A Data Intensive Reputation Management Scheme for Vehicular Ad Hoc Networks

Anand Patwardhan, Anupam Joshi, Tim Finin and Yelena Yesha
Department of Computer Science and Electrical Engineering, UMBC, Baltimore, MD 21250
{anand2, joshi, finin, yeyesha}@cs.umbc.edu

***Abstract*— In vehicular ad hoc networks individual vehicles can help each other locate resources and establish trustworthiness under highly dynamic conditions, lacking any centralized trust authority. To ascertain the accuracy and reliability of data aggregated in a distributed manner, we present a reputation management system for such networks that enables devices to quickly adapt to changing local conditions and provides a bootstrapping method for establishing trust relationships where only a few may exist *a priori*.**

Our scheme considers cooperativeness and accuracy of peer-provided data as two aspects of trust when evolving trust relationships and managing reputations. We use an epidemic data exchange protocol that incorporates reputation and agreement to ensure high reliability of data and stimulate proactive collaboration above and beyond stipulation, to enhance availability and reliability of data. We present preliminary simulation results which demonstrate the effectiveness of our data intensive reputation management scheme.

I. Introduction

Mobile ad hoc networks (MANETs) present the capability of providing local information in near-realtime, e.g. road closures, current traffic conditions, road conditions, etc. Such information has high utility only in a particular area and is time sensitive, making it important to be able to quickly ascertain its reliability. For most categories of such information we consider that softer security guarantees based on reputations of the participating entities are sufficient for practical purposes, in the absence of trusted third-party authentication.

Current systems providing location based services based on GPS receivers or mobile telephony, lack detail and timeliness in the information they provide. In the near future, sophisticated sensors embedded in the urban infrastructure will provide updates on current local conditions and directory services to locate nearby resources. Utilizing the mobility of vehicles and guided by their movement patterns, we seek to leverage the connectivity provided by multi-hop ad hoc networks to provide a convenient and effective way to disseminate reports of local events in order to improve their detail, availability, and timeliness. Our focus is on a broadcast data dissemination model in which anchored resources carousel current conditions encapsulated in data units, and vehicles in transit help propagate them. Validation of data units is achieved by consensus from multiple sources or via direct communication with a trusted source.

This research was supported by NSF awards IIS 0209001 and CNS 0203958

We propose a context-aware reputation management system that provides a bootstrapping process to build trust-relationships and stimulates proactive collaboration, to achieve a decentralized data management scheme that provides a high degree of reliability. Our focus is on enabling mobile devices to maximize their efficiency in locating, retrieving, and verifying personalized data. We utilize persistent identities, frequency of encounters, and a known set of anchored trustworthy sources to serve as nucleating points for building trust relationships with previously unknown devices.

We evaluate the effectiveness of our scheme in providing reliable, and timely data to the consumer devices while minimizing response time and cost, with improved reliability. The burden of collecting relevant information and verifying the data is placed on the consumer mobile devices. We present our simulation based on synthetically generated mobility patterns that conform to legal movement of cars along roads within an area of Washington DC, in order to achieve a sense of realism.

II. Motivation

Incidents like crashes, vehicle breakdowns, sinkholes, etc., lead to disruption of the normal flow of traffic. Such events when detected, should be quickly propagated for timely adaptation by affected vehicles. Existing methods of propagating traffic information like broadcast media (radio, TV) and dynamic message boards lack detail and personalization. Also, current mechanisms are centralized, i.e., event reports have to percolate back to some collection center and then rebroadcast over local radio or TV channels. Such information is usually periodically available and is often delayed. Moreover, notification of such traffic incidents often lack a detailed description of the event, which is likely to mislead the driver, or belated – leaving no time to adapt or fewer choices to adapt. Effective adaptation to traffic conditions will depend on the ability of Traffic Information Systems (TIS) to provide detailed, reliable, timely, and localized information to vehicles likely to be affected.

Two factors primarily determine what quality and detail of the data about a traffic event/incident can be reported in a timely manner, viz. (i) availability of sensors in range of the location of the event, and (ii) reliable and timely mechanisms

of propagating the data to potentially affected vehicles. Consequently, the extent to which existing TIS infrastructure can monitor such events is limited to areas where such sensors exist. Providing a pervasive infrastructure to monitor traffic and provide – localized and personalized information, as a free service may not be cost-effective.

In contrast, mobility of vehicles practically assures that traffic events leading to a congestion or other hazard, will have other vehicles nearby ensuring coverage with their own built-in sensors. Furthermore, MANET connectivity presents the capability of quickly and efficiently disseminating safety and hazard warnings and thus enable a timely choice of the most optimal driving routes based on current conditions.

