DisService: Network State Monitoring and Prediction for Opportunistic Information Dissemination in Tactical Networks

A. Mazzini, C. Stefanelli, M. Tortonesi
Department of Engineering
University of Ferrara
Ferrara, ITALY

G. Benincasa, N. Suri
Florida Institute for Human & Machine Cognition
Pensacola, FL, USA

Abstract—**Information dissemination is extremely difficult in tactical edge networks, which provide one of the most challenging environments for communication. The extremely dynamic nature of the tactical environment makes dissemination algorithms based on overlay networks or epidemic routing ineffective, and calls for ad hoc approaches that can dynamically reconfigure the dissemination process, adapting it to the current network conditions. This paper presents the algorithms for opportunistic information dissemination adopted in the DisService project. In particular, the paper investigates what kind of knowledge can be inferred about the interaction between nodes in a tactical network and how to use that knowledge to improve data dissemination algorithms. The algorithms are tested in a NS3 simulated environment and the results demonstrate the effectiveness of the DisService approach.**

Tactical edge networks; information dissemination, machine learning.

I. Introduction

Tactical networks, the basis for Network Centric Operations, provide one of the most challenging environments for communications, with mobile nodes connected via limited bandwidth and highly variable latency wireless ad-hoc links in hostile RF environments. The dynamic nature of military operations further complicates the scenario, leading to a frequently changing network topology and widely varying loads being placed on the network by users and applications.

Information dissemination, a critical function to enable essential tactical applications such as Blue Force Tracking, sensor data acquisition, target designation, and remote monitoring, is extremely difficult in this environment. On the one hand, the extreme dynamicity of network topology prevents the use of overlay networks to support dissemination, as they would be too expensive to setup and maintain in the tactical environment. On the other hand, epidemic dissemination algorithms, better suited for highly dynamic and mobile environments, have very high bandwidth and storage requirements, behave unpredictably, and do not always guarantee reliable delivery. As a result, developing efficient information dissemination algorithms that provide timely and reliable information delivery and that can cope with unreliable and bandwidth-constrained links is a very challenging task.

The peculiar nature of tactical edge networks calls for ad hoc information dissemination algorithms that can dynamically adapt to the current network conditions. For instance, the high mobility of some of the nodes that are typically found in tactical network suggests the opportunity to use those nodes as message ferries to improve the performance of the dissemination process. In addition, the significant topological differences across the network encourage to select the data caching and replication strategies better suited to local node density.

The aim of this paper is to study algorithms for opportunistic information dissemination in the context of the DisService project. DisService is an information dissemination middleware purposely designed for tactical network applications. More specifically, this paper focuses on the monitoring and prediction of contact windows between nodes, i.e., time intervals when they are within communication range. In addition, it details also how DisService can take advantage of this information to improve the dissemination algorithm performance.

We tested the DisService algorithms in the NS3 simulator. In the extensive simulations, we found that the adaptive information dissemination approach can lead to significant improvements in the timeliness and reliability of information delivery while reducing the overhead on the network.

II. Information Dissemination in Tactical Edge Networks

Reliable and timely information dissemination is essential in tactical edge networks. Perhaps the most fundamental need for tactical users is Blue Force Tracking (BFT), which provides presence and location of friendly forces, and applications that provide situational awareness information. BFT applications are indispensable to avoid friendly fire accidents. Timely exchange of tactical intelligence is critical to mission success. Also the retrieval of environmental data collected from sensors deployed on the ground or on unmanned moving robots is crucial to detect incoming threats.

Unfortunately, tactical edge networks are a highly dynamic environment that significantly complicates the realization of robust and efficient dissemination systems. In the tactical environment, it is common to find nodes that are disconnected or poorly connected to the rest of the network for most of the

978-1-4244-8180-4/10/$26.00 ©2010 IEEE

time. In fact, soldier platoons usually move across the battlefield and are connected to their Tactical Operation Center (TOC) through a low bandwidth tactical radio link. Soldiers may receive data from other sources, e.g., unattended ground sensor systems, only when they are in close proximity. Highly mobile Unmanned Air Vehicles (UAVs), and other airborne assets such as Joint Surveillance Target Attack Radar System (J-STARS) are also common in tactical networks. Besides their adoption for battlefield monitoring operations, UAVs are often used as information carriers to speed up the dissemination process between nodes on the battlefield, e.g., soldiers and ground sensors, and the TOC. Figure 1 summarizes the communication parties commonly found in tactical edge networks.

