

# Secure Local Data Aggregation and Delay Tolerant Dissemination in VANETs

Cesar Ghali   Sky Faber   Kyle Benson   Quan Nguyen  
 Group Number 1, Computer Science Department  
 University of California, Irvine  
 {cghali, fabers, kebenson, quann1}@uci.edu

**Abstract**—Recent advances in wireless networking and sensing technologies have brought about research trends in enabling vehicles to communicate with each other while driving. These vehicular ad-hoc networks (VANETs) could automate the vehicles themselves and/or greatly decrease dangers inherent with driving by warning each other of hazards and reacting to them. Due to the limited bandwidth and connectivity of wireless mediums and the large quantities of data being generated in such a system, novel techniques must be explored for aggregating and disseminating useful information so that it can be made available in a reliable and timely manner. In this paper, we propose a method of securely, quickly, and locally disseminating aggregate information about crash events and other hazards in VANETs. We show our technique’s efficacy through simulation in NS-3 [1].

## I. INTRODUCTION

Our problem of interest is distributed event detection in wireless sensor networks. Specifically, we are interested in detecting vehicle crash events in vehicular ad hoc networks (VANETs). When a vehicle is in any kind of collision it will affect both nearby vehicles and vehicles that may be rapidly approaching the crash site. Therefore, warning these other vehicles about such an event could enable them to take immediate evasive actions to mitigate the possibility of another crash. Each vehicle that witnesses the crash, as well as the one(s) involved in it, should each take part in notifying others so that as much detail about the nature of the event and possible affects can be cataloged as possible. Because many vehicles may be present during such an event, they may flood the wireless channel with information if it is not properly aggregated and disseminated to possibly affected entities.

We developed a protocol in which nodes broadcast aggregated packets to nearby nodes in order to forward them along to parts of the network that haven’t yet been notified. This quickly disseminates data about the event occurrence to vehicles that may be affected by the collision, whether from extra traffic congestion or possible road hazards/debris. At the same time, messages from different participants that witnessed the crash are aggregated together into a single packet, while still preserving information about which nodes detected the crash. This decreases overall bandwidth consumption while still delivering the same information about crash witnesses and potential dangers. By exploiting mobility patterns of VANETs, our protocol quickly covers nearby nodes with this information in a variety of traffic conditions, as will be shown in the discussion about simulation results.

As an additional consideration, security also becomes an important issue in this scenario. Detection of an event may have real world repercussions, such as causing vehicles to reroute or completely stop and so our protocol must ensure that no attacker can fake a crash event in order to cause such a reaction. We briefly catalog known attacks of this type and discuss possible protections against them as well as how our protocol takes them into consideration.

The remainder of the paper is structured as follows. The second section discusses related works in data aggregation and dissemination. In the third section, we discuss our protocol in detail. We present our simulation environment and results from the tests on it in the fourth section. In the final section, we conclude with a discussion about applications and future work.

## II. RELATED WORK

In this section, we present several works related to ours in the domains of data aggregation and dissemination. Much of it is related to wireless sensor networks but can be extended for use in VANETs.

### A. Data Aggregation

In the domain of wireless sensor networks, much recent research has focused on developing aggregation schemes to minimize the amount of power consumed by radio transmissions in order to extend the life expectancy of power-constrained devices that typically rely on batteries. Many approaches focus on minimum spanning trees (MST) or similar structures because this will require the minimum amount of total power across the whole network. As shown in [2], this is because the power required to transmit a radio message is directly proportional with the distance between the nodes. If sensor nodes can decrease the power of their radios to reach the next hop with minimal energy spent, the MST will give the optimal aggregation of data.

In addition to decreasing energy consumption, these approaches can also be used to reduce communication overhead, thereby freeing up bandwidth for other important messages. This becomes particularly important in VANETs where the vehicles are envisioned as constantly communicating detailed information with each other. Here, we discuss aggregation techniques applied to two possible topologies: static and dynamic (mobile nodes). The latter is likened to VANETs.

1) *Static Topologies*: In [2], the minimum-energy aggregation problem is solved by creating a MST in the network that has constraints on the maximum degree of any one node and the maximum number of hops from any node to the sink. The degree and hop bounds are in place because too high of a degree would put too much load on an individual node and too many hops would require too long for the information to reach the sink. As the determination of these optimal structures is NP-hard, this paper proposes several heuristics for approximating them. One of these heuristics, which was determined the best of the proposed three by their simulation results, iteratively builds the MST by adding the next node whose edge has the least weight (power consumption) as long as it doesn't violate any of the constraints. If it will, the node is joined to a different parent whose incident edge weight may be higher (thereby making the spanning tree only an approximate MST) but who does satisfy the degree and hop constraints. If no such parent exists, the parents are recursively checked for the possibility of choosing a new parent for them such that the constraint on the original node can be satisfied.

Another problem which is important in Wireless Sensor Network (WSN) is to schedule the aggregation [3]. The problem is if a node hears two messages at the same time then it can not receive the messages because of message interference. Therefore, in a given graph  $(V, E)$ , we need to find the schedule which consists of  $n$  sets of vertices  $S_i$  which satisfies the following properties:

- $S_i \cap S_j = \emptyset, i \neq j$
- $\cup_{i=1,n} S_i = V - \{r\}$
- Data are aggregated from  $S_k$  to  $V - \cup_{i=1,k} S_i$ .

