Improving Probabilistic Flooding Using Topological Indexes

Dawit Kifle
IT Doctoral Program
Addis Ababa University
Addis Ababa, Ethiopia
Email: dawit.kifle@aau.edu.et

Gabriele Gianini
EBTIC/Khalifa University of Science and Technology
Abu Dhabi, UAE
Università degli Studi di Milano
Milano, Italy
Email: gabriele.gianini@unimi.it

Mulugeta Libsie
Department of Computer Science
Addis Ababa University
Addis Ababa, Ethiopia
Email: mulugeta.libsie@aau.edu.et

Abstract

Unstructured networks are characterized by constrained bandwidth and battery power and require protocols that utilize them efficiently. Probabilistic flooding, allows nodes to rebroadcast RREQ packets with some probability p, thus reducing the overhead. The key issue in of this algorithm consists of determining p. The techniques proposed so far either use a fixed p determined by a priori considerations, or a p variable from one node to the other – set, for instance based on node degree or distance between source and destination – or even a dynamic p based on the number of redundant messages received by the nodes. In order to make the computation of forwarding probability p works optimally regardless of changing of topology, we propose to set p based on the node role within the message dissemination process. Specifically, we propose to identify such role based on the nodes' clustering coefficients (the lower the coefficient, the higher the forwarding probability). The performance of the algorithm is evaluated in terms of routing overhead, packet delivery ratio, and end-to-end delay. The algorithm pays a price in terms of computation time for discovering the clustering coefficient, however reduces unnecessary and redundant control messages and achieves a significant improvements in both dense and sparse networks in terms of packet delivery ratio. We compare by simulation the performance of this algorithm with the one of the most representative competing algorithms.

I. INTRODUCTION

The characteristics of unstructured networks (absence of fixed infrastructure, resource constraint, and, where it applies, mobility) pose severe challenges to the corresponding routing protocols [1], [2]. The simplest technique to disseminate messages in such networks is an epidemic protocol called flooding, where nodes simply rebroadcast any received message. With flooding messages reach any node in a number of steps of the order of the network diameter. Thus the algorithm has a high reachability and coverage and is characterized by extreme simplicity: it does not require complex computations, which also results in low latency. However its bandwidth consumption is highly dependent on the number of nodes in a network. As the latter increases the technique becomes expensive: flooding storms - in which redundant messages cause collision and contention - are the main issue of flooding especially in dense networks [3], [4], [5]. To address this problem several solutions have been proposed: probabilistic flooding (gossip-based routing) which is the focus of the present work represents, in all its variants, the most adopted solution.

The key issue in of this algorithm consists of determining p. The works in [6], [7] select a fixed value of p based on the best results obtained in the simulation of the network. Such algorithm performs better than flooding in terms of traffic overhead: in [7] it is reported that the gossiping protocol uses up to 35% fewer messages than flooding. Unfortunately gossip-based protocols – specifically those with fixed p – can be affected by the problem of *early gossip termination*: this originates when the value of a forwarding probability is insufficient to disseminate the message across the whole network. The probability of occurrence of such problem depends on the network connectivity and, for a given value of probability, is more likely of sparse networks.

Several strategies have been adopted to counter this problem in epidemic protocols. They will be reviewed in the Related Work Section. The techniques proposed so far either use a fixed p determined by a priori considerations, or a p variable from one node to the other – set, for instance based on node degree or distance between source and destination – or even a dynamic p based on the number of redundant messages received by the nodes.

In this work, in order to make the computation of forwarding probability p work optimally regardless of changing of topology, we propose to set p based on the node role within the message dissemination process. We propose to identify such role based on the nodes' clustering coefficients (the lower the coefficient, the higher the forwarding probability). In previous works [8], [9], we had considered the tuning of the forwarding probability based on local topology, by investigating the use of a purposely devised effective node degree; in the present work we use a general purpose metrics, the local clustering

coefficient: this metrics quantifies the degree by which how the neighborhood of a node is close to being a clique (i.e. a complete graph).