III. Related Work

Incentive based collaboration has been proposed in [3], [15]. Nuglets rely on a security module that uses tamper-proof hardware, whereas the Sprite systems relies on a centralized Credit Clearing Service and a software protocol to ensure fair sharing of bandwidth.

Michiardi and Molva [9] have proposed a game theoretic approach to evaluate the cooperation enforcement mechanisms in mobile ad hoc networks. In their approach energy conservation (battery power) is considered to be the primary reason for node selfishness, and each node seeks to maximize a utility function. Nodes are assumed to be rational i.e. nodes will behave only in selfish interest but malicious behavior for intangible benefits is not considered.

Srinivasan *et al*. [13] propose an analytical model on similar lines seeking an optimal operational point between cooperation and non-cooperation in relaying sessions for other devices. They too address the issue of optimized use of constrained energy resources to maximize device lifetimes. Their focus is on providing an analytical model for attaining an optimal operation point for the MANET and assume that authentic information about most other nodes in the MANET, e.g. the energy class acceptance rates, and other parameters involved in computing the utility function are available. Malicious behavior is however discounted, i.e. nodes are considered to act only in selfish interest and not for intangible benefits. The computation of such an optimal operation point depends on the assumption that sufficient information about the system is available. They acknowledge the need for a distributed mechanism to reliably acquire and disseminate all such required information. However, normally only partial information is available, furthermore judging its accuracy is limited by the reputation information available for the device providing it.

In realistic situations getting timely, accurate, and reliable information from unverifiable sources is difficult and remains a challenging problem. With lack of centralized systems to provide reliable data or provide security mechanisms. Reputation systems seek to maintain updated and correct reputations in a distributed manner based on observed behavior and recommendations from others. Several reputation systems have been proposed that are applicable in peer-to-peer and MANET environments e.g., [11], [7], [8] that provide trustworthiness metrics i.e. softer security guarantees when using second-hand information.

The analytical models like [13] and game-theoretic approaches like [9] provide insights on using reputations and incentives to promote cooperation and fair sharing. The focus of previous approaches has been on session based interactions, whereas the data interactions in our scenarios are predominantly based on those triggered by epidemic updates. In our scenario, ad hoc connectivity merely provides connectivity to the source once an information source has been discovered. We seek to promote collaborative behavior at the *application level*, over and above regulation, i.e. beyond mere conformance to the communication protocol specifications. Such collaborative mechanisms are necessary and justified when it is possible to identify a set of identities that reciprocate. To be able to identify such a workable set of identities we use local landmarks and context markers and use cooperation scores as incentives to reciprocate.

We assume that the battery power is no longer a problem, e.g. devices like cars where battery power is not a limitation. However the reliability, availability and quality of the data provided and the cooperation/collaboration offered by other devices in finding information is more important.

IV. Proposed Approach and Design Objectives

A. Trust and Reputation Management

Certain kinds of data are characterized by short-term utility and time-sensitiveness esp. traffic conditions, movement patterns, etc. and often the cost of encryption and decryption is not justified. Moreover some of the deployed sensors reporting the data are likely to lack the capacity to encrypt the data they are reporting.

Also, those capable of it, will still lack physical security increasing the possibility of compromise. Thus, behavior and the data being provided should be continuously monitored to be able to detect faulty or compromised entities.

In our approach we model two aspects of reputation, viz. cooperativeness and fidelity/reliability of provided data. Various cryptographic methods can be used when needed to provision the requisite security properties. For example, digital signatures ensure integrity and non-repudiation but do not provide any guarantee or assurance of the quality or reliability of the data itself. The quality of the provided data is chiefly dependent on the trustworthiness (and hence reputation) of the providing source itself. Cryptographic techniques can be meaningfully used in conjunction once such trust relationships are established.

Thus, to be able to build trust relationships we must first ascertain the cooperativeness of the other devices and the quality of information being providing by them, and maintain a history of such past experiences to evolve trust. Experiences from directly observed behavior are the most reliable method, which are further augmented by recommendations from other trusted entities to increase the scope and knowledge about trustworthy entities in the vicinity.

We assume that secure and persistent identifiers like cryptographically generated addresses [10] are used as means of identification for mobile devices. These are self-generated

addresses, and hence obviate the need for a trusted centralized distribution mechanism.

Further, we assume that the nodes in the network are capable of verifying that unique identities belong to distinct participants (thus preventing Sybil attacks [4]). Several schemes have been proposed to address the threat of Sybil attacks [4], [5], and are beyond the scope of this paper.