To solve this problem, we use the leader election [4] algorithm to build the rooted spanning tree. Based on that we will distributed rank each node based on its level (how far from the root) and its identity. Along the computation, we will compute the number of lower-ranked neighbors for each node. Using the computed information, we will build connected dominating set (CDS) [5]. Building CDS is the crucial step in different algorithms including the aggregation scheduling problem as it is considered the virtual backbone or spine of wireless ad-hoc network. Finally, we construct the schedule with linear time latency bound on network diameter and maximum node degree.

The above methods consider data aggregation in the network but do not address the issue of time-constrained data transmission. The authors of [6] address this problem by setting a deadline by which all of the aggregated data should reach the sink. This deadline is propagated towards the leaf nodes, with each intermediate aggregating node setting its deadline as a function of the number of nodes represented by the aggregated packets it will receive. Aggregation takes place when either all of the child nodes send their information or the timer expires, thereby ignoring any data not received by this deadline. This deadline is increased over time but is decreased if the sink does not receive the aggregated messages in time. In this manner, the timers will converge over time to a point at which the sink's deadline is respected but the preceding deadlines are late enough to maximize the number of nodes whose data is

received by the sink. The sink must therefore choose a deadline that allows every node in which it is interested to have its data aggregated. Otherwise, it may misjudge the actual state of the environment being monitored. Such a scheme is particularly important in time-sensitive applications, such as in VANETs where aggregated information may prevent dangerous crashes.

2) *Dynamic Topologies*: Some previous works build structures similar to MSTs but focus more on the ability of wireless sensors to broadcast data across an entire region, which may change as the network topology evolves. A common approach is to cluster nodes and elect leaders for these clusters so that each of the nodes can send their data to the leader, which will then aggregate these packets and send them towards the sink. This allows a subset of the nodes to transmit messages at longer ranges so that the other nodes can conserve power. Some such methods also allow for non-static topologies in which nodes may move, join, or leave the network.

The authors of [7] use the above approach to propose a system that allows for changing the aggregation node so as to distribute the burden of increased power consumption that comes from holding this role as well as tolerate mobile nodes. The network is broken into a number of concentric rings and sectors within those rings based on the density and size of the network. Nodes within each sector transfer their data to the node closest to the center of this sector so it can be aggregated and transferred to the next inner ring. The aggregated messages continue to be further aggregated and forwarded, eventually reaching the inner-most ring where they are received by the sink. By varying the locations of these sectors and rings, different aggregation nodes can be chosen and mobile nodes can be followed so as to maintain a balanced number of nodes within each sector. This balance is particularly important in mobile networks as changing node densities could cause issues of scalability in a static structure.

While the previously mentioned data aggregation methods build a structure over which communication is carried out, [8] takes an unstructured approach. Here a technique deemed Flow Updating is used to perform both broadcast- or gossip-based aggregation without the need of a spanning tree. Here aggregation functions are calculated by exploiting the symmetry of flows in the network. Flows in this context represent portions of the total aggregate value, rather than representing the capacity of an edge as in the traditional graph setting. A node's local calculation is based entirely on its initial state, and the flow entering the node. Nodes can update their flows with one another either by broadcasting to their neighbors, or using random unicast (gossip). The authors show that this converges considerably faster than older structured approaches. The key idea is that as flows are updated throughout the course of the algorithm, because of directional symmetry, the total aggregate value does not change. Furthermore, flow updating messages are idempotent, a new message simply replaces an old flow with the new value. Because of this idempotence, Flow Updating has good convergence properties, even in the event of link failure. Unfortunately, this method still relies on a static topology, and many rounds of updates, and would not work well with mobile nodes. Furthermore, the authors mention little about calculating aggregate functions that are

not based on sum.

Authors in [9] describe a technique for aggregating and disseminating travel times in VANETs. Their protocol performs hierarchical aggregation based on the distance from the origin of the data, in this case traffic times of a specific route. The intuition is that nodes physically further away from the origin receive coarser versions of the data. As a vehicle approaches the location, finer-detailed information becomes available so that it may plan a specific route. Nodes in the network share information via broadcast, but aggregate timings according to a hierarchy.

They first define a method for specifying routes with different granularity. A map of a physical street or highway is overlaid with landmarks; each landmark defines a specific point of interest on the map and is defined (semi-)manually. Landmarks are then organized hierarchically such that each level in the hierarchy has fewer and fewer landmarks. Times are calculated for landmarks close to each other on the same level in the hierarchy. Nodes in the network measure travel times between two landmarks at the lowest level, and broadcast these times to peers they meet in the same vicinity. As the information travels further from the origin, the nodes compute the distances between landmarks closer to the top of the hierarchy. Nodes receive only as much information as necessary to make accurate routing decisions.

The [9] paper defines very similar functionality as required for our problem. However, traffic data is generated more frequently and requires dissemination to virtually all nodes in the network. While we may not have these constraints, we may be able to apply similar techniques. Unfortunately, the authors simulations are done with additional infrastructure placement and does not compare with any other aggregation techniques.