We denote our algorithm by VTV(cc) (Variable probability Threshold Value based on Clustering Coefficient) and evaluate its performance in terms of routing overhead, packet delivery ratio and end-to-end delay. By simulation, ee compare its performance to the ones of reference algorithms. We show that, although the algorithm pays a price in terms of computation time for discovering the clustering coefficient, it reduces unnecessary and redundant control messages and achieves a significant improvements in both dense and sparse networks, especially in terms of packet delivery ratio.

The reminder of the paper is organized as follows. Section II reviews related work. Section III presents the proposed solution. The simulation setup and the results are discussed in Section IV. Finally, conclusions are drawn and future works are indicated in Section V.

II. Related Work

A forwarding probability – in which messages are forwarded with some probability value – is the focal point for all gossip-based routing protocols. Forwarding messages with some probability greatly improve the broadcast storm of flooding techniques. On the other hand, latency and early gossip termination are the main problems associated with gossip-based protocols [10], [11].

To improve the performance of gossip-based protocols in terms of latency and to reduce early gossip termination, several researchers propose different mechanisms to compute a forwarding probability. There are at least four major techniques used in practice: (1) Fixed forwarding probability, (2) probability set based on node degree, (3) probability set based on received number of duplicate messages, and (4) probability based on summaries of the local topology.

A. Fixed forwarding probability

The simplest gossiping protocol is a protocol with a fixed probability p. It has been studied for different values of p, but normally it is not considered in practice, since its problems (early gossip termination for small p and flooding storms for high p) are well known. To improve over that protocol, Haas et al. [6], [7], [12], proposed four gossip based algorithms, named Gossip1-Gossip4.

1) Gossip1 (p, k): In Gossip1, nodes forward the received packets for the first k (1 and 4 used as k's value in their experiment) hops with probability of $p = 1$. After k hops nodes that receive a RREQ forward with probability of p (whose value is set by the authors within the range of $[0.65, 0.86]$. In terms of reachability Gossip1 performs better than simple gossip: it reduces early gossip termination, especially in sparse networks. For Gossip1 to perform better the value of k should be chosen carefully. The authors set k to 1 for dense networks and to 4 for sparse networks.

2) Gossip2 (p_1, p_2, k, n): In Gossip2, two different forwarding probabilities p_1 and p_2 (where $p_2 > p_1$ always) are proposed respectively for dense and sparse regions of the network. Nodes with less than n neighbors forward a received RREQ with high probability p_2; the reminder, with low probability $p1$. The value of n is not stated in the paper. The simulation results show that Gossip2 achieves slight improvements over Gossip1 in terms of reachability, due to high forwarding probability for sparse areas of the network.

3) Gossip3 (p, m): In Gossip1 and Gossip2, a node whose p is greater than the random number generated for a RREQ packet drops the packet. On the contrary, Gossip3 lets a node store such a packet. If a node does not receive a duplicate of packets from at least m number of neighbors within a specified timeout t, the node forwards packets.

4) AODV+G: AODV+G replaces a flooding-based route discovery approach of AODV with Gossip3. The aim of the authors is to minimize the routing traffic due to flooding of RREQ during the route discovery phase of AODV, which normally would use flooding. Due to an improvement in routing load, AODV+G delivers better network performance in terms of end-to-end delay and packet delivery ratio and routing load.

B. Probability based on node degree and distance:

Cartigny and Simplot [13], [14], [15], proposed four works, that differentiate from Haas et al., in that p is determined by local information at a node.

1) Density Aware Probabilistic Flooding (DAPF): This allows each node to compute its forwarding probability - as shown in Equation 1 - based on node degree.

$$p = \frac{k}{n} \tag{1}$$

where $k \geq 1$, is an efficiency parameter and $n > 0$ is the node degree. A node with less neighbors forwards with high probability and vice-versa.