Reputation management involves efficient storage of identities and past experiences concerning those identities. Positive or negative experiences may be stored, based on satisfactory completion of transactions, fulfillment of expectations, or some other form of verifiable fiduciary action. In a distributed reputation management system, the privileges accorded to an identity are predetermined by the perceived reputation of an identity. Consequently, it is a fundamental requirement to protect reputation systems from misinformation and coercion by malicious (colluding) devices, either seeking to falsely boost their own reputations, or seeking to falsely malign others.

Existing reputation systems have focused on rewarding successful completion of transactions e.g. in peer-to-peer systems (structured or unstructured networks) the search mechanism itself is assumed to be correct, the security perspective is on prohibiting free riders who consume resources yet contribute nothing in return. Consequently these systems provide specific deterrence mechanisms – or preventive measures that discourage cheating. In case of MANETs, cooperation in creating and maintaining routes – and forwarding packets (relaying sessions) for other devices is necessary to achieve a stable network. Further, it is necessary to stimulate nodes to cooperate sufficiently in order to achieve reasonable data throughput [13].

Selfish interest lies in maximizing limited (battery) lifetime by not relaying packets for others. Similarly, altruism or collaboration in storing data for others is also against selfish interest. However in serendipitous environments where locating timely and reliable data is difficult, proactive collaboration becomes a necessity. We have shown the effectiveness of collaborative querying in our previous work [12]. Though existing reputation systems provide the fundamental basis for evaluating trust and fostering cooperation, they do not help improve the reliability or availability of data.

The cost of verification of data depends on the available sources of data. Since we consider data to be trustworthy either if the source itself is deemed trustworthy, or if there is agreement amongst copies from distinct multiple sources. Thus the cost of evaluating data depends on how much is known about the available data sources. Hence, the reputation management system should ideally have a high success rate when evaluating identities.

It is thus important to be able to group devices of similar interests and using distinguishing characteristics available from the available context or from prior experiences, in determining which identities and reputations to store.

Buchegger and Boudec [2] have presented work on a robust reputation management system. More sophisticated approaches on reputation management involve mechanisms to prevent against collusion and false praise or accusations. We seek to achieve similar capabilities with a lower cost and by disclosing limited reputation information.

We choose a simplistic reputation management approach to test our hypothesis and focus on the ability and effectiveness of stimulating proactive cooperation yet seek to make the reputation management system system resilient to attack.

B. Data Content and Dissemination

We assume that the data being served locally is context specific and its utility is limited to a particular area and thus we focus on increasing its availability and reliability only in that particular area. Devices may be able to learn about the information sources from farther off, however they need not worry about the timeliness or availability. Context specific data needs to be delivered and consumed in near realtime for it to be useful. Moreover, devices need to be able to locate reliable sources and be able to verify/provide metrics on the quality of the data (timeliness, reliability, trustworthiness, and integrity).

C. Suitability of the broadcast model

We choose a broadcast model for data dissemination, where anchored resources continuously carousel current local conditions, whereas mobile devices monitor these feeds. This approach is suitable for its simplicity and the advantage of being highly scalable, since no request-response mechanism is involved with the original sources themselves. Other devices can query each other for data. This mechanism is sufficient and well-suited for scenarios where other devices caching portions of past streams can serve as secondary sources for the same data. To prevent flooding the entire network the data streams are limited by hop-limits and lifetimes and context (which helps in caching only locally relevant information).

Data sources are anchored and provide various kinds of realtime context specific information including local traffic conditions, weather, road conditions, parking availability etc. Each anchored data source provides its context information like latitude,longitude position, time (including that for validity of data and a proposed decay – hop count).

D. Data Collection and Verification

Mobile devices need realtime or near-realtime context specific information. This information can be cheaply propagated by anchored resources broadcasting local updates and other mobile devices in the neighborhood can collaborate to propagate and collect such information of interest. This leads to the issue of the reliability of data if it is retrieved from an unknown source. The risk of data corruption/fabrication or corruption can be mitigated by availability of multiple secondary sources and the ability to verify data. Data received directly from the primary source is trivially reliable/trustworthy. As shown in fig. 1, as a device moves further away from the primary/reliable source of the data, the reliability of the data is expected to decrease. Our primary focus is in enabling a data management mechanism which provides assurance on the quality of the data received even when the primary source of the information is not in direct radio-range.