The aggregation technique used in [10] is based on Flajolet-Martin sketches. In such sketches a set of elements is represented by a string of bits using a hash function as follows:

- A string of bits is first initialized to zero.
- Each element of the set is fed into a hash function.
- The output of the hash corresponds to an index in the bit array which is then set to 1.

When two sketches need to be combined, an OR operation is performed. Moreover, a TTL-based sketch can be used to express lifetime for each entry. Each bit is replaced with a positive integer that is initialized from a maximum value and then is decremented every time the sketch is broadcasted until the bit reaches zero. When two such sketches are aggregated, the maximum value of a specific index is taken, indicating the newest observation of that element in the set. The authors used the proposed Flajolet-Martin-based dissemination to survey the empty spots in a parking lot. Several entities are scanning the parking lot and marking the empty spots using the previous algorithm. At the end, all the collected sketches are aggregated. This algorithm can be used in other applications such as aggregating the existence of some event into a small string of bits data structure.

## B. Data Dissemination

In networks where nodes represent agents within a system, nodes must often disseminate pertinent information to

interested parties quickly and reliably. Constraints such as available bandwidth and network connectivity often encumber these goals and so novel approaches must be taken to ensure delivery of accurate data. Precision and recall are two metrics that should be taken into considerations while designing such a protocol. Precision is defined as the rate of people getting the content they are interested in to the total number of people getting that specific content, while recall denotes the quantity of interested entities that get the disseminated content. This section focuses on delay-tolerant dissemination and a study of its application to VANETs.

Flash dissemination requires the distribution of data content to a large number of recipients in a very short time. Such a technique should take into consideration the unpredictable topological changes in the network, scalability to serve large network sizes and the ability to adapt to heterogeneity in the network setup. In [11], the authors designed an intelligent flash dissemination protocol, CREW, that outperforms its counterpart protocols in different network setups. CREW is a gossip-based dissemination protocol that is fault-tolerant, decentralized, reduces network overhead.

Since gossip-based dissemination protocols lack the ability of good scalability in large network scenarios, CREW designers modify the push content concept into a pull content concept. In the later approach, each node selects a node at random and exchanges with it the list of all the chunk-ids that it needs. Then, the selected node sends a random chunk to the requester node. This approach reduces the number of broadcasts which fits large networks, and eliminates receiving multiple copies of the same chunk from different gossiping nodes.

Delay Tolerant Networks (DTNs) are new network structures composed of different nodes connected in an ad-hoc fashion. Packet delivery in such networks are based on the best effort model. In which, a node broadcasts the packets it receives to all its neighbor hoping that they will eventually be delivered to their destination, even if an intermediate node must wait for a network connection to be established with the next hop.

On the other hand, a data dissemination protocol for DTNs, Habit is proposed in [12]. It is to build two directed graphs and share them among all the nodes in the DTN network. The first graph is called the regularity graph which models the physical connections among DTN nodes, and the second graph is called the interest graph which models the shared interests between the nodes. All the nodes in the network share their information about other nodes in order to build an approximate vision of the full graphs of the network. When a node has a message or content to disseminate, it consults with the network representation that it has in order to obtain the shortest paths that the content should traverse to all interested node.

Content dissemination in mobile ad-hoc networks (MANETs) is quite different than in static environments described earlier. Patterns of interspersed network connectivity and mobility can be exploited to facilitate the spread of information. VANETs are a special case of MANETs in which the nodes are assumed to move in relatively straight lines, at similar speeds, and with another group of nodes

usually travelling in the opposite direction very nearby. These specific characteristics are leveraged in [13] to develop, compare, and contrast different techniques for dissemination of aggregated data in VANETs.

This particular work advocates for forwarding of messages through a broadcast mechanism as opposed to a unicast to a neighbor of the node. This is because the data about many vehicles within the network that are close to each other is aggregated together so the broadcast storm problem is mitigated by only having a subset of the vehicles disseminate this information throughout the local network. The authors consider three different schemes for forwarding this information:

- Forwarding to vehicles moving in the same direction
- Forwarding to vehicles moving in the opposite direction on the other side of the road
- A combination of both techniques (forwarding both directions)

By simulating a VANET environment, the authors discovered that the second technique works best for their scenario. This is because they aggregate exact locations of vehicles on the road, compressing their coordinates in order to fit within a single message. This results in more information contained within each message and therefore more lossy compression is required. By exploiting the mobility of the traffic moving in the opposite direction, these messages can be ferried more quickly and so contain less information, thereby requiring less compression. In actuality, the bi-directional approach is more suitable if less accurate vehicle locations are tolerated because there are more avenues through which the messages can be forwarded, more intermediate nodes whose data can be aggregated into each message, and the compression does not affect the desired data quality.

### III. PROPOSED PROTOCOL

In this section, we discuss the details of our protocol and the decisions that led to this design. Please make special note that we do not assume the presence of any dedicated infrastructure, be it roadside radios or cellular towers. While such infrastructure may increase the efficiency and effectiveness of our protocol, it may not always be available and, in the case of cellular networks, may not provide the latency that such fast dissemination requires. Therefore, we aim to accomplish our goal only through the ad-hoc network of vehicle-mounted radios wherever possible.

