mobile Ad hoc Networks", Proceedings of the 2006 IEEE International Conference on Networking, Sensing and Control (ICNSC), 23-25 April 2006 Page(s):1-5.

[24] G. Angione, P. Bellavista, A. Corradi, E. Magistretti, "A k-hop Clustering Protocol for Dense Mobile Ad-Hoc Networks", 26th IEEE International Conference on Distributed Computing Systems Workshops (ICDCS), 04-07 July 2006 Page(s):10-20.

[25] H. Nakagawa, K. Ishida, T. Ohta, Y. Kakuda, "GOLI: Greedy On-Demand Routing Scheme Using Location Information for Mobile Ad Hoc Networks", 26th IEEE International Conference on Distributed Computing Systems Workshops, (ICDCS), 04-07 July 2006 Page(s):1-10.

[26] W. Choi, M. Woo, "A Distributed Weighted Clustering Algorithm for Mobile Ad Hoc Networks", International Conference on Internet and Web Applications and Telecommunications (AICT-ICIW), 19-25 Feb. 2006, Page(s):73-73.

[27] M. Xue, E. Inn-Inn, W. Sah, "Analysis of Clustering and Routing Overhead for Clustered Mobile Ad Hoc Networks", 26th IEEE International Conference on Distributed Computing Systems, (ICDCS), 04-07 July 2006 Page(s):46-46.

[28] T. Chiu, S. Hwang, "Efficient Fisheye State Routing Protocol using Virtual Grid in High-Density Ad-Hoc Networks", 8th International Conference on Advanced Communication Technology, (ICACT), Volume 3, 20-22 Feb. 2006 Page(s):1475-1478.

[29] T. Liu, S. Hwang, "On Design of an Efficient Hierarchical Location Service for Ad Hoc Network", 1st International Symposium on Wireless Pervasive Computing, Jan. 2006, Page(s):1-6.

[30] B. Liang, Z. J. Haas, "Hybrid Routing in Ad Hoc Networks with a Dynamic Virtual Backbone", IEEE Transactions in Wireless Communications, Vol. 5, No. 6, June 2006.

[31] P. Santi, "Topology Control in Wireless Ad Hoc and Sensor Networks", ACM Computing Surveys, Vol. 37, No. 2, June 2005, Page(s): 164194.

[32] Roy Friedman Guy Korland, "Timed Grid Routing (TIGR) Bites off Energy", Proceedings of ACM MobiHoc05, May 25-27, 2005, USA.

[33] Y. Wu et.al., "Spreading code assignment in an ad hoc DS-CDMA wireless network", ICC 2002. Volume: 5, 2002 pp: 3066-3070.

[34] Jamal N. Al-Karaki, "Infrastructureless wireless networks: Cluster-based architectures and protocols", http://archives.ece.iastate.edu/archive/00000062/, Ph.D. thesis, Iowa State University, 2004.

[35] A. E. Kamal, J. N. Al-Karaki, "RMM: A Realistic Mobility Model for Mobile Ad hoc Networks", in the proceedings of the IEEE International Conference on Communications (ICC) 2007.

[36] M. R. Garey and D. S. Johnson,"Computers and intractability, a guide to the theory of NP-Completeness", W. H. Freeman and Company, 1979.

[37] D. Goodman, "Wireless personal communications systems", Reading, MA: Addison-Wesley, 1997.

[38] K. Rosen, "Discrete mathematics and its application", Chapter 6, Fourth edition, McGraw Hill Publishing, 1999.

[39] "The ns-2 simulator", http://www.isi.edu/nsnam/ns (date retrieved: April 7, 2004).

[40] CPLEX optimization Incorporation, "Using the CPLEX Callable Library", V4.0, Published 1995, http://www.ilog.com/products/cplex/ (date retrieved: April 7, 2004).

[41] J. W. Mark, W. Zhuang, "Wireless Communications and Networking", Prentice Hall, 2002.