A MANET Architecture for Airborne Networks with Directional Antennas

William Huba, Nirmala Shenoy
Golisano College of Computing and Information Sciences
Rochester Institute of Technology, Rochester NY 14623, USA
nxsvks@rit.edu, bill.huba@gmail.com

Abstract—Surveillance using unmanned aerial vehicles (UAVs) is an important application in tactical networks. Such networks are challenged by frequent link and route breaks due to highly dynamic network topologies. This challenge can be addressed through robust routing algorithms and protocols. Depending on the surveillance area to be covered and the transmission range of the transmitters in the UAVs, several of them may have to be deployed, requiring solutions that are scalable. The use of directional antennas mitigates the challenges due to limited bandwidth, but requires a scheduling algorithm to provide conflict free schedules to transmitting nodes. In this article we introduce a new approach, which uses a single algorithm (i) that facilitates multi hop overlapped cluster formations to address scalability and data aggregation; (ii) provides robust multiple routes from data originating nodes to data aggregation node and (iii) aids in performing distributed scheduling using a Time Division Multiple Access (TDMA) protocol. The integrated solution was modeled in Opnet and evaluated for success rate in packet delivery and average end to end packet delivery latency. High success rates combined with low latencies in the proposed solution validates the use of the approach for surveillance applications.

Keywords—Airborne Surveillance; Network of Unmanned Aerial Vehicles; Directional Antennas; Time Division Multiple Access; Distributed Scheduling

I. INTRODUCTION

Surveillance networks comprising of airborne nodes such as unmanned aerial vehicles (UAVs) are a category of mobile ad hoc networks (MANETs), where nodes are travelling at speeds of 300 to 400 Kmph. Surveillance requires aggregation of data captured by all nodes in the network at few nodes, from where the data is then sent to a center for further action. Due to high mobility of nodes and varying wireless environment, the topology in surveillance networks is subject to frequent and sporadic changes. Such MANETs thus face severe challenges when forwarding data from node to node, which is the task of the medium access control (MAC) protocol and also in discovering and maintaining routes between source and destination nodes, which is the task of the routing protocols. Another challenge faced is the scalability of the protocols to increasing number of nodes.

In this article, a unique solution for surveillance networks comprising of UAVs, equipped with directional antennas is proposed and investigated. The solution uses a single algorithm for several operations such as (i) multi-hop overlapped cluster formation, (ii) routing of data from cluster clients to cluster head for data aggregation, and (iii) scheduling concurrent time slots to transmitting nodes using a Time Division Multiple Access (TDMA) based MAC protocol, in UAVs that use directional antenna systems. To best leverage the strengths of this approach, the MAC, clustering and routing functions were implemented as processes operating using a single address generated by the algorithm to collaboratively address the challenges faced in surveillance networks. This leads to a new MANET architecture. Due to the critical nature of the application the new architecture and a unified approach is justified. Performance evaluations conducted in airborne networks with twenty, fifty and seventy five UAVs validate these justifications.

Surveillance applications require low packet loss and low packet delivery latencies hence in this work the analysis was directed primarily towards these performance metrics. Other performance metrics such as MAC and routing operational overhead were also recorded. The architecture introduced in this article achieves the performance goals. Due to lack of similar published work and the availability of evaluation models of implementations in such application scenarios, the presentation in this article is limited to the results from simulations of the proposed architecture.

The rest of the paper is organized as follows. Section II describes related work in the area of TDMA MAC, routing in large MANETs and clustering. The benefits of the integrated approach are highlighted in the light of these discussions. Section III describes the Integration Architecture, the rationale for the same, the components of the architecture and the interworking principles. Section IV describes the scheduler and the link assignment strategy. Section V provides the simulation details in Opnet and the performance analysis based on data collected. Conclusions and possible enhancements are discussed in Section VI.

II. RELATED WORK

The topic areas of major contribution in this article relate to routing, clustering and medium access control for use with directional antennas in MANETs. The significance of the proposed solution lies in the closely integrated operations of routing, clustering and medium access control coordinated by a control entity which has intelligence to coordinate their operations based on the applications requirements. To the best of our knowledge integrated clustering, MAC and routing solutions to MANETs have not been investigated though integration of clustering and routing have been researched. One of the main goals in this approach was to break down the limitation of protocol
layering towards an efficient MANET solution. Cross layered approaches, which break down such limitations in inter-layer communications, also facilitate a more effective integration and coordination between protocol layers. However, the proposed solution is not a cross layered approach, as one main problem encountered in such approaches is still the integration framework that has to work across different techniques and algorithm used by the different routing, clustering and MAC protocols. If the MAC uses a scheduled TDMA approach that has to work with directional antennas, the challenges are compounded. It was felt that for dedicated and critical MANET applications, one should not be constrained by the protocol layers or stacks, and other existing norms in the regard. What is important though is that such solutions should co-exist and interwork with networks that use current protocol structures.

Due to the uniqueness of the approach, it is not possible to cite and discuss related work that adopts similar techniques. Hence related work in each of the component topic areas are discussed under several subsections. Subsection A describes related work in the area of directional antennas and scheduled MAC protocols. This is followed by routing protocols and algorithms to address scalability in Subsection B. Hierarchical and hybrid routing protocols fall under this category. Routing combined with clustering is another approach to address scalability in MANETs and are discussed in Subsection C. Clustering, especially multi-hop clustering is very important to address scalability in MANETs; they also aid in data aggregation which is important in surveillance applications and is discussed in Subsection D. Lastly in Subsection E the significance of the proposed approach in the light of the related work is discussed.