2) Border Retransmission Based on Probabilistic Flooding (BRBPF): BRBPF lets each node compute p based on the distance μ between sender and receiver nodes.

$$\mu = \frac{N_b}{N_a + N_c} \qquad (2)$$

where $N_b \geq 1$ is the number of neighbors that are unique for the broadcast message receiver, $N_a \geq 1$ is the number of neighbors unique for the sender of broadcast message, $N_c \geq 0$ is the common neighbors of the receiver and sender. a receiver estimates the distance of a node that forwards RREQ packet by comparing its neighbors to the neighbors of the sender. If both of them have common neighbors, the RREQ is forwarded with low probability. Since the node degree is not considered in BRBPF the algorithm is not very performing, especially in sparse networks.

3) Density Aware and Border Node Retransmission Based on Probabilistic Flooding (DABNRBPF): The forwarding probability computation which is proposed in DABNRBPF is

$$p = \frac{A - \alpha}{M^\sigma} \mu^\sigma + \alpha \qquad (3)$$

where A and α are used to bound the probability, σ is the convexity coefficient of the graph, and M is a constant representing the largest possible value for μ.

In the experiment M is set to 0.601 and A and α are set to 1.0 and 0.0, respectively. Thus $p = (\mu/M)^\sigma$. The simulation results show that if the parameters are selected carefully, DABNRBPF performs better than DAPF and BRBPF in terms of reachability.

4) Density Aware and Border Node Retransmission Based Probabilistic Flooding with Neighbor Elimination (DABN-RBPFNE): In DABNRBPFNE, a forwarding probability is computed by setting $A = k/n$, i.e.

$$p = \frac{k/n - \alpha}{M^\sigma} \mu^\sigma + \alpha \qquad (4)$$

Equation 4 combines Equation 1 and 3.

Simulation results show that DABNRBPFNE performs best in terms of reachability and reduction of unnecessary broadcasts compared to the other three works.

C. Probability based on duplicate message count

Mohammed et al., [16], [17], proposed two gossiping routing variants, namely, the Efficient Counter Scheme and Adjusted Counter-Based Broadcast Scheme (ACBS).

1) Efficient Counter Scheme (ECS): ECS computes p based on the number of received duplicates: nodes that receive a RREQ, store it for a period of Random Assessment Delay (RAD). During RAD, nodes count the number of duplicates d that they receiveand if d is less than a threshold m, the packet is rebroadcasted. The authors replaced the route discovery of AODV – flooding approach – by ECS (AODV-ECS). As the simulation results show, AODV-ECS performs better than AODV in terms of reachability, number of transmitted packets and latency. Since nodes can forward packets only after the RAD timer has expired, its performance in terms of end-to-end delay depends on the choice of RAD timer.

2) Adjusted Counter-Based Broadcast Scheme (ACBS): To improve over ESC in terms of reachability, the authors of [17] proposed ACBS: here nodes are allowed to broadcast the stored packet when RAD timer expires even if the number of received duplicates d is greater than the threshold m, but using a low broadcast probability p_1; if the count d is lower than the threshold the packet is forwarded with high probability p_2. The simulation results show that for small value of p_1 ACBS performs slightly better than ECS.

D. Local Topology Aware Tuning of Probabilistic Routing

In [9] the forwarding probability is tuned based on nodes' role in their local topology in terms of message dissemination: the salient features of the topology are captured by some synthesis metrics. Nodes important in terms of message dissemination are identified as effective nodes, nodes with little role in dissemination are identified as non effective. A given node v identifies its neighbors nodes as effective or non-effective based on its effective node degree $effd(v)$, defined as the number of neighbors of v with more than one neighbor.