We seek to investigate the effectiveness of collaborative methods by consumer mobile devices to collect and propagate reliable information, as opposed to using *a priori* trusted sources only. The search for particular kinds of information can

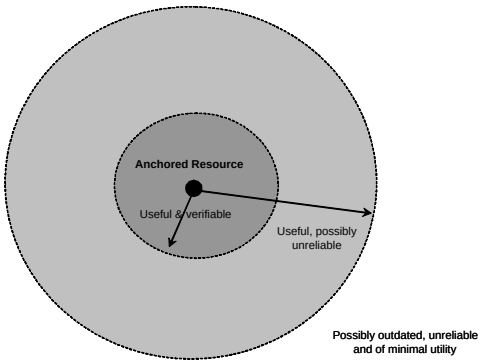

Fig. 1. Decay of Trust with distance

be pipelined, the quality of information can be improved as we approach the primary/reliable source of information. Further, performance can be improved by organizing groups of devices of similar interests to collaborate and to serve as secondary sources/caches of information that improve the availability, reliability, cost and response times. The price paid collaborative caching and searching is in terms of trust evaluation, reputation management, data caching and communication.

We wish to identify an area where the data relayed via intermediaries from some anchored resource is useful and reliable. This will be helpful for the devices to determine the trustworthiness of the relayed information. Moreover in most cases since the utility of the data is limited to a particular area, the need to verify it is limited to only the area of its applicability. The ability to explore information sources in adjacent areas gives devices a headstart to lookup information and verify it when they arrive in the area of utility.

We assume that the queries being made – or the kind of data to be fetched is known beforehand. Thus we primarily seek to address the following: (i) content location (locating and receiving data streams in near-realtime), (ii) meeting realtime requirements (low response times), (iii) efficient and scalable operation, data dissemination and retrieval, and (iv) verifying the authenticity and integrity of the data received.

V. System Model

Our focus is on modeling pervasive environments in urban/metro areas, though it is conceivable that similar approaches will be applicable in other pervasive environments. Urban areas can be considered to be composed of roads and intersections (city blocks). Buildings and other physical obstructions not only limit/constrain mobility they may also affect connectivity, however we currently focus only on constrained mobility. The participant devices composing the MANETs are the resources embedded in the environment (e.g. sensors, parking meters, traffic cameras etc.) and the consumer devices (e.g. car computers and other mobile devices capable of communication). Persistent identities are necessary for associating reputations, and we assume that such identities are available for all the participating devices. Further, we also utilize the persistent identities of the fixed resources to identify and markup geographic context, which then forms the basis of personalizing reputation management systems.

VI. Reputation Management

A. Device Classification

Since in these scenarios the objective is to obtain reliable information and to verify trustworthiness, faulty nodes are counted as malicious nodes since distinguishing a faulty node from a malicious node is of no significant benefit to the individual device. Validation of data units (cached streams) provided by other devices are used to evaluate trust and building reputations. When providing a cached stream to others, the device should validate it first. The device is held responsible for the data that it chooses to provide to others. As such, if data is unverified, there is a prospect of propagating falsified data. Devices that are vulnerable/susceptible to falsified information due to corrupt reputation information or subversion, themselves are equally harmful as malicious devices. Hence devices that provide portions of cached streams are held responsible for those – making them accountable for verifying data before propagating it further. Trust is then a measure of combination of observed and acceptable behavior and associated history of cooperation.

Each node maintains device identities classified into one of: (i) Encountered, (ii) Observed, (iii) Cooperative, and (iv) Malicious, such that each set is mutually exclusive and the perceived ranking is first determined by category followed by the level of cooperation.

The set of identities maintained by individual devices is based on one or more personalization features, e.g. (i) frequently visited locations (context markers), (ii) mobility patterns, and (iii) information categories. Thus individual devices can choose to maintain reputations about devices that have common interests which have provided relevant and useful information in the past.

B. Data Intensive Reputation Management Protocol

Individual devices observe the behavior of other encountered devices and interact with them in order to locate and retrieve information and also to exchange reputation information. Hello messages contain a trace digest of the recently seen anchored devices. This enables individual devices in locating relevant data and identifying opportunities for data exchange. Segment messages are periodically broadcast by the anchored resources and contain updates on the current local conditions for that area. Segment messages are also used by individual devices to answer queries by other devices.

The trace digests in the hello messages provide indicators of which contexts the sender was in and enables recipient devices to minimize redundancy in epidemic updates when querying for new information.