A. Scheduled MAC in Directional Antenna Systems

To achieve higher capacity and improved delay guarantees in the network, Spatial reuse TDMA (STDMA) scheme is employed at the MAC layer [1, 2, 3]. In STDMA, which is an extension of TDMA, time is divided into time slots; and multiple transmissions can be scheduled as long as the receiving nodes do not get their packets interfered with. In this manner, STDMA takes advantage of the spatial separation between nodes to reuse the time slots. Generally, such schemes require strict time synchronization among participating nodes for efficient transmission and reception among the nodes. In addition, as a result of mobility of nodes in MANETs, periodic changes in the network require that STDMA schedules, which describe transmission rights of nodes in the network, be updated with minimal computational complexity. Furthermore, the updated schedule must be propagated to all nodes in the network in timely and efficient (using less resources) manner.

One of the most challenging tasks in such schemes is generating the STDMA schedule(s) that efficiently use the network resources. Since multiple nodes can simultaneously transmit in the same time slot, an optimal STDMA scheduling algorithm must allow high reuse of time slots with minimal interference while minimizing frame length (i.e. number of time slots per frame). Multiple algorithms have been proposed in literature [4-10]. The scheduling function can be performed by one of the participating node – a centralized scheduler. Centralized scheduling requires all information about the network such the number of nodes and links at the central scheduler, which is difficult to achieve. On the other hand, distributed scheduling can be done at the expense of increased complexity. In distributed scheduling, only nodes in the region of the change will act on it and update their schedules on network changes. In cluster based solutions centralised STDMA scheduling [6, 7], is less complicated and more efficient since each cluster head has all information about nodes in its cluster. Unfortunately, the overhead costs due to re-distribution of schedule whenever the network changes, are higher than that of distributed STDMA scheduling.

B. Routing in MANETs

Literature is rich with work conducted in the area of routing and clustering for MANETs. Several survey articles published on MANET routing and clustering schemes from different perspectives indicate the continuing challenges in this topic area. In [16] the authors present a survey of routing protocols and cross layer design effects. The survey presented in [17] is under the three broad categories of proactive, reactive and hybrid routing. A comprehensive technical report on MANET routing protocols [18, 20] covers them under the categories of uniform and non-uniform routing protocols, hierarchical (topology and cluster based), position based and so on, with performance comparisons. Reference [22] is an early review article that covers the characteristics of several routing protocols.

1) Proactive Routing Protocols

Proactive routing protocols require dissemination of link information periodically so that a node can use standard algorithms such as Dijkstra’s to compute routes, to all other nodes in the network or in a given zone [27]. Link information dissemination requires flooding of messages that contain link information. Depending on the node mobility and wireless media conditions and the periodicity in link information dissemination, in large networks, such transmissions can consume significant amount of bandwidth making the proactive routing approach not scalable. Several proactive routing protocols thus target mechanisms to reduce this control overhead. Fisheye State Routing (FSR) introduces multi-level fisheye scope with reduced routing packet sizes and update frequency [28] to remote nodes. Fuzzy Sighted Link State uses the optimal routing algorithm, Hazy Sighted Link State [30] to further reduce link message dissemination. Multi scope approaches work well when the network grows in terms of number of hops end-to-end. Optimized Link State [25] reduces flooding of messages by using selected one hop nodes as multi point relays, to propagate link messages. Topology Broadcast Reverse Path
forwarding [29] propagates link-state updates in the reverse direction on a spanning tree formed by the minimum-hop paths from all nodes to the source node. The last two schemes achieve high efficiency in a dense network.

2) Reactive Routing Protocols

Reactive routing protocols avoid the periodic link information dissemination and allow a node to discover routes to a destination node only when it has data to send to that destination node. The reactive route discovery process can result in the source node receiving several route responses which it may cache. Routing overheads in reactive routing protocols can thus be considerably low if the number of simultaneously communicating nodes is not high. As mobility increases, route caching may become ineffective as pre-discovered routes may become stale and unusable. Dynamic Source Routing (DSR) [24] protocol, after the discovery, requires each data packet to carry the full address of every hop in the route, from source to the destination, and hence faces scalability problems as the addresses could be MAC (48 bits) or IP (32 bits) or IPv6 (128 bits). Ad Hoc On-demand Distance Vector (AODV) [23] routing protocol overcomes this problem by using intermediate nodes to maintain the forwarding information. Temporally Ordered Routing Algorithm (TORA) [18], [19] protocol uses link reversal, route repair and creation of Directed Acyclic Graphs (DAGs), similar to Light-Weight Mobile Routing (LMR) [34] and inheriting its benefits but reducing far-reaching control messages.

3) Hierarchical Routing

Partitioning a MANET physically or logically and introducing hierarchy can limit message flooding and also address the scalability. Mobile Backbone Networks (MBNs) [35] use hierarchy to form a higher level backbone network by utilizing special backbone nodes with low mobility to have an additional powerful radio to establish wireless link among them. LANMAR [34] was extended to route in the MBN.

4) Hybrid Routing

Scalability in MANET routing protocols have been addressed by combining proactive and reactive routing in a hybrid approach, where the use of proactive routing is restricted to a limited area or zone and reactive routing is used when communicating with distant nodes. Zoning requires some form of partitioning mechanism. Sharp Hybrid Adaptive Routing Protocol (SHARP) [36] is application adaptive and automatically finds the balance point between proactive and reactive routing. In SHARP, a hot destination node that receives data from many sources determines a proactive zone, and outside of the zone any reactive routing algorithm like AODV or DSR could be used. Hybrid Routing for Path Optimality (HRPO) [32] combines proactive route optimization to a reactive source routing protocol to reduce average end-to-end delay in packet transmissions. The Zone Routing Protocol (ZRP) [33]] is a hybrid routing protocol, where each node has a pre-defined zone centered at itself. Any proactive routing can be used within the zone and any on-demand routing can be used for inter zone communications. ZRP provides a route discovery mechanism outside the zone through a Bordercast Resolution Protocol (BRP), where BRP establishes a Bordercast tree to send the discovery messages to the border nodes in a given zone.

C. Routing and Clustering

Nodes physically close to each other form clusters with a cluster head communicating on behalf of the cluster. Multi Hop clustering techniques such as the d-hop or k-hop clustering [8] algorithms can offer flexibility in terms of controlling the cluster size and cluster diameter, but are often complex to implement.