The forwarding probability p of node v is computed as

$$p = \frac{N_v - effd(v)}{N_v} \qquad (5)$$

The forwarding probability p of node v is high when most of its neighbor nodes are non-effective neighbor nodes. The simulation [9] confirms the intuition that this definition works better in loosely connected networks. However, this definition
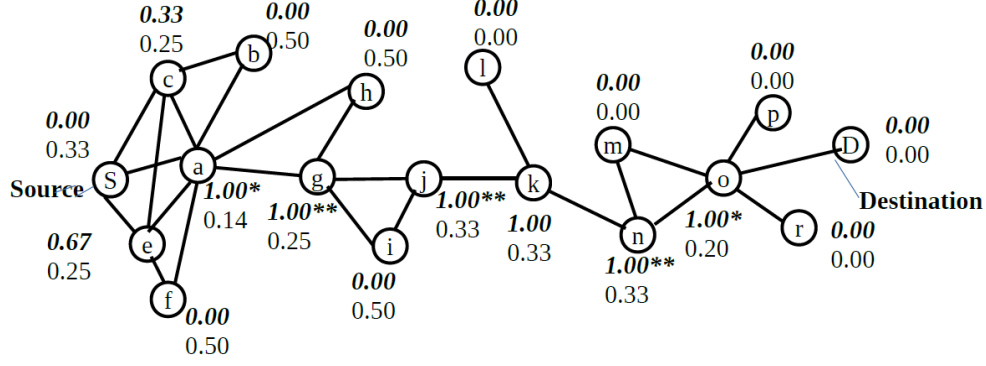
Figure 1. *A stylized example of unstructured network annotated with: the forwarding probabilities computed by VTV(cc) (in boldface) and those computed by VTV(non-effd). S indicates a source node, D a destination node for an hypothetical path. The number of asterisks near the probability value denotes the level of criticality of the node. The criticality here indicates how much the removal of the a can get the network close to partitioning in a part including S and a part including D (equivalently, how the failure of the node to broadcast a message can impair the path discovery process).*

has almost no impact in well connected topologies and does not identify critical nodes (whose removal can partition the network into subnetworks). The weakness of this algorithm, called VTV (non-effd), is addressed by the algorithm VTV (cc), presented in the next section, by means of the use of topological indexes.

III. VARIABLE THRESHOLD VALUE AS A FUNCTION OF CLUSTERING COEFFICIENT VTV (CC)

In well connected topologies, the role or importance of nodes can be rated by their ability to diffuse the received RREQ packets out of their local topologies and to avoid unnecessary redundant RREQ packets circulation in their neighborhood.

A. Tuning p based on the Clustering Coefficient

To complement VTV(non-effd), we propose to compute a forwarding probability as a function clustering coefficient cc_v of the node v, defined as

$$cc_v = \frac{NE_v}{maxNE_v} = \frac{2 * NE_v}{N_v(N_v - 1)} \tag{6}$$

where NE_v is the number of edges between pairs of the N_v neighbors of node v and $maxNE_v$ is the maximum number of pairs possible among the neighbors. The clustering coefficient is endowed with good capabilities for describing the degree of connectivity of a graph. We found it a useful topological index able to quantify the importance of a node in terms of message dissemination.

In the proposed algorithm, which we call VTV(cc), first the nodes compute their clustering coefficient, based on information received from the neighbors, using Equation 6, then they set their forwarding probability as:

$$p = 1 - cc_v \tag{7}$$

except for the case in which $0 \neq cc_v \neq 1$ and v has at least a neighbor with clustering coefficient 1: in that case $p = 1$. In mobility scenarios we assume for the sake of simplicity that the clustering coefficient of nodes is recomputed every time a node receives a RREQ packets and has to make a forwarding decision. This choice has a computational cost and a communication cost (which is $O(N^2)$ where N is the number of neighboring).

B. An example comparing VTV(non-effd) to VTV(cc)

Here we clarify the rationale behind VTV(cc) by comparing it to VTV(non-effd) on a stylized example network, represented in Figure 1. Assume that the source node S has launched a route discovery to establish a path to a destination node D. The nodes shown have different impact on message dissemination. For instance, nodes a, g, j, k, n, and o (shown in Figure 1) significantly disrupt the dissemination of RREQ packets if they decide not to rebroadcast. On the other hand, some other nodes have little impact. For instance, if nodes like $b, c, e, f, h, i, l, m, p, r$ rebroadcast the received RREQ packets, they disseminate only unnecessary redundant messages (i.e., messages that are already received by nodes).