The algorithm in table I is used to process each segment message received. A segment is denoted by $d^{t_i}_{s,p}$, where t_i denotes the timestamp, s denotes the original source of the segment and p denotes the proxy providing it. Unless data source is provided by a trusted source $s \in \mathcal{T}$, where $\mathcal{T}$ is the set of trusted devices. $\mathcal{V}$ is the set of cached data

```
ProcessSegmentMessage ( d^{t_i}_{s,p} )

    if s ∈ T and (s, t_i) ∉ V
        V ← V ∪ {(s, t_i), d^{t_i}_{s,p}}
        Validate( U, d^{t_i}_{s,p} )

    elseif s ∉ M
        if (s, t_i) ∈ V
            Validate(d^{t_i}_{s,p})
            return
        elseif verification session U_{s,t_i} exists
            U_{s,t_i} ← U_{s,t_i} ∪ {d^{t_i}_{s,p}}
        else
            create verification session U_{s,t_i}
            U_{s,t_i} ← U_{s,t_i} ∪ {d^{t_i}_{s,p}}
        CheckConsensus(d^{t_i}_{s,p})
```

```
Validate(d^{t_i}_{s,p})

    if d^{t_i}_{s,p} = d^{t_i}_{s,s}
        Promote(p)
    else
        Demote(p)
```

```
CheckConsensus(d^{t_i}_{s,p})

    if | P |> AGREE_MIN
    for some P ⊂ U_{s,t_i}
    s.t. d^{t_i}_{s,m} = d^{t_i}_{s,n} for some m, n ∈ P
        V ← V ∪ {(s, t_i), d^{t_i}_{s,m}}
        for each p ∈ U_{s,t_i}
            Validate(d^{t_i}_{s,p})
        remove verification session U_{s,t_i}
```

TABLE I

SEGMENT VALIDATION ALGORITHM

segments that have been verified. $\mathcal{U} = \{((s_i, t_j), \{d^{t_i}_{s,p}\})\}$ where (s_i, t_j) uniquely identifies a segment by its source s_i and timestamp t_j. $\{d^{t_i}_{s,p}\}$ is the set of all the versions of the same segment received so far. A parameter $\mathcal{AGREE}_{MIN}$ is used to determine if agreement has been reached for a value of a data segment. After the data value for a segment has been validated either by receiving it from the original source, or by achieving consensus, the actual or agreed value is used to `promote` or `demote` the cooperation index and reputation of the providing devices in the verification session for that segment. The functions promote and demote are used to provide positive and negative reinforcement to existing perceptions of trust in those devices.

The algorithm basically works as follows. Unless a source is trusted, the segments provided by it are cached in verification sessions $\mathcal{U}_{s,t_i}$(uniquely identified by the original source s and timestamp t_i. When a segment is received from one of the *a priori* trusted sources belonging to $\mathcal{T}$, it is added to the set of verified set of segments $\mathcal{V}$. Further, any corresponding verification session for that particular segment is closed (since now the correct value is known). Any segments received from sources in the set $\mathcal{M}$, i.e. classified as malicious (by reputation) are discarded immediately. Every time a segment from an unknown source is received, and added to an existing or new verification session, the verification session is checked to see if a set of distinct sources agree upon the value of the data. If the condition for agreement based on the value of the parameter $\mathcal{AGREE}_{MIN}$, that value is deemed to be true; the the majority of agreeing sources in that verification session are accordingly promoted, the remaining disagreeing sources are demoted, and the session is closed. Verification sessions are discarded if they are not closed by a validating segment from a trusted source or by agreement, within a timeout period.

VII. SIMULATION

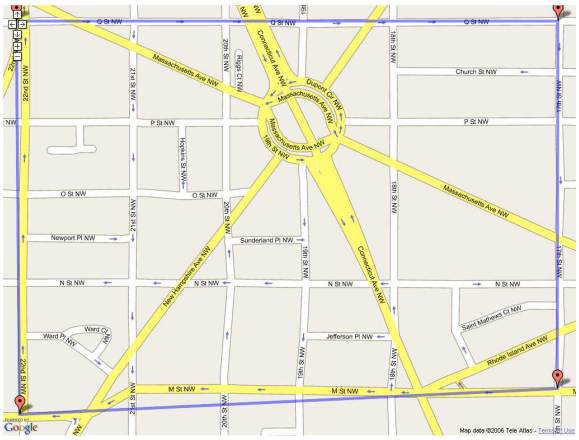

Fig. 2. Street map of the Dupont Circle area in Washington, DC (modeled area is demarcated).