1) Clustering and Zoning

Clustering or zoning can be efficiently employed for the type of convergecast traffic encountered in surveillance networks, were the primary traffic flow is from cluster clients (CC) to cluster head (CH) [11- 15]. In such cases proactive routing approaches are recommended as the routing is limited to the cluster or zone and will also reduce stale routes. However proactive routing algorithms require the dissemination of link state information to all routers in the network or zone, which can introduce latency in realizing or breaking a route, and high overhead.

2) Cluster Based Routing

Different routing strategies can be used inside and outside the cluster. Several cluster based routing were designed to address scalability in MANETs. Cluster Head Gateway Switch Routing (CGSR) [15] is a cluster based hierarchical routing scheme. A mobile node belonging to two or more clusters acts as a gateway connecting the clusters. CGSR uses distance vector routing and maintains a cluster member table and a routing table at each node. Hierarchical State Routing (HSR) [16] is a multi-level, clustering based link state routing protocol that uses the clustering scheme recursively. In HSR, Hierarchical ID (HID) is used which is a sequence of MAC addresses of nodes on the path from the top of the hierarchy to the node.

D. Significance of the Architecture

From the above discussions it would be clear that clustering, routing and scheduling are different operations and hence normally are based on different algorithms or techniques. When combining the different operations, it becomes essential to define an interworking mechanism for the different algorithms. This adds processing complexity. It also results in added overhead for the operation of the combined functions. If all these operations can be based off a single algorithm, the complexity and overhead can be reduced significantly as demonstrated in this work.

If the above approach were possible, and if the MAC, routing, clustering and scheduling can use a single address
for their operation (unlike our current protocol stack, where MAC protocol uses 48 bit MAC addresses for its operation and routing protocols use 32 bit IP addresses (or 128 bits If IPv6)), we can achieve a solution, where the processes can closely interact and also avoid issues and overhead due to protocol layering, handling different headers and complex cross layered techniques. This would also make the solution compact and efficient and foster close interworking among the different operations.

III. The Integrated Approach

Given the challenges faced by MANETs, the authors decided to approach the solution from a holistic perspective. Towards this the essential functions required to support communications among the mobile entities in a MANET were identified. An architecture that would aid in best organizing the functions, taking into account the application demands and the challenging wireless media was then designed. The architecture would continue support for the existing protocol layered structure by either bypassing them during operation, interwork with them or replace them with a provision to bridge with networks based on these protocol structures.

The new architecture proposes a communications layer that bridges the application and the physical layers directly, bypassing other protocol layers. The communications layer includes routing, clustering and medium access functions whose operations are coordinated by an intelligent entity that incorporates the needs of the application taking into consideration the physical layer constraints.

A. The Rationale

Protocol layering introduces operational overhead. It also reduces efficient interworking among the protocols. Cross-layered techniques to address the communications needs of the wireless ad hoc networks were crafted for the purpose. Such techniques however introduce complexity as they overlay on the existing protocol structures. A few points to consider at this time is; (i) given a new wireless networking scenario and environment and the ensuing challenges, is there a need to continue with structures, algorithms and protocols that were developed for less challenging network situations such as the wired networks; (ii) secondly is there a need to continue with the two addresses in a bandwidth constrained environment? (iii) how about the complexity and resulting unreliability and lack of robustness?, and (iv) lastly how does this impact on the weight and power constraints faced by mobile devices? There is undoubtedly need for networks to interwork with one another, which does not however impose the condition that they have to use the same protocols, structures and so on.

B. The Architecture

A schematic of the architecture is shown in Figure 1. The light colored box indicates the use of either the TCP/IP protocol suite just below the Applications layer or the implementation of thin dummy protocol to incorporate port functions. The approach is similar to the Multiprotocol Label Switching used for tunneling to bypass IP layer often adopted in wired networks. It is different however as the communications layer now has all functions required for MANET operation, the MAC to enable sharing the wireless medium, the routing functions to discover routes reactively or proactively and clustering which is needed to address the scalability demands of MANET applications. All these functions are now coordinated by an intelligent Operation Control (OC) entity.

![Figure 1 The Integration Architecture](image)

The crucial entity in this architecture is the OC. Any MAC, routing or clustering protocol could be used in the other blocks. However, if these protocols operate on different techniques and address schemes, the effectiveness of the OC unit is reduced and it could become very complex balancing off the benefits of the approach. Note the positioning of the OC unit in the architecture (without TCP/IP suite) would provide information on the applications traffic and their quality requirements, whilst also collecting data on the Physical layer to control and coordinate the operations of the other entities in the communications layer. If TCP/IP were included then the information from the application can be passed through the DiffServ field in the IP (v4) header. However IP routing would be bypassed and the solution operates transparent to layer 3 protocols.

C. The Components

In this section, the different components used in the communications layer of the architecture in this work will be described. As the OC unit is crucial to the architecture, this will be the first component to be discussed. To make the OC unit efficient it is important to adopt an algorithm or technique that would allow coordinated operation of the other three entities in the communications layer. The significance of the coordinated operation would be clear at the end of this section and will be justified when the performance is discussed. The Multi-Meshed Tree (MMT) algorithm was selected for this purpose. This algorithm modified accordingly has already been used to support MAC, routing and clustering [37-41]. MMT is also amenable to optimization based on the network communications needs and is discussed under future work in the Conclusion Section. In this work, the algorithm was...
enhanced to maintain several connections between nodes in a meshed tree cluster and the cluster head which is the root of the meshed tree. For completeness, the MMT algorithm is first briefly described. This is followed by clustering supported by MMT, the proactive route maintenance and then the establishment of reactive routes for communication among nodes between clusters. This is followed by the interworking principles between the scheduler, MAC and the directional antenna system.