The extreme dynamicity of network topology prevents the use of overlay networks to realize dissemination in the tactical environment. In fact, the high mobility of nodes would make the process of setting up and maintaining overlays too expensive [1]. While better suited than overlay network approaches for highly dynamic and mobile environments, epidemic dissemination mechanisms have very high bandwidth and storage requirements and do not guarantee reliable delivery [2][3].

As a result, there is the need to develop ad hoc information dissemination algorithms that adapt to the network environment, leveraging potentially useful peculiar characteristics. For instance, there is the opportunity to take advantage of fast moving nodes to carry information to its final receivers, as well as to reduce the replication of data in portions of the network with a high node density. In order to save bandwidth and storage resources, less aggressive replication and caching strategies can also be adopted in areas where the links between nodes are relatively stable. In addition, there is often a direct relationship between physical proximity and the requirements for precision in the information exchange between nodes. An adaptive dissemination system can exploit this to reduce bandwidth utilization, especially for long communication path. Finally, not all consumers require reliable and/or sequenced delivery of information. The dissemination system should be flexible in supporting these properties when needed and saving resources when they are not necessary.

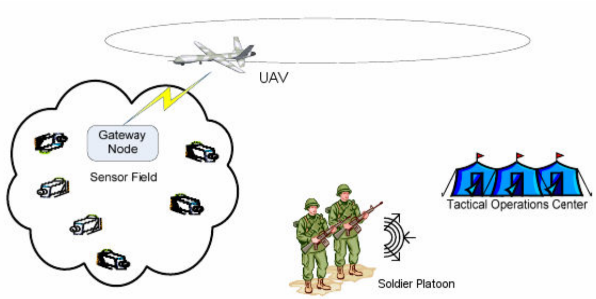

Figure 1. Typical tactical edge network scenario.

III. DisService

DisService is an information dissemination service for tactical edge networks that monitors and opportunistically manages communications, storage, and processing capacity in a distributed network to improve the performance of information dissemination [4]. DisService supports the storage and forwarding of data and caches data throughout the network, thereby enabling disruption tolerant dissemination and improving the availability of data.

DisService publishes information in the context of a group. Every node in the field can subscribe to many groups, and information sent in the context of a specific group may also be tagged to differentiate between multiple types of data. Each node in the network running DisService operates in a distributed, peer-to-peer manner while processing and communicating the published information and requested subscriptions from/to neighboring nodes.

DisService also supports multiple dissemination algorithms, including reliable flooding, epidemic protocols, and heuristic protocols, which can be selected based on the type of information being disseminated. DisService chooses the best dissemination algorithm according to the characteristics and preferences of the applications: the expected number of subscribers to a group, the nature of their subscription (reliable or not), and the relative priority. Reliable flooding is an extremely expensive algorithm, but is appropriate for high priority messages that must be delivered reliably. Epidemic protocols, instead, limit the number of data (re)transmissions in case of a large number of neighbor nodes, trading off reliability for a more efficient use of bandwidth and storage resources. Heuristic protocols further enhance epidemic approaches by exploiting domain-specific knowledge, e.g., knowledge about scheduling of new information generation that renders currently managed messages obsolete, to guide the probabilistic dissemination.

The multiple dissemination strategies supported by DisService allow addressing the significantly different requirements and priorities of applications, e.g., Blue Force Tracking, sensor data acquisition, tactical information exchange, and remote monitoring that must typically coexist in tactical networks and compete for scarce bandwidth resources.

While DisService provides many dissemination strategies and allows applications to choose the preferred one, currently it implements only a limited support for automated strategy selection according to the current network topology and state. There is a need to extend the DisService adaptation capabilities to consider environmental peculiarities and dynamicity as well as the application requirements.