Evaluation and simulations of networking protocols have so far primarily used Random Waypoint Model, which however do not realistically model vehicular traffic. To simulate our proposed approach and data management model for cars in a metro area we chose to use a more realistic scenario of modeling the actual road network in a 700m by 900m area around DuPont Circle, Washington DC, as shown in fig. 2. Using geocoding services we obtained the latitude and longitude of each of the road intersections in selected area and used the Haversine formula [1] to convert the latitude and longitude positions to Cartesian coordinates. Further, we represented this as a directed graph taking into account one-way streets, to impose a realistic model. The resultant directed graph represents an approximation of the city block geometry and valid mobility patterns. Valid transitions include traversal of the graph along the directed edges from some source within the simulation area.

We use Glomosim v2.0.3 for our simulations [14]. We simulate movement of vehicles in this area with the above mobility constraints, with number of nodes varying from 50 to 200, and transmission range of 100m (802.11b has a transmission range of 300ft = approx. 91m). The random waypoint model

was adapted with constrained mobility (allowed only legal directions of motion as per the street map) with speeds ranging from 15 m/s to 25 m/s and pause times from 0 to 30s, for a simulation period of 30 mins.

One anchored resource is placed at each of 38 major intersections, that periodically broadcasts data segments. These fixed (primary) sources act as landmarks and also broadcast information streams. Other mobile devices listen for all available data streams of interest from these sources and also serve to propagate this information further.

Bad nodes(other than primary sources, only in relaying information) will try to provide falsified/misleading information, however information can later be verified when the primary source or other trusted secondary source becomes available. We measure the availability of data, the reliability of data, and the response times of data using these ad hoc networked devices propagating information from the primary sources. Anchored resources monitor device activity within their radio-range and record devices that they have heard from (seen). Mobile devices on the other hand remember which landmarks they have seen and what information has been gathered so far.

The identities of the 38 anchored resources are implicitly known and trusted by all the mobile devices, however the mobile devices do not have any trust relationships amongst each other. Each of the 38 anchored resources send out a new data segment every 10 seconds, each minute thus has upto 6 unique segments. Each unique segment is broadcast once every second for a 10 second interval. Thus, 0 to 6 unique segments can be received by a device in each of the one minute interval.

Fig. 3 (a) and (b) show the contour plots of information foraged by one mobile device as it travels around the Dupont Circle area for a duration of 30 mins, when there are total 50 mobile devices and none of them are faulty or malicious (i.e. relayed data is never falsified or faulty). The Y-axis shows the anchors ordered by their designated identifier, and X-axis shows time in minutes. Fig. 3 (a) shows a contour plot denoting the amount of information received directly from the 38 *a priori* trusted anchored resources, as it navigates the area around Dupont Circle. Fig. 3 (b) shows the total information including data from unknown sources possibly not validated. The additional information in Fig.3 (b) also shows that the mobile device is benefiting from information exchanges with other mobile devices, that would not have received if restricted to collecting information solely from pre-enumerated trusted resources.

Fig. 4 shows a contour plot of number of segments collected from unknown sources that were later validated. This demonstrates the amount of information that was available at hand even before the device received it from the actual source. When the device is able to make trusting decisions based on its current situational awareness and reputations it has built using the data intensive reputation management protocol described in table I, it will be possible to evaluate trustworthiness of this information without relying on the actual source.

The bar charts in fig. 5 show the averages of number of segments received from trusted sources (TD), number of segments validated (VD), number of segments invalidated (ID) and number of segments that remained unverified/timed-out

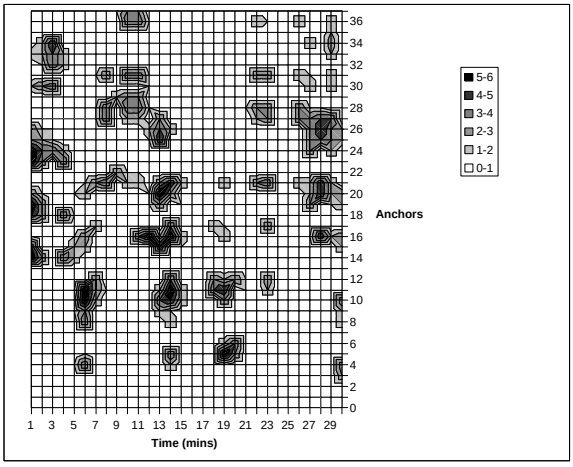

Fig. 4. Information available via unknown sources, later validated

(TM) in the 30 minute simulation. The bar charts of averages with 0%, 30% and 60% bad nodes respectively, for four simulation setups with 50, 100, 150 and 200 mobile devices.

It can be observed that the average number of segments declines with the increase in total number of mobile devices. This is expected since the amount of contention increases leading to a decline in overall throughput [6]. Moreover increased contention can, in part, be attributed to the constrained mobility patterns, which cause the mobile devices to converge at road intersections (unlike the Random Waypoint Model).