1) Meshed Tree

It is a traditional approach to have mesh connections among communicating nodes for redundancy purposes, but the mesh is then logically configured (by blocking some of the physical connections) into a tree using either the spanning tree approach or the Dijkstra approach (or other tree algorithms) to avoid looping packets. Logical tree creation adversely impacts the strengths of the meshed topology, as on any physical link changes the tree connections can break and the tree will have to be recreated. This is true of wired and wireless networks. In wireless ad hoc networks such link breaks can occur more often and the latency introduced in the network connectivity establishment and convergence can be detrimental.

Contrary to this approach, the meshed tree algorithm allows building several tree branches that exist concurrently from a single root by leveraging all the connections possible in the meshed structure without logically blocking any links. The tree branches so formed are limited only by criteria specified for the meshed tree creation. Looping is avoided through the use of a smart numbering scheme to define the tree branches. Thus under the meshed tree algorithm a node resides in multiple tree branches unlike the trees formed by the Dijkstra or the spanning tree algorithms. On the failure of one path (tree branch) the node remains connected to the root on another path, without the need to rebuild the tree. Meshed tree construction is dynamic and the tree branches evolve continually based on the decisions by the nodes to join a branch. In the case of mobile nodes this feature allows the nodes to remain connected to the root with a high probability despite link breaks. Time lags and their impact to reconstruct the tree and their resultant performance impacts are avoided.

2) Meshed Tree Clusters

As per the proposed solution the meshed tree created around a designated or elected root is a cluster; the root node is the cluster head. The creation of a meshed tree is explained with the aid of Fig. 2. The dotted lines link nodes that are in communication range with one another at the physical layer. The root of the meshed tree is labeled ‘CH’ for cluster head. Nodes A to G are the cluster clients (CC). For simplicity in explanation, the meshed tree formation is restricted to nodes that are connected to the CH, by a maximum of 3 hops. At each node several values or IDs are noted. These are the virtual IDs (VIDs) assigned to the node as they join a meshed tree branch in the cluster. Let the CH be assigned a VID 1, the CCs have l as a prefix in their VIDs. Any CC that attaches to a branch is assigned a VID, which inherits the prefix from its parent node, followed by an integer, which indicates the child number under that parent. In this work we limit the number of children to nine and use single digits to identify the children nodes. This does not eliminate the possibility of the scheme to have more than nine children under one node. It was not used in this case, as having too many paths going through a single node could create bottlenecks.

D. Routing in the Architecture

(i) Proactive Routes in the Cluster

In Fig. 2, each tree branch (shown by the dotted-dashed lines with an arrow head) is a sequence of VIDs that is assigned to CCs connecting at different points of the branch. The branch information of the meshed tree provides the route to send and receive data and control packets between the CCs and CH. For example, the branch denoted by VIDs 14, 142 and 1421, connects nodes C (via VID 14), F (via VID 142) and E (via VID 1421) respectively to the CH. To forward a packet from CH to node E, its VID 1421 will be used as the destination VID. When such a packet is broadcast, enroute nodes C and F receive the packet and forward to E. This is possible as the VIDs for nodes C and F are contained in E’s VID. The VID of a node thus provides a virtual path vector from the CH to itself. Note that the CH could have also used VIDs 143 or 131 for node E, in which case the path taken by the packet would have been CH-C-E or CH-D-E respectively. Thus between the CH and node E there are multiple routes identified by the multiple VIDs. The support for multiple routes through the multiple VIDs, allows for robust and dynamic route adaptability to topology changes in the network and the cluster. Nodes can request for new VIDs and join different branches as their neighbors change.

To send a packet from node E to CH, the packet has to be directed to destination VID 1, which is its first digit. To send packets to other nodes in the cluster, the packet can be passed via the CH, a common parent node or to a child node.
forward packet to other nodes either the cluster head will be used, or the packet can be sent directly on the branch if the source and destination node have (grand) parent or (grand)child relationship.

G. Scalability in the Architecture

1) Inter-Cluster Overlap and Scalability

A surveillance network can comprise of several tens of nodes; hence the solutions for surveillance networks have to be scalable to that many nodes. We assume that several ‘data aggregation nodes (i.e., CHs)’ are uniformly distributed among the non-data aggregation nodes during deployment of the surveillance network. Meshed tree clusters can be formed around each of the data aggregation nodes by assuming them to be roots of the meshed trees. Nodes bordering two or more clusters are allowed to join the different meshed trees and thus reside in the branches originating from different CHs. Such border nodes will inform their CHs about their multiple VIDs under the different clusters. When a node moves away from one cluster, it can still be connected to other clusters, and thus the surveillance data collected by that node is not lost. Also, by allowing nodes to belong to multiple clusters, the single meshed tree cluster based data collection can be extended to multiple overlapping meshed tree (MMT) clusters that can collect data from several tens of nodes deployed over a wider area with a very low probability of losing any of the captured data. This addresses the scalability requirements in surveillance networks.

Figure 3 shows 2 overlapped clusters and some border nodes that share multiple VIDs across the two clusters. The concept is extendable to several neighboring clusters. Nodes G and F have VIDs 142, 132 under CH1 and VIDs 251 and 252 under CH2, respectively.

2) Flexible Multi-hop Cluster Formation

Except for the CH, each node in Fig. 2 is a CC that will send the captured surveillance data to the CH. The size of the tree branch can be limited by limiting the length of the VID, which in turn allows control of the diameter of the cluster. Each node that joins the cluster has to register with the CH, by forwarding a registration request along the branch of the VID. This confirms the path defined by the VID and also allows the CH to accept /reject a joining node to control the cluster size. The number of VIDs allowed for a node can control the amount of meshing in the tree branches of the cluster.

Note that a node is aware of the cluster under which it has a VID as the information is inherent in the VIDs it acquires, thus a node has some intelligence to decide which VIDs it would like to acquire – i.e. it can decide to have several VIDs under one cluster, or acquire VIDs that span several clusters and so on. Moreover, a VID also contains information about number of hops it is from the CH, an attribute inherent in the VID length. This information can be used by a node to decide the cluster branch it would like to join based on the hops.