The nodes in Figure 1 are annotated by their forwarding probability computed as a function of cc within VTV(cc) (boldface) and within VTV(non-effd) according to equation (5). Focussing on the high-impact nodes (denoted by asterisks) one can see that all of them forward the received message with probability $p = 1$ in the case of VTV (cc); in the case of VTV(non-effd) the same nodes are allowed to forward messages with a forwarding probability $p < 1$. Focussing on the

Table I
SIMULATOR AND THE PARAMETERS' VALUES USED IN THE EXPERIMENT

Simulator	NS2
Mobility Model	Random Way-Point
Speed	15 m/s
Transmission Range	250m
Packet Size	512 bytes
Area	1500m by 1500m
Source-Dest. Connectivities	5, 10, 15, 20, 25, 30
Number of Nodes	900
Simulation Time	100 sec.

low-impact nodes one can see that VTV(non-effd) allows them to rebroadcast messages with some probability $p >$ whereas VTV(cc) bars them from forwarding the message.

C. Ad-hoc On-Demand Distance Vector (AODV)

AODV it a well documented and well tested MANETs routing protocol. As a result, we found it an ideal protocol to be used as a benchmark.

Within the protocol AODV, the first time a node u requests a route to node v, it uses an expanding-ring search to find the route. That is, it first tries to find the route in a zone of small radius, by flooding. Then tries to find the route in zones of larger and larger radius. If all these attempts fail, it resorts to flooding the message through the whole network. The exact choice of zone radii to try is a parameter of AODV. Typically, not too many radii are considered before resorting to flooding throughout the network [18].

By means of simulation experiments, we are going to compare VTV(cc) and VTV(non-effd) to AODV in terms of some standard performance metrics.

IV. SIMULATION RESULTS AND ANALYSIS

TWe tested the performance of VTV (cc) in sparse networks (SN) and dense networks (DN), in a mobility scenario set according to the random way-point model. The parameters used in the simulation are shown in Table I. We compared to VTV (non-effd) and AODV, using the metrics packet delivery ratio, routing overhead and end-to- end delay.

- The Average End-to-End Delay refers to the time taken for a packet to be transmitted across a network from source to destination, averaged over all the possible source-destination node pairs.
- The Packet Delivery Ratio refers to the number of Packets successfully received by all receivers.
- The Routing Overhead is the the average number of routing packets required to deliver a single data packet.

A. Case 1: Dense Networks

a) Average End-to-End Delay: Computing the clustering coefficient value on fly helps nodes to make a forwarding decision based on the updated local topology information. However, it entails a high overall computational effort in dense networks which increases the latency.

This situation makes VTV (cc) different from VTV (non-effd) and AODV. The main reason for end-to-end delay of VTV (non-effd) in dense networks is the number of collisions of the redundant traffic. On the contrarythe , the latency of VTV (cc) is due to the clustering coefficient computation time. This is apparent from Figure 2, where the average end-to-end delay is plotted against the number of nodes in the network.

b) Packet delivery Ratio: The packet delivery ratio of VTV (non-effd) and AODV starts to decline in the presence of high congestion - which is caused by high redundant control messages during route discovery. Due to the limitation of VTV (non-effd) in dense networks, some non-effective nodes are undetected and participate in forwarding RREQ packets. As a consequence, VTV (non-effd) in dense networks is affected by broadcast storms. For instance, on average 34.44% more redundant RREQ packet are generated by VTV (non-effd) compared to VTV (cc). This situation isapparent in the plots of Figure 3. In dense networks, the minimum packet delivery ratio achieved by VTV (non-effd) and VTV (cc), respectively, is 55% and more than 80%, .
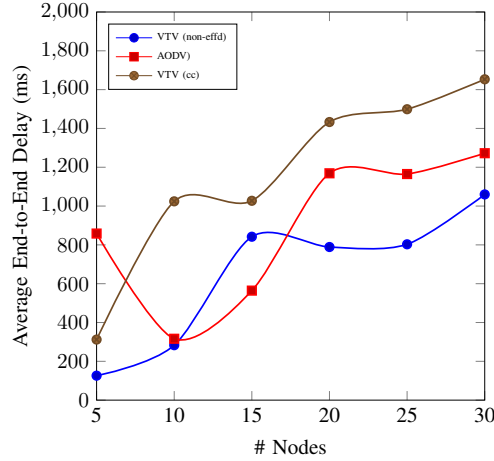
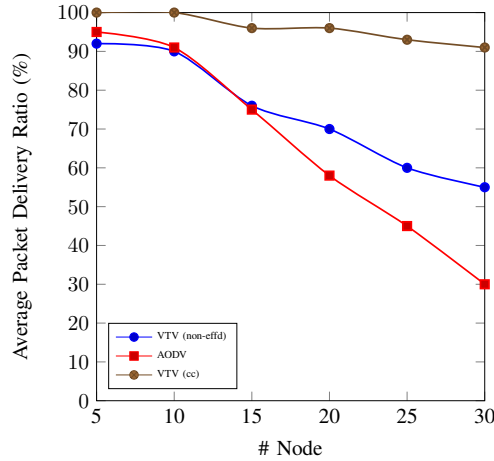Figure 2. Dense networks: End-to-End Delay.