IV. Opportunistic Information Dissemination in DisService

In the context of the DisService project, this paper focuses on the development of an adaptive system to select the most appropriate dissemination algorithms, among the ones supported by the middleware, according to the current network conditions. The algorithm selection requires explicit support from network state and topology monitoring mechanisms that must provide accurate and timely information to exploit for effective decision making in the context of dissemination strategy tuning. As a result, we have carefully designed the DisService middleware components that continuously monitor the network state using both passive and active measurements, and process metrics that could be useful for the tuning of information dissemination strategies.

DisService can adapt the dissemination strategies according to the current network conditions in order to achieve significant performance and reliability improvements. For instance, information that must be delivered to a small set of nodes that get periodically disconnected from the rest of the network might be best handled by dissemination strategies with aggressive caching and conservative forwarding. DisService integrates an adaptive system that chooses and tunes the dissemination strategies to adopt according to the current environmental conditions.

In this context, a particularly interesting metric to analyze is the *contact window* between different nodes. We define a contact window between nodes N1 and N2 as the tuple containing the start time and the duration of a time interval in which N1 and N2 are in communication range and can be considered neighbors. By monitoring and analyzing contact window information, DisService can discover cyclic mobility patterns and predict when a certain node might be in range.

DisService-enabled nodes implement the monitoring of contact windows by exchanging presence messages. Each node explores its neighborhood by broadcasting periodically one "Hello" message containing its unique node identifier. When a neighbor receives a "Hello" message from a node, it updates the corresponding status information of that node, creating a new contact window or updating the currently open one. If a node does not receive any message from a known neighbor for a specific amount of time, the neighbor is considered disconnected and the corresponding contact window is closed.

Each node stores a contact window history table, containing information about its current and previous contact windows to neighbor nodes. DisService then processes the collected information to build a model that enables predictions on future contact windows to nodes, especially the ones exhibiting cyclic mobility patterns. To this end, DisService can select the best forecasting algorithms among simple ones based on moving average models, such as Exponentially Weighted Moving Average (EWMA), as well as more complex algorithms that consider autoregression too, such as Autoregressive Integrated Moving Average (ARIMA), and also algorithms based on multivariate time series forecasting.

We have considered the EWMA algorithm as it is often used in distributed systems (for example, to calculate RTT in TCP) because of its simplicity. The performances of the EWMA heavily depend on the smoothing parameter, that has to be configured according to the dynamicity of the environment under control. We have also implemented ARIMA, a statistical technique for understanding and predicting future values of time series. ARIMA models are heavily used in econometrics for their capability to predict non-stationary processes with emergent trends, e.g., constant, linear, or quadratic. Finally, since the contact window is a two-dimensional metric, we have considered a multivariate time series forecast approach based on autoregressive moving average models.

We have subsequently developed heuristics that enable DisService to exploit the forecasts and put in place the best information dissemination strategy. More specifically, we have implemented an opportunistic epidemic dissemination strategy that leverages the contact window prediction provided by the algorithms discussed previously. When using this opportunistic strategy, nodes that have a new message for subscription S check whether any neighbor node belongs to the corresponding group. If there is no node with subscription S within range but a contact prediction to one such node is available, the originating node stores the message in its internal cache and delays the transmission until the other node comes into range. To improve the reliability, the originating node can also decide to transmit some messages anyways. If no prediction is available, the originating node forwards the data to all of its neighbors, assuming that one of them might be closer to a node with subscription S.

We designed the opportunistic epidemic dissemination strategy to favor the use of highly mobile nodes, such as UAVs, as intermediate nodes for information propagation, given their increased reach and visibility. The strategy is also effective in cases when a node acts as a relay to a neighbor that has no other communication link (which implies that the first node is the only way for the second node to receive any information).

V. Experimental Results

We have created an accurate testbed in a simulated environment in Network Simulator 3 (NS3) [5] in order to test the algorithms implemented in DisService in a reproducible and controllable environment. We used version 3.7.1 of the NS3 simulator for all the experiments presented in this paper.

We decided to simulate a typical battlefield scenario, where two platoons of soldiers move in groups across the battlefield, where ground sensors collect data, and a UAV speeds up the data dissemination process by carrying data between the sensor field and the TOC. The NS3 testing environment was set up with 80 nodes in total: 40 nodes – divided in 2 groups of 20 nodes each – representing sensors that collect environmental information, one node representing the TOC, one node representing an UAV which "harvests" information from sensors deployed on the field and delivers it to soldiers and to the TOC, and 38 nodes representing two patrols of soldiers – the first one composed of 18 nodes and the other of 20 nodes. All the nodes are connected through a wireless 802.11 connection at 6Mbps with non QoS-enabled MAC layer; the TOC is also connected to each of the soldier patrols through a point-to-point tactical radio link at 1.5Mbps.