H. Inter-cluster Reactive Routing

This feature though not used in the work is described for completeness of the proposed architecture and its capabilities. Nodes bordering two or more clusters are allowed to join the branches originating from different CHs, and will accordingly inform their respective CHs about their multiple VIDs under the different clusters.

A node that has to discover a route to a distant node sends a ‘route request’ message to its CH(s). The CH then identifies the neighboring clusters based on updates from border nodes and forwards a copy of the ‘route request’ message to the border node, so that they can forward to the CH in the next cluster. The ‘route request’ message however has an entry for all the clusters that will be receiving the message, to avoid looping of the message. Thus the route request is not forwarded by all nodes, but only by all clusters and follows a path CH-border node-CH and so on.

When the CH of the destination node receives the route request, it will forward the route request directly to the destination node. The clusters forwarding the route request record the original sending node and the last cluster that the route request came from; this information is useful in forwarding the route response message when it returns. The destination node generates the route response and sends to its CH, which then forwards it back to the CH in the originating cluster and the source node along the same cluster path the route request took. Along the path back, all forwarding CHs will record the previous cluster and original sender of the route reply. The route between the sender and the destination node is thus initially set up as a sequence of CHs, but maintained as next cluster information. Mobility of
nodes does not impact the reactively discovered route, as long as the CHs exist. Note that movement of CHs also does not impact the reactive routes.

1) Robustness of the Reactive Routes

The route between nodes L in cluster 2 and A in cluster 1 while there are having an active sessions will be maintained at CH2 and CH1. If there were other clusters they would not maintain information for the route between the two nodes. Thus the reactively discovered route between L and A is maintained as a sequence of CHs and at the CHs as described earlier. The proactive route between L and CH2 and A and CH1 can change continually as the nodes move. Also the border nodes used between CH1 and CH2 to forward packet under the session can change, which change is recorded and maintained by the two CHs. Despite all the changes in the proactive routes, the reactive route which is the sequence CH2-CH1 does not change. They will change only if the CHs die. Thus the probability of a reactive route failure depends now on only two nodes as compared to the several numbers of nodes that normally define the reactively discovered path. With node mobility a single node movement in a path results in the path failure and rediscovery. In the proposed scheme as the reactive routes are a concatenation of the proactive routes between node-CH-border node-CH node and these proactive routes are dynamically updated as the nodes move, reduces the probability of the reactive route failure considerably.

I. Highlights of the Architecture

Under the related work section we highlighted several routing schemes, and frameworks that combined different types of routing algorithms and cluster based routing. From the meshed tree based clustering and routing scheme described thus far, it should be clear that the scheme adopts a proactive routing approach, where the proactive routes between CCs and CH in a cluster are established as the meshed trees or clusters are formed around each cluster head. Thus a single algorithm and through process of joining a cluster nodes automatically also acquire routes to the CH. There is flexibility in dimensioning the cluster in terms of CC in a cluster and the maximum hops a CC is allowed from a CH. The tree formation is different from other tree algorithms as a node is allowed to simultaneously reside in several branches, and thus allowing for dynamic adaptability to route changes as nodes move. This also enhances robustness in connectivity to the CH. We know of no work in the literature with such unique properties. Though multiple overlapped clusters have been discussed in the literature [15], [16], the proposed meshed tree cluster achieves this in a simple way.

J. Interworking of Modules in the Architecture

It is important to understand the interworking of the modules and their interaction with the directional antenna system. Hence, the directional antenna system is first described followed by the interactions among the modules and their use of the directional antenna systems.

1) Directional Antenna System

All nodes in the surveillance network are assumed to be equipped with four phased array antennas capable of forming two beam widths. One beam width is focused with an angle of 10° and the other is defocused with an angle of 90°. The defocused beams are used for sending broadcast packets, while the focused beams are used for unicast or directed packets. Each antenna array covers a quadrant (90°) and is independently steerable to focus in a particular direction within that quadrant in the focused beam mode.

We also assume that each node is equipped with a Global Positioning System (GPS) which is used for time synchronization and to provide node position. The latter information is used in a tracking algorithm to estimate the location of a receiver node, so transmitting nodes can direct their beams to the destination node.

2) Interworking Principles

The surveillance data collected by the nodes is passed to the routing module, which will decide on the route or VID to use to forward the data to the CH based on directions provided by the OC. The OC unit in this case decided on routes with the least hops. When there is a backlog in the packet to a particular destination the OC unit informs the scheduler to negotiate for more slots. The meshed tree cluster formation and its parameters are maintained by the OC unit. The unit also decides on the overlap and number of VIDs to be maintained, the cluster size and so on. The OC unit can monitor Physical layer parameters to decide on the routes, this feature was not used in this work.

Once the route has been decided, the node knows the address of the next hop node which will forward the packet. This information is then passed to the STDMA scheduler to schedule slots, taking as input the number of slots, slot time and control slots. This information is then passed to the MAC to create the frame and forward to the next node. Before forwarding, the MAC, locates the destination node position and controls the antenna array to transmit the packet using a directed beam.
IV. SCHEDULING AND LINK ASSIGNMENT

The VIDs carry link information between a pair of nodes that share a parent-child relationship. Thus a link assignment strategy was adopted in this work. The structure of the VIDs, also allows each node in a cluster to be aware of its neighbors due to the parent-child relationship defined by the VIDs. This allows a node to schedule time slots with its neighbors (parent or child) taking into consideration its current committed time slots to its other neighbors.

A. Scheduler Operations

The scheduling algorithm has to schedule time slots for (1) cluster formation after deployment of the UAV nodes, (2) subsequent cluster and route maintenance, and (3) data aggregation. It should also send updated schedules in a timely manner as network topology changes. For all of these operations different categories of time slots as described below were used.