Also, we refer to messages in one of three ways, although they all use the same format:

- 1) A message originating from one of the vehicle(s) involved in the crash, which we refer to as the *victim*
- 2) A message originating from one of the vehicle(s) that *witnessed* the crash
- 3) A message that is being forwarded by any node in the network, possibly including a victim or witness

#### A. Assumptions

The design of our system assumes that several underlying hardware and software components have already been developed. First, we of course assume that a fleet of vehicles have

all been equipped with sensors and radios to intelligently transmit information related to driving conditions between each other. These vehicles may or may not be autonomous, as such a system could also be used to warn drivers of upcoming road hazards and aid in safely navigating passed them. Second, to facilitate crash event detection, the sensor inputs, such as video and radar, are assumed to be enough information to intelligently determine that a crash has taken place nearby the vehicle. While the details of actually implementing such event detection is outside of the scope of this paper, we envision it to involve computer vision and machine learning algorithms that associate sudden changes in motion of nearby vehicles to a likely crash event. Third, for the purpose of preventing spoofed crash data, we assume that the vehicle sensors are secure and tamper-proof. When they detect that the vehicle has crashed, similar to the way airbag sensors work currently, they must generate unforgeable data that can be authenticated by other vehicles that overhear it during transmission. This raw crash data is assumed to be included in a message from the victim as a means of guaranteeing the authenticity of an event. Without such sensors, the security of this system is in jeopardy as it relies solely on the participants' honesty for detecting only real events.

#### B. Dissemination

When a crash occurs, any nearby vehicles that detect it should attempt to notify others that will be approaching it immediately. By broadcasting the message to all vehicles within transmission range, more will be aware of the situation than were previously. These newly aware vehicles will forward the information by broadcasting it again to all their neighbors. In this manner, each possibly impacted vehicle will be notified if an ad-hoc path exists between it and the original detecting vehicle.

To address the issue of missing ad-hoc links, whether due to vehicles being out of range or the message being corrupted in the channel, these broadcasts are repeated several times. We noted that messages could, and indeed in our simulations did, propagate as far as possible in the local ad-hoc sub-network without reaching all vehicles possibly affected. We decided that each vehicle should broadcast the message a total of five times, with one second between each iteration, in order to ensure delivery. Additionally, we added a random wait time so that each node would immediately re-broadcast a received packet, but then wait for a random period between 0.5 and 1.5 seconds before beginning the one-broadcast-per-second regimen. The intuition of this optimization is that nodes near to each other will receive the packet and therefore likely rebroadcast at the same time. If they re-broadcast their subsequent messages at the same time, they are likely to be reaching the same additional nodes. By instead staggering these re-broadcast times slightly, while still maintaining an interval of one second within a node, it is more likely that rebroadcasts will reach nodes that are only very briefly in range of this cluster. It may also serve to reduce channel contention and spread the network transmission more evenly to prevent spikes in channel usage.

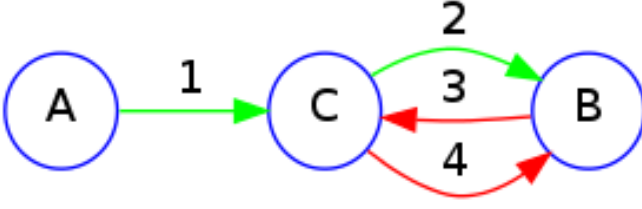


Fig. 1. *Broadcast Loops*: Edges are labeled in order of message transmission. After receiving message 2, node B forwards it back to C as message 3, who will then forward it back to B as message 4 and so on...

We chose these values because it appeared feasible for a vehicle moving in the opposite direction to enter and exit transmission range quickly enough that one or two corrupted packets might result in it not receiving the message. Utilizing these vehicles allows for the message to be propagated behind the event more quickly and effectively, as described in [13]. Five seconds provides adequate time for vehicles on the opposite side of the highway to move far enough away from the event scene that they can give ample warning to nearby vehicles. Furthermore, we assume that after five seconds dedicated infrastructure such as cellular networks or satellite uplinks will be able to notify any nodes that have not been contacted already.

As shown in Figure 1, broadcasting can easily create loops in a network that lead to an infinite amount of traffic being generated. To mitigate these *broadcast storms*, which flood the network with many messages, and prevent them from saturating the channel, we include mechanisms for limiting the forwarding of the messages. Specifically, we include a time-to-live (TTL) and a unique broadcast ID. A TTL is an integer number that represents the maximum number of *hops* that a message should take, that is the number of times it is forwarded by intermediate nodes before it should no longer be forwarded any further. The TTL was set to 10 in order to prevent the message from persisting too long and reaching too many nodes that were not interested in its contents. If we set it to much less than 10, the messages did not seem to reach far enough. Also, with an effective range of approximately 100 meters, a TTL of 10 meant that the message should reach vehicles up to one kilometer away, which was our goal distance as anyone within this extent would approach the crash site very soon. Similar to AODV [14], the unique broadcast ID, a combination of the source node's network address and a unique sequence number that it generates locally, is used to prevent loops. A node ignores any messages that it has already seen instead of forwarding them. Repeat messages are identified by storing the broadcast ID of recently seen messages and comparing the broadcast IDs of the received message to this collection.