In every experiment the TOC is placed in the same fixed position, the sensors are positioned randomly in two 300x300-meter areas on the opposite sides of the battlefield, the UAV moves according to a Waypoint Mobility Model, with three waypoints set respectively near the TOC and in the middle of each ground sensor field. The soldier patrols move with a Reference Point Group Mobility model, with soldiers closely following their patrol leader across waypoints chosen randomly on the battlefield.

By analyzing the data collected from extensive NS3 simulations, we observed that the number of neighbors changes significantly as a function of time and network location. Figure 2 demonstrates this by showing how the mean number of neighbors that each node has throughout the simulation. The results are grouped by node type. As expected, the number of neighboring nodes for sensors and soldiers varies significantly as they come within communication range. The variation is even more appreciable in the case of the UAV. These observations suggests considering local topology characteristics, e.g., node density, to modulate the aggressiveness of data caching and replication policies in adaptive dissemination strategies in order to save storage and bandwidth resources.

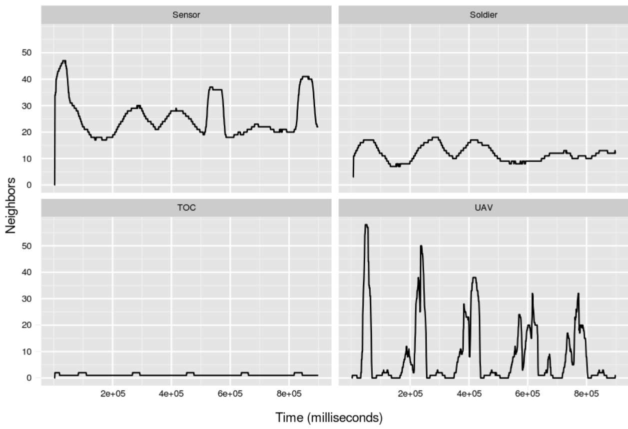

Figure 2. Mean number of neighbor nodes. Results grouped by node type.

We then studied contact window data, and observed that contacts between nodes exhibit an interesting behavior. In fact Figure 3 shows the density estimation of contact window duration for each type of neighbor node, in order to verify which kind of node exhibits predictable contact window duration patterns.

Notice that the contact window distribution is multimodal. These results suggest the opportunity to predict the duration of contact windows by leveraging previously sampled values. In some cases, such as contact windows to the TOC, conditional prediction seems particularly interesting: when a certain threshold is passed, there is a high probability that the contact will last significantly longer. On the contrary the contacts to the UAV are mostly bounded by a 5-second threshold.

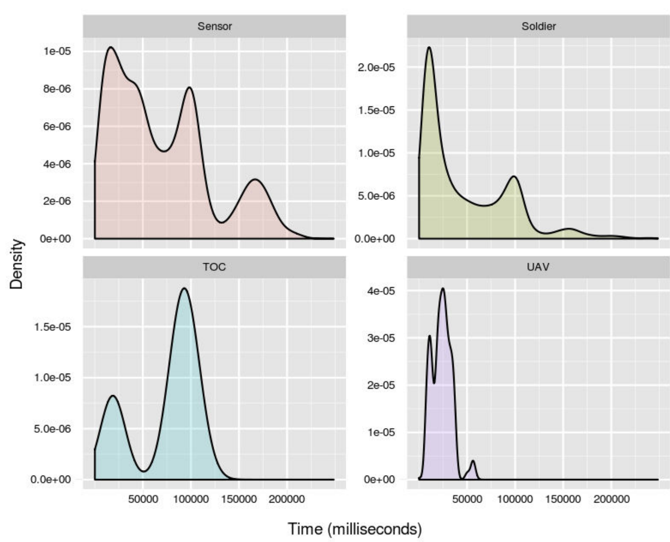

Figure 3. Distribution of contact window duration with neigbor nodes, grouped by neighbor node type.