![Distributed Scheduling Among Neighbors](image)

**Broadcast Slots:** Some slots are preselected as broadcast slots in which they announce their VIDs, location, and their acceptance into a cluster by the CH. Nodes individually schedule data slots in a distributed manner with their one-hop neighbors making the scheme truly distributed. The end

<table>
<thead>
<tr>
<th>Slot</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH</td>
<td>TX to B</td>
<td>RX to B</td>
<td>CTRL</td>
<td>TX to A</td>
<td>RX to A</td>
<td>TX to D</td>
<td>RX to D</td>
<td>TEMP</td>
<td>TX to C</td>
<td>RX to C</td>
</tr>
<tr>
<td>A</td>
<td>TEMP</td>
<td>CTRL</td>
<td>TX to B</td>
<td>RX to CH</td>
<td>TX to CH</td>
<td>TX to B</td>
<td>TEMP</td>
<td>TEMP</td>
<td>TEMP</td>
<td>TEMP</td>
</tr>
<tr>
<td>B</td>
<td>RX to CH</td>
<td>TX to CH</td>
<td>RX to A</td>
<td>TX to C</td>
<td>RX to C</td>
<td>TX to A</td>
<td>TEMP</td>
<td>CTRL</td>
<td>TEMP</td>
<td>TEMP</td>
</tr>
<tr>
<td>C</td>
<td>TX to E</td>
<td>TX to F</td>
<td>RX to E</td>
<td>RX to B</td>
<td>TX to B</td>
<td>CTRL</td>
<td>RX to F</td>
<td>TEMP</td>
<td>RX to CH</td>
<td>TX to CH</td>
</tr>
<tr>
<td>D</td>
<td>CTRL</td>
<td>TX to E</td>
<td>TX to G</td>
<td>RX to E</td>
<td>RX to G</td>
<td>RX to CH</td>
<td>TX to CH</td>
<td>TEMP</td>
<td>TEMP</td>
<td>TEMP</td>
</tr>
<tr>
<td>E</td>
<td>RX to C</td>
<td>RX to D</td>
<td>RX to C</td>
<td>TX to D</td>
<td>CTRL</td>
<td>TEMP</td>
<td>TEMP</td>
<td>RX to F</td>
<td>TX to F</td>
<td>TEMP</td>
</tr>
<tr>
<td>F</td>
<td>TEMP</td>
<td>RX to C</td>
<td>CTRL</td>
<td>TEMP</td>
<td>TEMP</td>
<td>TEMP</td>
<td>TX to C</td>
<td>TX to E</td>
<td>RX to E</td>
<td>CTRL</td>
</tr>
<tr>
<td>G</td>
<td>TEMP</td>
<td>RX to D</td>
<td>TX to G</td>
<td>TEMP</td>
<td>TEMP</td>
<td>TEMP</td>
<td>CTRL</td>
<td>TEMP</td>
<td>TEMP</td>
<td>TEMP</td>
</tr>
</tbody>
</table>

**Table 1 Sample Schedule Generated by the Distributed Scheduler**

current schedule, in a configuration (conf) packet, so neighboring nodes can listen and decide to join the cluster.

- **Directed Slots:** All other slots are used in a directed mode, where one node is transmitting using the directed beam to its listening neighbor. Directed slots can be assigned slots or temp (unassigned) slots.

- **Temp Slots** are used by nodes to negotiate for a common time slot for data transfer.

- **Assigned Slots:** Temp slots become assigned slots after a mutual negotiation by a pair of nodes. In the assigned slots control information for cluster and route maintenance, link maintenance (lnk_mnt) control packet generated by the MAC and data packets are sent and received. Assigned slots are unidirectional and are used either for transmitting (data-tx) or receiving (data-rx). If there are data packets to be sent in such slots, the control packets are sent first, followed by the data packets. At least one packet must be sent by a node during in a data-tx slot each frame to every neighbor that it is associated with. When there are no data packets to send, the MAC sends lnk_mnt packets to monitor the link status. The link will be dropped between two nodes if these transmissions are not maintained every frame. Unidirectional links that can only send or receive data but not both are not supported in this scheme. Acknowledgement of received packets and retransmission of unacknowledged packets are handled by the MAC, but only route requests, route replies, and data packets are acknowledged. Lnk_mnt packets are implicitly acknowledged when the neighboring node sends its own lnk_mnt back. Each explicit ACK contains a low and a high sequence number which represent the range of packets than are being acknowledged. If the sender of the packet does not receive the corresponding ACK by the following frame, it will attempt to resend the packet up to a maximum of three attempts. At that point the link is considered failed and the VID is no longer valid. Any queued packets are rerouted after the VID is dropped.

B. Link Assignment

The approval of a new node by the CH is an indication that the CC has both a physical and logical path towards the CH. Scheduling slots for the new node starts subsequent to its acceptance into a cluster by the CH. Nodes individually schedule data slots in a distributed manner with their one-hop neighbors making the scheme truly distributed. The end
It then forwards the registration request from node B towards the CH. During the next frame, node A will send a \textit{link\_mnt} packet to node B with the updated schedule. Thus a set of slots for transmitting and receiving between nodes A and B are decided. No other node’s schedule is taken into account unless it directly affects the current link between two negotiating nodes. Time slots are reused across different sets of nodes by taking advantage of the spatial separation between nodes.

The process of allowing a new requesting node a VID to reserve a data\_rx slot in which the parent node can transmit allows the parent node to resolve conflicts in case the suggested data\_rx slot is not available. If two nodes attempt to assign themselves the same data\_rx slot for a third node, the third node will accept the data\_rx allocation from the first schedule that it receives. When it gets the second schedule, it will not make any schedule changes and just send a link maintenance packet to the sender. The denied node will see the conflict and choose a different temp slot to allocate as a data\_rx slot. This also prevents any lost packets during the link establishment process. Since nodes choose their own receiving slots, but not transmit slots, there is certainty that the neighboring node is available during a transmission. Assigning data slots in this manner allows for dynamic asynchronous links.