Each message, including the originals, will include the ID of the vehicles that witnessed (and participated in) the crash as well as the location of the event in order to facilitate authenticating the information and avoiding the scene, respectively. Because each vehicle that witnesses the crash begins disseminating messages, there could be a large number of duplicate messages detecting the same event that are treated

as distinct for the purposes of forwarding. This is where data aggregation improves the efficiency of the system.

### C. Aggregation

Instead of allowing each original message to be disseminated separately, we decided to aggregate them together into a single message whenever possible. This saves a lot of network overhead by greatly reducing the number of packets that must be transmitted to disseminate the same information. When a node receives more than one message that pertains to the same crash (as identified by the victim ID), the IDs of the witnesses are combined together into a single packet before forwarding. In this way, each packet contains a subset of IDs corresponding to nodes that witnessed or participated in the event. When a node receives such a packet, it may learn more than one ID as compared to only learning one without any aggregation.

In the case of multiple rebroadcasts, nodes will continuously update the packet they are broadcasting during each interval. Therefore, if a node learns of another witness, or learns the ID of any crash victim(s), the next packet broadcasted will include this information. This does not mean that the counter of how many rebroadcasts should be sent is reset, however, as this could prolong the life of the message to a wasteful point. In this manner, forwarded messages typically contain only one or a small number of IDs initially. As the message propagates further, aggregation will take place and each message will tend to contain more information about the event.

## IV. SIMULATION AND RESULTS

In this section of the report, we describe the implementation and simulation of the proposed protocol. We used NS-3 [1] to simulate the proposed crash model and measure some performance parameters. A set of classes are integrated with NS-3 to simulate the existence of a highway with several lanes and vehicles driving along them. The mobility model used in our experiment is the ns3-highway-mobility model presented in [15]. Vehicles are created and initiated to drive on the highway at a starting point and with some velocity constraints such as minimum and maximum allowed speed. The existing model was very basic with vehicles mostly driving the same speed and on the same lane, which renders it unrealistic. We improved this model by adding some randomness of the vehicles speed limits and allowing them to pass each other by driving on different lanes.

We also add the existence of the crash event. We initialize the simulation with two additional parameters, the crash vehicle and the time of the crash. The first parameter specifies the ID of the vehicle that will crash at a time specified by the second parameter. When a crash happens, the specified vehicle will stop and send a "I have crashed" message (only one time) and the neighbor vehicles will start disseminating the event information to the rest of the network.

It is worth mentioning that at each time step of the simulation, we write the position of each vehicle and the messages transferred to log files so we can verify the correctness of the implementation. The vehicle positions log file is used to feed a Java-based tool to vitalize the highway with all the vehicles

driving on it during the whole simulation time. The transferred messages log file, is used to verify the validity of the dissemination and aggregation techniques used in our proposed protocol. We were able to visualize that the predefined vehicle is really having a crash and all other vehicles are successfully passing it on the road.

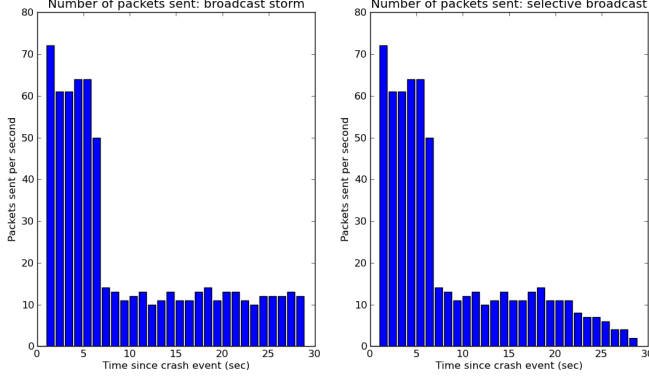


Fig. 2. Broadcast storm versus selective broadcast

Figure 2 shows a comparison between the two cases of using broadcast storm, that is always forwarding a broadcast message even if it has already been seen, and selective broadcast. We notice that in the case of using selective broadcast, the number of packets transferred in the network starts decreasing after 20 seconds. The reason is because most vehicles have already received the crash message and start ignoring all message that are already received. On the contrary, vehicles that use broadcast storm keep on forwarding crash messages even if they already received the same messages before. This explains why after 20 seconds from the crash event, these vehicles continue forwarding messages.

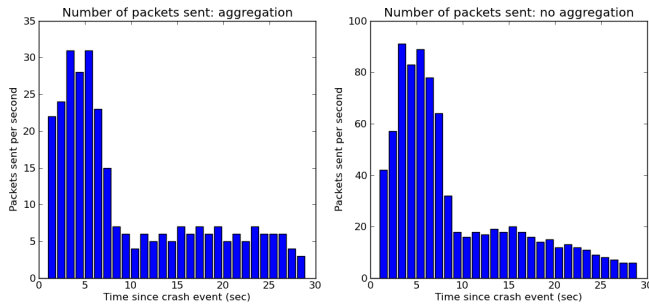


Fig. 3. Aggregation versus non-aggregation

Figure 3 illustrates the case where vehicles use aggregation versus not using aggregation. One can notice that the number of messages sent when aggregation is not present is approximately three times more the number of messages sent when aggregation is used.