In addition, Figure 4 shows the density estimation of time to next contact, grouped by neighbor node type. The figure demonstrates that time to next contact is a predictable metric. For instance, times to the next contact with the UAV fall mostly in a short range around the 20 seconds value and times to the next contact with the TOC are mostly grouped around the two modes, respectively at 30 and 220 seconds.

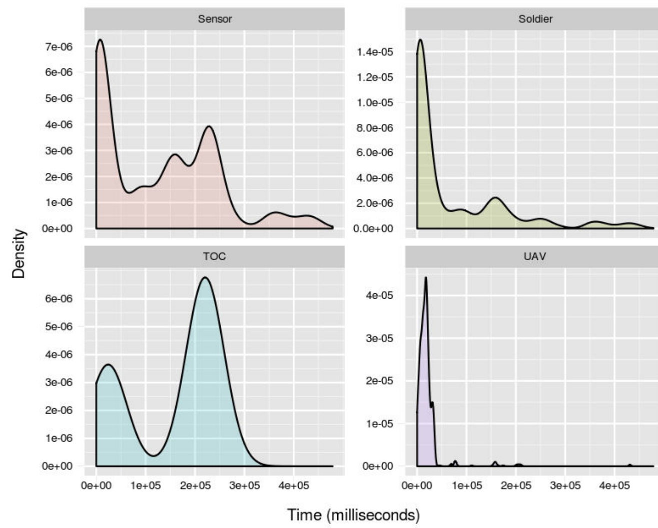

Figure 4. Distribution of time to next contact, grouped by neighbor node type.

These observations raise the opportunity to study whether we are able to predict the next contact window information of a specific node with the algorithms described in the previous section. In addition, we are interested in finding which algorithm provides the most accurate results, and can effectively provide reliable input for dissemination strategy selection and tuning decisions.

First we evaluated one-step ahead forecasts provided by the Exponentially Weighted Moving Average (EWMA), using the Zoo package [6] in the R Statistical Framework [7]. Figure 5 shows the results we have achieved. An intermediate value like 0.5 for the smoothing parameter seems to offer a good tradeoff for both the fast-varying time to next contact and the low-varying next contact duration variables.

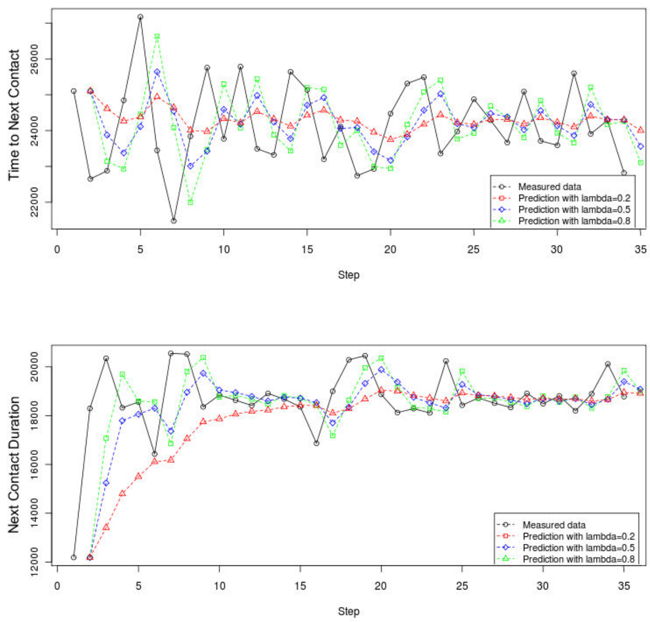

Figure 5. One-step ahead forecast of time to next contact and next contact duration (in milliseconds) using the EWMA algorithm, with several values for the smoothing parameter.

We have then evaluated the performance of the ARIMA forecasting algorithm, by leveraging on the standard library functions of the R Statistical Framework [7]. The results are shown in Figure 6.