It can be seen that number of invalidated packets increases and validated packets decreases as the number of bad nodes increases. However the resilience and merit of the collaborative data exchange is highlighted by the fact that even though the average number of validated segments can be seen to be declining (approx. from 750 segments for 0% bad, down to 450 segments with 60% bad, for 50 devices). This is the data received from unknown sources that was received prior to encountering the actual source itself, but was later validated (and was thus a forewarning). In case of 0% bad nodes almost half of the data received from trusted sources was initially received from secondary sources. Also the segments that could not be verified are roughly half in number of those received from trusted sources. This data is in fact represents the potential for increased scope of situational awareness beyond the radio range and traveled path of an individual device that can be filtered by the reputations of the providers. In the current version it is possible to detect bad nodes using the data they provide, depending on whether it was validated or invalidated.

VIII. Future Work and Summary

Urban areas and surface transport networks are increasingly seeing a proliferation of embedded and networked sensors that can interact with other mobile devices. Moreover vehicular technology is rapidly incorporating advanced sensor systems that can monitor the health of the car and moreover allow it to interact with other such mobile and anchored devices. Efficiently distributing crucial data to vehicles in a timely and

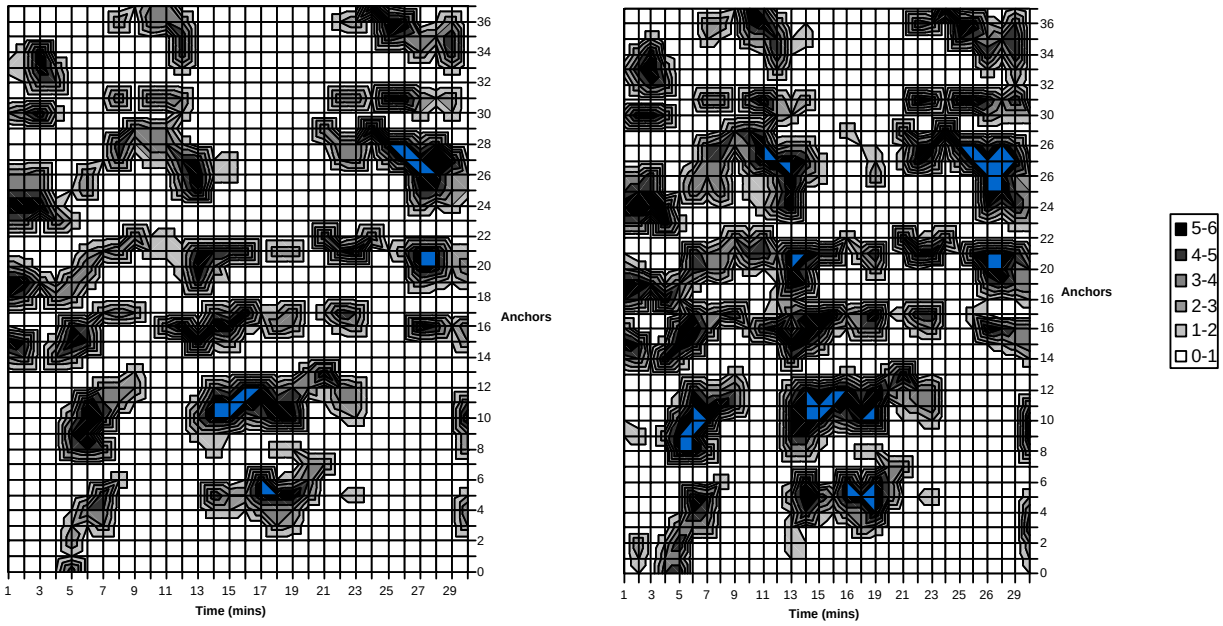

Fig. 3. (a) Contour plot of information cached via trusted sources only (b) Contour plot of information cached via trusted and other secondary sources

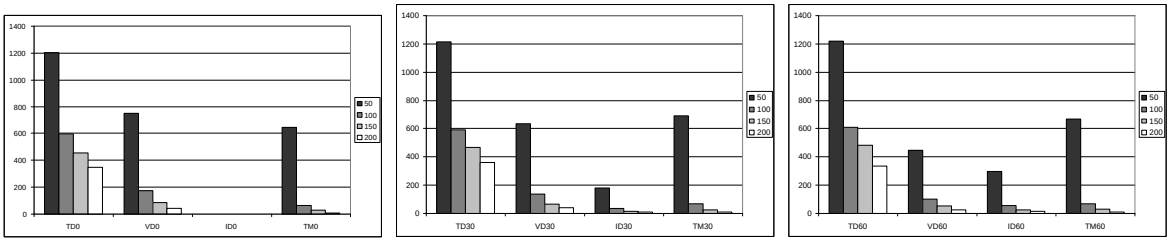

Fig. 5. 0%, 30%, and 60% malicious/faulty mobile devices, for 50, 100, 150, and 200 mobile devices

secure manner is one of the challenging problems in vehicular ad hoc networks.