In Figure 4, we demonstrate the different dissemination patterns when low, medium and heavy traffic are present. It shows that after roughly 25 seconds of simulation, all of the

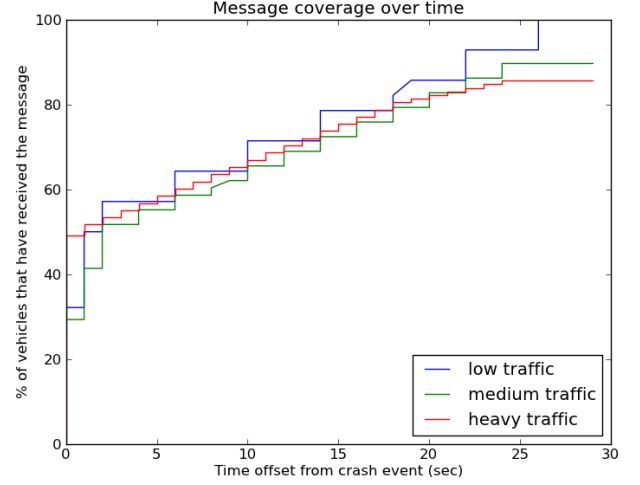


Fig. 4. Message coverage over time

vehicles that are suppose to have received the crash message already have. At time 0 seconds, or right after the crash occurs, we notice that 50% of the vehicles have received the crash message in the heavy traffic case whereas only 30% of them received the message in low and medium traffic. It is worth mentioning that the vehicles receiving the messages are the ones that are close to the crash area which means that the more recipients there are, the better the locality of the dissemination. As time progresses, the message coverage curves for the three traffic density cases start the converge together. Moreover, it can be noticed that the cases where heavy traffic exists have smoother message coverage curve due to the presence of a lot of vehicles close to each others. Thus vehicles carrying the message do not need to travel for very long before coming into the vicinity of other vehicles to transfer the message to.

## V. SECURITY CONCERNS

In VANETs, the safety messages including traffic information messages (e.g. traffic condition), general safety-related messages (e.g. collision avoidance) and liability-related messages (e.g. when accidents happen) are critical to the security of drivers, their physical safety, and law enforcement. Therefore, VANETs security should be considered at the design phase, rather than ad-hoc fixes after deployment. In the next paragraphs, we will describe different types of attacks on VANETs [16], [17] and security requirements to prevent such attacks.

### A. Attacks on VANET

In this section, we will describe different attacks on VANETs. The attacks are classified according to the following criteria:

- Membership: Insider(I)/Outsider(O)
- Motivation: Malicious(M)/Rational(R)
- Method: Active(A)/Passive(P)
- Scope: Local(L)/Extended(E)

For each attack we will use the following notation: Membership.Motivation.Method.Scope, e.g., I.M.A.L. means that the

adversary is an insider, behaves maliciously and mounts an active attack in a local area.

- *Jamming attack*: In this attack, the adversary wants to interfere with radio transmission. One simple jamming attack is the adversary continuously sends radio signal on a valid channel, thus preventing the legitimate vehicle from sending the radio signal. This is O.M.A.L.
- *Hidden vehicle*: This attack works as follows: whenever a vehicle broadcasts a warning message, the adversary will pretend to be in a better position to broadcast the message. Thus the message originator would stop broadcasting message to save traffic. The adversary in turn doesn't broadcast the warning message further. Basically, the origin of the warning message is hidden by the adversary. This is O.M.A.L. attack.
- *Tunnel attack*: This attack exploits the fact the GPS is disabled when the vehicle is going through the tunnel. The adversary can send fake information of GPS to the vehicle when it leaves the tunnel. This attack works as GPS is not authenticated. This is O.M.A.L. attack.
- *Wormhole attack*: This is a collusion attack when we have 2 vehicles far apart but have a fast communication channel. The message from one vehicle is tunneled to another remote vehicle then the remote vehicle broadcasts the message. This way, an authenticated message is broadcast to another location. Essentially, there is a wormhole between 2 vehicles which helps broadcast a message to another location far apart. This is an I.M.A.E attack.
- *Bush telegraph attack*: This attack exploits the fact that if in every step there is a small error then in the end, the result would be totally different. The idea is if the error in each step is small, i.e., within the tolerance threshold then the neighbor would accept it. To implement this attack, the adversary needs to compromise a set of vehicles. This is I.M.A.L. attack.
- *Sybil attack*: In this attack, a compromised vehicle claims multiple fake identities. This way, he can create fake positions, attract all the messages to him, etc. This is I.M.A.L.

#### B. Security requirements for VANET

To prevent the attacks described above, the following security requirements are needed:

- *Authentication*: The sender identity must be authenticated as well as the sent messages.
- *Message Integrity*: There must be a mechanism to prevent (or detect) messages integrity, i.e., detect the change of message by unauthorized vehicle.
- *Availability*: VANET must have the capability to prevent denial-of-service attack.
- *Non-repudiation*: The sender can not deny the message that he sent. This is really important in case of accidents, i.e., the vehicle that causes the accidents can not deny its responsibility.
- *Privacy*: This is really important as the drivers do not want the companies or government to trace all of his sent

messages or locations.