Figure 3. Dense networks: Packet Delivery Ratio.

c) Routing Overhead: Shown in Figure 3 is the Routing overhead of the three considered protocols. The generation of control messages, by VTV (cc), based on the clustering coefficient, yields an acceptable trade-off in packet delivery ratio. VTV (cc) is able to reduces routing overhead significantly compared routing VTV (non-effd) and AODV. Besides, the performance of VTV (cc) less sensitive to the reliability of source-destination connections, as also apparent in the example shown in Figure 3.

B. Case 2: In Sparse Networks

a) Average End-to-End Delay: As discussed earlier, the time complexity of VTV (cc) (which is $O(n^2)$) makes the end-to-end delay of VTV (cc) high. However the delay due to computational time of clustering coefficient is lower in sparse networks than in dense networks. As shown in Figure 5, the difference in end-to-end delay between VTV (non-effd) and VTV (cc) is marginal.

b) Packet Delivery Ratio: Early gossip termination especially in sparse networks is a common event in gossip-based routing protocols and it represents the main reason for their possible poor performance in terms of packet delivery ratio. For example, the average packet delivery ratio achieved by AODV and VTV (non-effd) are 77.67% and 80.67%, respectively. However, VTV (cc) is able to achieve 88% of packet delivery ratio in the same network scenario. One can also observe in Figure 6, that the behaviour of VTV(cc) it is stable despite with respect to the number of nodes.

c) Average Routing Overhead: Routing overhead is not a serious issue in sparse networks. But this doesn't mean that there is no redundant RREQ packets. In our experiments, VTV (cc) was able to reduce the routing overhead by 6.6% compared to VTV (non-effd). The results are shon in Figure 7.
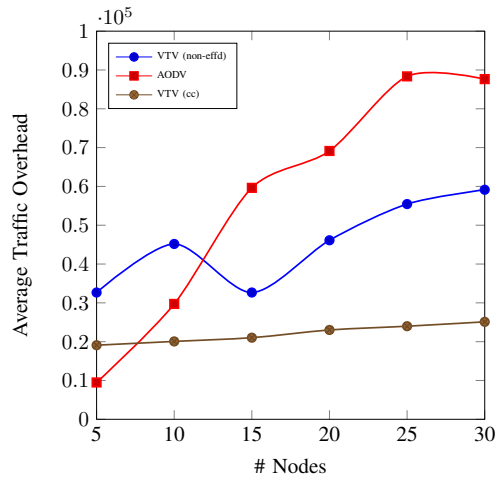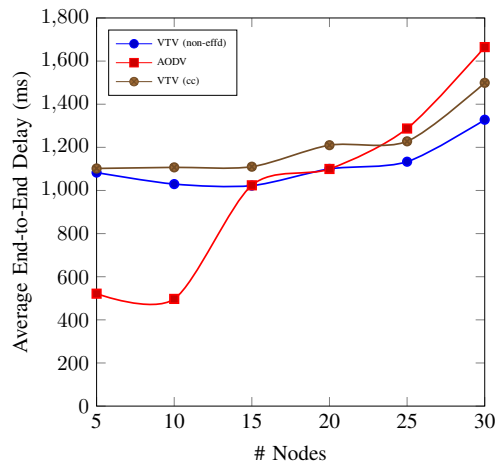
Figure 4. Dense networks: Traffic Overhead.
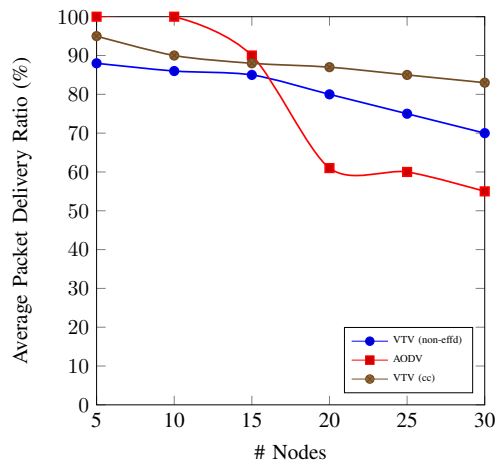