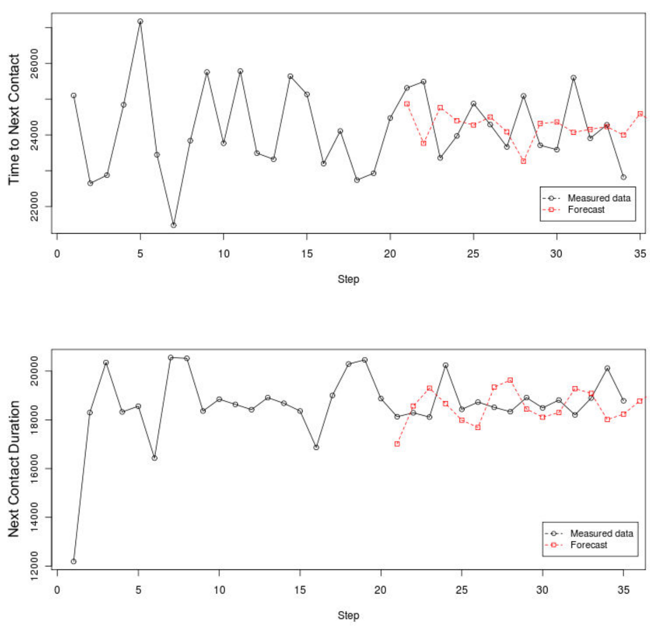

Figure 6. Forecast of time to next contact and next contact duration (in milliseconds) using the ARIMA algorithm.

Finally, we have evaluated the performance of the multivariate time series forecast algorithm, using the Dynamic Systems Estimation package [8] in R. Figure 7 shows the results of the one-step ahead prediction.

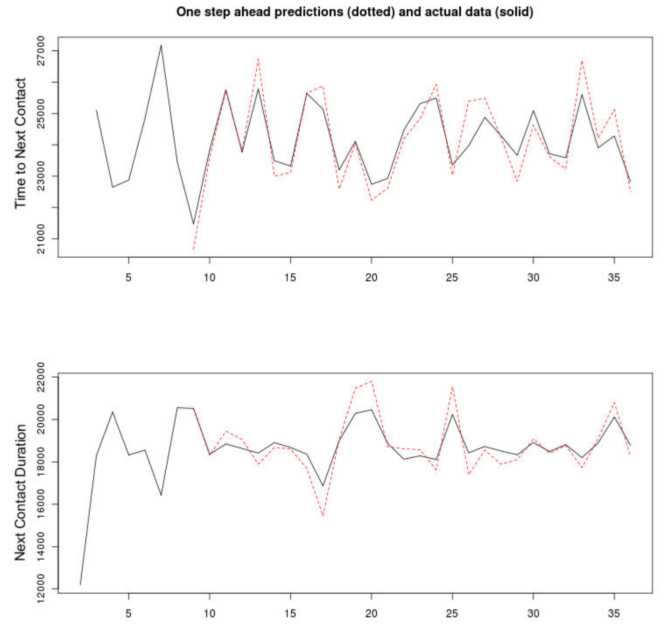

Figure 7. One-step ahead forecast of time to next contact and next contact duration (in milliseconds) using the multivariate time series algorithm.

Table I presents the Standard Error of the Estimate (SEE) for the forecast algorithms discussed above. The multivariate time series forecast algorithm proved to be significantly more accurate than the others, although it is more complex and requires several samples for bootstrapping. As a result, it represents a viable solution for nodes with high computational capabilities. The ARIMA algorithm produces reasonably accurate results, but is not suited for application in the tactical environment as it needs a large number of samples for bootstrapping and has a significant computational complexity. Finally, while the accuracy of EWMA-based forecast is limited, we believe that the lightweight computational requirements of the EWMA algorithm make it well suited for nodes with low computational capabilities. In addition, the simplicity of EWMA enables implementing several estimators with different smoothing parameters at the same time, and to dynamically switch to the most accurate one.

TABLE I. STANDARD ERROR OF THE ESTIMATE (SEE) OF THE PROPOSED FORECAST ALGORITHMS

	Time to Next Contact	Next Contact Duration
EWMA (λ=0.2)	1328.2	2160.0
EWMA (λ =0.5)	1463.1	1739.3
EWMA (λ =0.8)	1583.2	1678.5
ARIMA	1014.4	866.0
Multivariate forecast	565.0	638.8

Finally, we studied how to leverage the knowledge acquired about the prediction algorithms to improve DisService performance and reliability, through the adoption of adaptive dissemination strategies.