We presented our proposed approach of a decentralized data management system that incorporates a variety of anchored and mobile devices and yet provides timely and accurate data to consumer devices. To ensure the reliability and trustworthiness of the aggregated data we presented a data intensive reputation management system suitable vehicular ad hoc networks. Our reputation systems aims to build up reputations where only a few prior trust relationships may possibly exist *a priori*.

Our preliminary simulation results show that even though a single device is restricted by its radio range and other interference, when foraging for information, increased scope, availability can be achieved by collaborative data exchanges with other devices. To ascertain the quality and reliability of the information thus aggregated, it is necessary to dynamically build and maintain reputations in order to make trusting decisions.

For future work we are investigating the use of adaptive and reciprocative cooperation levels in the collaborative data exchanges. Also, it is important to be able to determine the level of cooperation based on the current situation awareness of the device, e.g. in areas of high data availability the individual devices may scale back on the amount of information being exchanged. Another observation from our simulations was that as the number of devices in the neighborhood increase, the throughput decreases. Hence, we are also looking at combinations of adaptive transmission signal strengths and adaptive collaborative behavior to decrease contention and increase useful throughput.

REFERENCES

[1] The Haversine Formula. [Accessed on May 1, 2006. `http://en.wikipedia.org/wiki/Haversine_formula`].

[2] S. Buchegger and J.-Y. L. Boudec. A robust reputation system for p2p and mobile ad-hoc networks, 2004.

[3] L. Buttyan and J.-P. Hubaux. Enforcing service availability in mobile ad-hoc wans. In *MobiHoc '00: Proceedings of the 1st ACM international symposium on Mobile ad hoc networking & computing*, pages 87–96. IEEE Press, 2000.

[4] J. R. Douceur. The Sybil Attack. In *Proceedings of the First International Peer to Peer Workshop (IPTPS 2002)*, pages 251–260, March 2002.

[5] P. Golle, D. Greene, and J. Staddon. Detecting and correcting malicious data in vanets. In *VANET '04: Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks*, pages 29–37, New York, NY, USA, 2004. ACM Press.

[6] P. Gupta and P. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, 2000.

[7] R. Jurca and B. Faltings. An incentive compatible reputation mechanism. In *IEEE International Conference on E-commerce*, pages 285–292, 2003.

[8] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina. The eigentrust algorithm for reputation management in p2p networks. In *WWW '03: Proceedings of the 12th international conference on World Wide Web*, pages 640–651, New York, NY, USA, 2003. ACM Press.

[9] P. Michiardi and R. Molva. A game theoretical approach to evaluate cooperation enforcement mechanisms in mobile ad-hoc networks. 2003.

[10] G. Montenegro and C. Castelluccia. Statistically Unique and Cryptographically Verifiable(SUCV) identifiers and addresses. citeseer.ist.psu.edu/montenegro02statistically.html, 2002.

[11] J. Mundinger and J.-Y. Le Boudec. Analysis of a Robust Reputation System for Self-Organized Networks. *European Transactions on Communication*, 16(5):375–384, 2005.

[12] A. Patwardhan, F. Perich, A. Joshi, T. Finin, and Y. Yesha. Active Collaborations for Trustworthy Data Management in Ad Hoc Networks. In *Proceedings of the 2nd IEEE International Conference on Mobile Ad-Hoc and Sensor Systems*, Washington, DC, November 2005. IEEE.

[13] V. Srinivasan, P. Nuggehalli, C.-F. Chiasserini, and R. R. Rao. An analytical approach to the study of cooperation in wireless ad hoc networks. *IEEE Transactions on Wireless Communications*, 4(2):722–733, March 2005.

[14] X. Zeng, R. Bagrodia, and M. Gerla. GloMoSim: A Library for Parallel Simulation of Large-Scale Wireless Networks. In *Workshop on Parallel and Distributed Simulation*, 1998.

[15] S. Zhong, J. Chen, and Y. R. Yang. Sprite: A simple cheat-proof credit based system for mobile ad-hoc networks. In *INFOCOM 2003: Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies*, pages 1987–1997. IEEE, March 2003.