Designing VANET to satisfy all the described security requirement is not a simple task. However, the fundamental step toward the security of VANET is to have a secure tamper resistant hardware to authenticate the vehicle and to sign the message by private key. Due to space constraints, we will not go into detailed solutions for all the listed security attacks. Instead, we will describe a solution for the authenticating disseminated data which is relevant to our project.

Itani et al. proposed in [18], a Secure Bundle protocol, Sec-Bundle, to authenticate disseminated data sent through the networks. The proposed protocol is based on using an authentication structure, called a Bloom filter [19]. Such structure is used to verify the membership of a set of elements. In other words, using a Bloom filter, it is feasible to verify if a specific element belongs to a set or not. First, the source feeds all the data chunks in the Bloom filter, signs it with its private key using Identity-Based Encryption (IBE), and disseminates it. When an intermediate node, say a router, receives a data chunk, it compares it against the membership structure to see if it belongs to the set or not. If yes, the chunk is forwarded to the next hop, or dropped otherwise. Membership verification at intermediate nodes level is beneficial to drop bogus packets as soon as possible. The wireless link on which such packets are injected is the only link that carry them. The previous authenticated dissemination protocol can be used as a base infrastructure to address our problem.

#### VI. CONCLUSION

In this paper, we proposed a secure method for quickly aggregating and disseminating information about crash events in VANETs. The protocol was implemented in a simulation environment to show its efficacy. We determined that our method is effective at reaching all local nodes very quickly even in a variety of traffic conditions. By extending the ns3-highway-mobility project, we have also furthered VANET simulation research as our modifications are publicly available for others to benefit from. To conclude the paper, we discuss some optimizations and possible future work as well as alternative applications for our protocol.

#### A. Optimizations and Future Work

Although we did not have time to implement them, we considered a few optimizations and extensions to our protocol. One involved adapting the forwarding and re-broadcasting mechanism to the number and location of other nodes nearby. If a node is located within the center of a node cluster that it knows has already received a message, by overhearing each of them broadcast the message or an acknowledgement of it, it does not need to rebroadcast the message itself as it will not reach any additional nodes that the others cannot already. This could prevent some excess broadcasts and reduce the load on the network, although it would add a large degree of complexity to the protocol and require careful negotiation between the nodes. Not only would each node need to keep track of what nodes around it have received the message, it must also know that there are no other nodes present that



have not yet. This could be accomplished by overhearing other protocols that would likely be utilized in VANETs, such as range-finding and movement negotiations. However, it is still difficult to say with certainty whether another node has entered one's broadcast radius or not.

We also considered adjusting the time between rebroadcasts, and even the actual number of rebroadcasts, based on the local node density of the network. Our simulation results indicate that event data can be disseminated far quicker with higher traffic densities due to the decreased space between each vehicle. This indicates that some parameters might be relaxed in order to decrease the number of messages in the network and frequency of rebroadcasting while still maintaining the same dissemination speed.

Although we did not consider utilizing dedicated infrastructure in our system, it could certainly be used for improvements. If a VANET-equipped road has a series of radio towers along it to aid communication between the vehicles, they could be used to more rapidly spread information without adding load to wireless channel. By using higher-powered road-side antennae, each tower could cover more vehicles in a shorter amount of time than ad-hoc communications. Furthermore, the towers could pass the message further back the road to reach vehicles coming towards the crash by communicating with each other either wirelessly or through a wired link. A cellular network could be used in a similar manner, although, as previously discussed, the latency on cellular towers makes them less suitable for our particular application. Even so, they might be used to notify a large area of vehicles about the event and perhaps stop the ad-hoc dissemination of it once the dedicated infrastructure becomes aware and can take over the dissemination process. A combination of all three (ad-hoc, road-side, and cellular) networks may also prove to be the most effective and flexible when properly managed.

### B. Other Applications

While our technique was intended for detection of and notification about crash events in VANETs, it could be applied in other domains as well. For example, it could be used for notification of other impending hazards, such as road debris, weather, or simply traffic congestion. Some modifications to the protocol details could make it suitable for exchanging basic information as well. For example, our method of aggregating messages could be used to efficiently determine the approximate speed of traffic and negotiate which vehicles should be in which lane and how fast for them to travel. It could also be used to notify very close vehicles (perhaps with a TTL of 2 or 3 hops) about one's intentions, such as changing lanes or making turns.

In addition to VANETs, this protocol might be applied in other areas, such as wireless sensor networks. Indeed, we discussed how data aggregation is often used in this domain, especially for the purpose of reducing power consumption. One can envision a system in which sensor devices and/or sinks only have intermittent connections, either due to movement or perhaps powering down their radios and only periodically coming on-line to listen for any interesting data. In

this case, the application of staggered repeated broadcasts may help ensure data delivery and even reduce power consumption if the nodes effectively coordinated.