For example if node A’s buffer indicates packet (to be sent to node B) accumulation beyond a threshold value, then in the next \textit{link\_mnt} packet, A can request node B to set aside $x$ data\_rx slots, where the value $x$ is capped to avoid one node taking up all available slots. Node B will respond with the updated schedule by setting aside the $x$ slots provided it has no such similar demands from its other neighbors. If there are similar demands, it will allocate slots proportional to the demands of its neighbors. The on demand allocation can result in increased number of data\_rx slots at B (to receive from node A) though the single data\_tx towards node A will be maintained unless changed by a demand. The tuning of the on-demand slots is executed every frame. If the amount of traffic being sent to node B decreases, the link will be reduced to having one data\_rx and one data\_tx slot again.

Table 1 is a sample schedule generated for the cluster in Fig 1. Nodes A, B, C, and D receive the initial configuration packet from the cluster head and schedule their data\_rx (RX) slots; 4, 1, 9, and 6 respectively. This decision is a random allocation of matching temp slots based on the sequence in which the configuration packets were received. The cluster head accepts these \textit{data\_rx} packets which were sent in the registration request messages of these nodes and sets the corresponding slots as data\_tx slots in its own schedule. It then allocates data\_rx slots to each of these nodes on slots 5, 2, 10, and 7. Node A receives a configuration packet from Node B and decides to use slot 6 as its data\_rx slot for receiving from node B. Node B then chooses slot 3 as the data\_rx slot for A. At the same time Node B receives Node C’s configuration and chooses slot 5 as its data\_rx slot. Node C selects slot 4 as the complementary slot. This is the same slot that the CH is transmitting to Node A, but due to the directional antennas there will be no interference. The process continues branching outward until every link has a pair of slots allocated.

**V. SIMULATIONS**

The performance evaluations of the surveillance network using the proposed solution was carried out using Opnet (version 14.5) simulation tool. All the processes explained above were modeled in Opnet. For surveillance data, each CC generated a 1 MByte file, which was then sent to the CH for aggregation. Normally UAVs travel in elliptical trajectories. In the models, we used circular orbits, to introduce more route breaks and thus stress test the solution. These circular orbits had a diameter of 20 Km (which defines the areas for each scenario), while the maximum transmission range was limited to 15 Km. The overlap between trajectories is seen in Fig. 4. A maximum of 5 UAVs were allowed in one circular trajectory, thus the UAVs were deployed over a wider area, which was covered with several trajectories. For example, in the 20 node scenario, there were four circular trajectories with slight overlap in their trajectories, to avoid physical network segmentation as shown in Figure 6. In the trajectories, the speed of the UAVs varied between 300 to 400 Kmph; hence, the different colors for the trajectories.

![Figure 6. Typical Deployment and UAVs](image)

The physical layer parameters were maintained invariant. Packets with 1 bit error rate were dropped and no \textit{Forward Error Correction} was implemented. In the focused beam mode the data rate is 50 Mbps and in the defocused mode the data rate is 1.5 Mbps. A single frame had 50 timeslots each of 4 ms duration and 0.5 ms guard time. These values were optimized based on our prior work [4, 5].

Due to the lack of similar published work and models in Opnet (the evaluation tool used) the performance of the presented solution is analyzed with respect to the performance goals stated for surveillance networks earlier.
namely success in packet delivery, and latency in packet and file deliveries. Included in the performance graphs are the overhead incurred by the MAC and routing protocols, and the average hops encountered during packet delivery, which is useful in explaining some results.

**Success Rate** was calculated as the percentage of packets received at a destination node with respect to the number of packets generated by the sender paired with that destination node.

**Overhead** for both MAC and routing was calculated as the percentage of control traffic to all the traffic in the network. This was determined only when data sessions were active. The bits contributing to overhead calculations was discussed earlier.

**Packet latency** was recorded as the end to end latency i.e. from the time the packet was sent by a sender node till it was received by the CH in seconds. File delivery latency was calculated similarly in seconds.

In each of the test scenarios, a certain number of nodes were randomly selected to send a 1 MByte file to the CH. These selected nodes sent the files simultaneously, thus stress testing the solution. Furthermore the number of sending nodes was increased to include all of the nodes except the data aggregation nodes, which is a highly stressful test scenario. Each test scenario was repeated with 20 different seeds (high prime numbers) and the results averaged over these seeds. The simulations were limited 20 runs in each case due to the stable outcomes noticed with different seeds.

### A. 20 Nodes Scenario

Figures 7A to 7C are the plots for the twenty UAV scenario with 4 clusters. The x axis in all plots shows the number of nodes that are simultaneously sending aggregation traffic, i.e., 1 MByte file to the 4 CHs. The number of sending nodes was varied from 4 to 16. In the last case all 16 CCs were sending a 1 MByte file simultaneously to the CHs.

With increasing number of senders, the success rate hardly dropped below 100%. This shows the efficiency of the scheduler to successfully schedule all the packets that are arriving simultaneously. The average hops recorded in graph 1 however shows a decrease when the number of sending nodes was increased. When 20 nodes were selected to send traffic they encountered an average hop distance of 1.8 hops; which dropped to 1.4 hops when all 16 nodes were sending traffic. This is because of the random way in which the sending nodes were selected. The average hops graph can be interpreted thus – the first four nodes that were selected were farther away from the CHs, but as more nodes were randomly picked they were closer to the CH. The impact of this is noticeable in the packet and file latencies recorded in graph B, which shows a decrease with increasing number of senders.

In Fig. 7B, the average packet latency recorded was less than 0.8 seconds. Acceptability of packets arriving at this latency depends on the criticality of the surveillance application. If an upper limit was specified then that could be used as a cut off to drop packets arriving late. The file delivery latency is only slightly higher at around 1.2 seconds, which shows that all packets in the 1 MByte file were transported from the data collection node to the aggregation nodes, i.e., the CH within the time.