More specifically, we considered a typical sensor data acquisition application, in which each sensor pushes data periodically to its neighbors, channeling each message through the "data" subscription identifier. One of the ground sensor nodes acts as a data collector, applying simple data fusion algorithms to the content of the messages received by its neighbors, in order to limit the bandwidth required for the transmission of data to the final consumers while providing additional contextual information. The data collector node transmits the processed data on the "data fusion" subscription, effectively handing it over to the UAV that, in turn, delivers it to the TOC.

We then implemented this application in the scenario described above and evaluated the performance and reliability improvements that could be achieved by exploiting in this context the opportunistic epidemic dissemination strategy described in the previous section. First, we configured the testbed to adopt a probabilistic epidemic dissemination strategy, where every node, upon receiving a message, forwards the incoming packet to its neighbors with a 40% probability. We then performed another simulation run using the opportunistic epidemic dissemination strategy, with the same retransmission probability parameter. With regards to the forecast algorithm for contact window predictions, we decided to adopt the EWMA (with λ =0.5) algorithm, because of its simplicity.

Analyzing the experimental results, we observed that the opportunistic epidemic dissemination strategy significantly improves the reliability of the system while limiting the retransmission of packets, since the messages are forwarded only when the destination can be reached. More specifically we have observed an average reduction of about 3% of the network overhead, measured as the number of duplicate messages received by each node. The reliability of the system, instead, measured as the number of messages received from the UAV, significantly raised from an average of 75%-80% to almost 100%.

VI. CONCLUSIONS

The results presented in this paper demonstrate that adaptive dissemination strategies that react to network state and topology changes can effectively reduce bandwidth consumption while supporting faster and more reliable delivery of data.

The promising results we have achieved so far call for further work in order to develop even more efficient heuristics to better adapt the information dissemination strategies to the current application requirements and network environment state.

In the future, we are planning to analyze metrics related to local network topology characteristics, e.g., number of neighbor nodes, as an input for adaptive dissemination strategies. In addition, we will also consider the optimization opportunities brought by the integration of data fusion with information dissemination. Finally, we would like to evaluate cross-layer solutions that could leverage feedback from the MAC layer for more lightweight contact window monitoring.

ACKNOWLEDGMENT

This work is supported in part by the Office of Naval Research under grant N00014-09-1-012 and by the U.S. Army Research Laboratory under Cooperative Agreement W911NF-04-2-0013.

REFERENCES

[1] C. Gui, P. Mohapatra, "Overlay Multicast for MANETs using dynamic virtual mesh", Wireless Networks, Vol. 13, No. 1, pp. 77-91, January 2007.

[2] A.Vahdat, D. Becker, "Epidemic Routing for Partially-Connected Ad Hoc Networks", Duke University Technical Report CS-200006, April 2000.

[3] A. Lindgren, A. Doria, O. Schelen, "Probabilistic Routing in Intermittently Connected Networks", in Proceedings of the First International Workshop on Assurance with Partial and Intermittent Resources (SAPIR), Fortaleza, Brazil, August 2004.

[4] N. Suri, G. Benincasa, S. Choy, S. Formaggi, M. Gilioli, M. Interlandi, J. Kovach, S. Rota, R. Winkler, "DisService: A Peer-to-peer Disruption Tolerant Dissemination Service", in Proceedings of IEEE MILCOM 2009 Military Communications Conference, ISBN 978-1-4244-5239-2, 18-21 October 2009, Boston, MA, USA.

[5] NS3 Network Simulator Homepage: http://www.nsnam.org/ (retrieved on April 23rd, 2010).

[6] A. Zeiles, G. Grothendieck, "zoo: S3 Infrastructure for Regular and Irregular Time Series", Journal of Statistical Software, Vol. 14, No. 6, May 2005.

[7] R Development Core Team, "R: A language and Environment for Statistical Computing", ISBN 3-900051-07-0, Vienna, Austria, 2009, URL: http://www.R-project.org (retrieved on April 23rd, 2010).

[8] P. Gilbert, "A Note on the Computation of Time Series Model Roots" Applied Economics Letters, Vol. 7, No. 7, pp. 423-424, July 2000.