Outside of the wireless domain, our protocol might find use in a peer-to-peer system. By using the technique of staggered rebroadcasting, peers might send some important information to nodes that either join the network or move around within it. This could help to address the issue of churn that is so common in these systems, ensuring that each node is aware of the pertinent data and allowing them to more efficiently distribute it as more pieces become known through aggregation.

### REFERENCES

- [1] "ns-3." [Online]. Available: <http://www.nsnam.org/>
- [2] H. Cheng, Q. Liu, and X. Jia, "Heuristic algorithms for real-time data aggregation in wireless sensor networks," in *Proceedings of the 2006 international conference on Wireless communications and mobile computing*, ser. IWCMC '06. New York, NY, USA: ACM, 2006, pp. 1123–1128.
- [3] B. Yu, J. Li, and Y. Li, "Distributed data aggregation scheduling in wireless sensor networks," in *INFOCOM 2009, IEEE*, april 2009, pp. 2159–2167.
- [4] I. Cidon and O. Mokryn, "Propagation and leader election in a multihop broadcast environment," in *Proceedings of the 12th International Symposium on Distributed Computing*, ser. DISC '98. London, UK, UK: Springer-Verlag, 1998, pp. 104–118.
- [5] P.-J. Wan, K. M. Alzoubi, and O. Frieder, "Distributed construction of connected dominating set in wireless ad hoc networks," *Mob. Netw. Appl.*, vol. 9, no. 2, pp. 141–149, Apr. 2004.
- [6] J. Y. Choi, J. W. Lee, K. Lee, S. Choi, W. H. Kwon, and H. S. Park, "Aggregation time control algorithm for time constrained data delivery in wireless sensor networks," in *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*, vol. 2, may 2006, pp. 563–567.
- [7] Y. Zhu, R. Vedantham, S.-J. Park, and R. Sivakumar, "A scalable correlation aware aggregation strategy for wireless sensor networks," *Information Fusion*, vol. 9, no. 3, pp. 354–369, 2008.
- [8] P. Jesus, C. Baquero, and P. Almeida, "Fault-tolerant aggregation by flow updating," in *Distributed Applications and Interoperable Systems*, ser. Lecture Notes in Computer Science, T. Senivongse and R. Oliveira, Eds. Springer Berlin / Heidelberg, 2009, vol. 5523, pp. 73–86.
- [9] C. Lochert, B. Scheuermann, C. Wewetzer, A. Luebke, and M. Mauve, "Data aggregation and roadside unit placement for a vanet traffic information system," in *Proceedings of the fifth ACM international workshop on VehiculAr Inter-NETworking*, ser. VANET '08. New York, NY, USA: ACM, 2008, pp. 58–65.
- [10] C. Lochert, B. Scheuermann, and M. Mauve, "Probabilistic aggregation for data dissemination in vanets," in *Proceedings of the fourth ACM international workshop on Vehicular ad hoc networks*, ser. VANET '07. New York, NY, USA: ACM, 2007, pp. 1–8.
- [11] M. Deshpande, B. Xing, I. Lazardis, B. Hore, N. Venkatasubramanian, and S. Mehrotra, "Crew: A gossip-based flash-dissemination system," in *Distributed Computing Systems, 2006. ICDCS 2006. 26th IEEE International Conference on*, 2006, p. 45.
- [12] A. Mashhadi, S. Ben Mokhtar, and L. Capra, "Habit: Leveraging human mobility and social network for efficient content dissemination in delay tolerant networks," in *World of Wireless, Mobile and Multimedia Networks Workshops, 2009. WoWMoM 2009. IEEE International Symposium on a*, june 2009, pp. 1–6.
- [13] T. Nadeem, P. Shankar, and L. Iftode, "A comparative study of data dissemination models for vanets," in *Mobile and Ubiquitous Systems - Workshops, 2006. 3rd Annual International Conference on*, july 2006, pp. 1–10.
- [14] C. Perkins and E. Royer, "Ad-hoc on-demand distance vector routing," in *Mobile Computing Systems and Applications, 1999. Proceedings. WMCSA '99. Second IEEE Workshop on*, feb 1999.
- [15] H. Arbab and M. Weigle, "Highway mobility and vehicular ad-hoc networks in ns-3," in *Simulation Conference (WSC), Proceedings of the 2010 Winter*, dec. 2010.
- [16] M. Raya and J.-P. Hubaux, "Securing vehicular ad hoc networks," *J. Comput. Secur.*, vol. 15, no. 1, pp. 39–68, Jan. 2007.
- [17] X. Chen, K. Makki, K. Yen, and N. Pissinou, "Sensor network security: a survey," *Communications Surveys Tutorials, IEEE*, vol. 11, no. 2, pp. 52–73, quarter 2009.



- [18] W. Itani, A. Tajeddine, A. Kayssi, and A. Chehab, "Slow but certain wins the race: authenticated bundle communication in delay tolerant networks," in *Proceedings of the 6th ACM workshop on QoS and security for wireless and mobile networks*, ser. Q2SWinet '10. New York, NY, USA: ACM, 2010, pp. 90–97.
- [19] B. H. Bloom, "Space/time trade-offs in hash coding with allowable errors," *Commun. ACM*, vol. 13, no. 7, pp. 422–426, Jul. 1970.