Fig. 7C is the plot of a very important parameter as it shows the channel bandwidth used by the control traffic both by the MMT based routing protocol as well as the MAC protocol. The MAC and routing overhead were recorded to show the ratio of messages used for control purposes by two operations.

The MMT routing overhead was below 20% while the MAC overhead reduced from 10% when there were 4 sending nodes to less than 5% when there were 16 sending nodes. It should be noted that the MMT routing traffic also includes the cluster formation control traffic.

The MAC overhead shows a decrease with increasing number of senders, because when there are fewer data packets to send (with less senders) the MAC still sends maintenance packets, thus the ratio of control bits to the...
total bits that travelled the network, shows a decrease when there are more data packets in the network. The routing overhead records a very slight increase (around 1%) with increasing senders, which can be attributed to more route maintenance which will be triggered to correctly route the high amount traffic generated.

B. 50 Nodes Scenario

Figures 8A to 8C are the plots for the 50 UAV scenario with 10 clusters. The number of UAVs sending 1 MByte file simultaneously was varied from 10, 20 to 40. Thus in the case of the 40 senders, all CCs were sending 1 MByte files to the CHs simultaneously.

The success rate in graph A shows a slight drop to around 99.7 % as the senders increased, which shows the reliability in data transfer of the proposed solution and its scalability as the number of surveillance nodes and data sending nodes increased. The average hops which is plotted along with success rate graph does not show a linear decrease as in Fig 5 graph A. This is again attributed to the random selection in sending nodes. The first 10 senders were on an average of 1.7 hops from the CH, the added 10 senders for the 20 node case reduced the average hops to slightly above 1.5, and the last 20 senders brought the average hops to 1.5.

Figure 8B reflects the impact of the average hops in the packet and file delivery latency. There is drop when the senders increase from 10 to 20, this is because the average hops has a steep decrease from 1.7 to 1.5. However the average hops drops very slightly when senders are increased from 20 to 40 nodes, this and the fact that there is more traffic and more buffering by the nodes, the packet and file latency increase with increase in senders from 20 to 40.

The MAC and routing overhead in Figure 8C show a similar trend as observed in Figure 7. Though the number of nodes has increases, control traffic is calculated as a ratio of control traffic to total traffic in the network during the time that the files are being delivered.

C. 75 Nodes Scenario

Figures 9A, 9B and 9C are the performance plots for the test scenario with a total of 75 UAVs and 15 clusters, the number of sending nodes was varied from 15, 30 to 60. Hence again when 60 nodes are sending 1 Mbyte file it is the case of all CCs sending traffic to the CHs. The success rate dropped to around 98.7% with increasing number of senders - reflecting the robustness of the proposed solution and its scalability to increasing UAVs and increasing number of senders. The plot of the average hops again shows a decrease from 1.55 to 1.47 as the number of senders selected randomly to send the traffic to the CH was increased.

Figure 9B is the plot for the packet and file latency. The plot shows an increase because the change in the average hops was 0.06 as the number of senders was increased. The latency trends reflect the average hops trend. Figure 7C which is the plot of the MAC and routing overhead has a similar trend as noted for the 20 and 50 node scenarios.

Summarizing, the performance graphs indicate the high robustness of the proposed solutions to highly mobile and stressful MANET conditions. The continually high value of success rate despite the increase in the network size and the increase in the number of sending nodes indicate the reliability of the proposed solutions and its scalability. The packet and file latencies never exceeded 0.8 seconds and 1.2 seconds respectively in the three network setups. This indicates the robustness of the scheduling algorithm.

The overheads noted have similar trends and show very little difference as they were calculated as a ratio of the traffic in the network. The senders in each case were a quarter of the CCs, half of the CCS and the rest of the CCs. The control traffic increases with the increase in the number of nodes in a scenario, but as it is expressed as a ratio of all the traffic in the network including the data traffic, and due to the ratio of senders being consistent in all scenarios, this value can be noticed to be very close in all scenarios.
VI. CONCLUSION

Surveillance networks are critical tactical applications, and hence require special consideration during solution design. The primary goal in surveillance networks of UAVs is to collect the captured data reliably at few nodes, and with low latencies. In this work we presented a solution that uses an integrated approach where MAC, routing and scheduling are based off a single algorithm and use a single address. The design leads to a new MANET architecture that has performance advantages over traditional approaches and provides a low complexity yet robust and scalable solution.

The solution was evaluated in a UAV surveillance network of varying sizes of 20, 50 and 75 nodes. In each case the numbers of simultaneous 1 MByte file senders were increased from one quarter to one half to all of the remaining nodes besides the aggregation nodes. This was a highly stressful test case. The results achieved under such stress situations were very good. The drop in reliable and timely delivery was very low as the numbers of senders were increased. These results thus validate the use of the solution to such critical tactical applications.

The proposed solution has several tunable parameters as the MMT algorithm allows such capabilities. These capabilities are optimizing the cluster size, determining the number of VIDs to allow for nodes, decisions by nodes to join different clusters or have several branches under one cluster, length the tree branches and so on. The architecture has the feature to allow considering applications criteria and physical layer constraints while determining the paths. The information could be used for improved system design. This is due to the structure and positioning of the communications layer between the applications layer and physical layer. The solution is transparent to layer 3 and hence will not be impacted during IPv4 to Ipv6 transition or to any other layer 3 protocol. It can thus interwork with existing systems and their protocol structures.

ACKNOWLEDGMENT

This work was partly by funded by AFRL, Rome NY under contract no. 30822, and partly from ONR.

REFERENCES


[38] Hua W., Martin N., Al-Mousa Y., Orakwue C. and Shenoy N., “A Distributed Scheduler for Airborne Backbone Networks with Directional Antennas”, ComNetS, Bangalore India 2011