Figure 5. Sparse networks: End-to-End Delay.

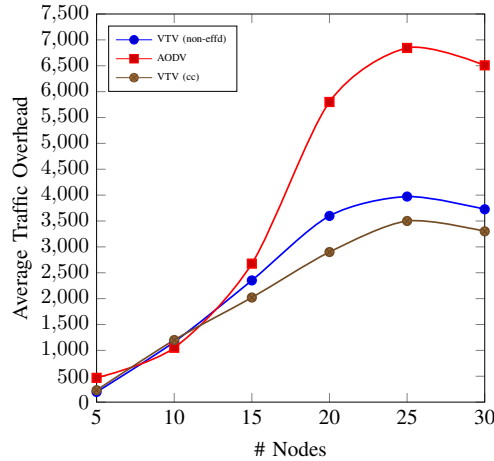Figure 6. Sparse networks: Packet Delivery Ratio.

Figure 7. Sparse networks: Average Traffic Overhead.

V. Conclusion

In this work, to further improve the performance of gossip-based routing protocols, we proposed to tune the forwarding probability based on the node's clustering coefficient. In the proposed algorithm, called VTV (cc), the nodes compute their clustering coefficient, then use it to make a packet forwarding decision. Nodes having small clustering coefficient typically interconnect two or more subnetworks: their can removal partition the network into two or more subnetworks; those nodes are required to forward messages with high probability p so as to ensure sufficient message dissemination.

Nodes whose clustering coefficient is zero are considered as a bridge nodes: to guarantee sufficient message dissemination, these nodes forward messages with a probability 1 (i.e., flood the received messages).

Nodes having a maximum clustering coefficient in their local topologies are less important in terms of disseminating control messages: to minimize their negative impact on performance, they are allowed to forward the received control messages with low probability.

So, the efficiency of clustering coefficient in identifying nodes capable of disseminating control messages sufficiently has been demonstrated by simulating VTV (cc) on different networks: VTV (cc) improves the packet delivery ratio by 21.17% and the traffic overhead by 34.44% in dense networks compared to VTV (non-effd).

References

[1] S. Joy and S. GeethaPriya, "Manets - review on broadcasting techniques," in "IJARIIE, www.ijariie.com", vol. 2, pp. –2395–4396, 2016.

[2] A. Najiya and K. Fameela, "Routing protocols in mobile ad hoc networks: A literature review," in "International Journal of Advanced Research in Computer Engineering and Technology (IJARCET)", vol. 3, pp. 2278 – 1323, 2014.

[3] M. Toqeer, N. Tabassam, A. Rehan, and M. Adnan, "Gossip based routing protocol design for ad hoc networks," in "International Journal of Computer Science", no. 1, 2012.

[4] A. Anderiya and P. Krunalkumar, "A study on the comparison of routing protocols of mobile ad hoc networks," in "International Journal of Mobile Computing and Application (SSRG-IJMCA)", vol. 4, pp. 2393–9141, 2017.

[5] E. Edwin and R. Latha, "A comparative study of routing protocols for mobile ad hoc networks," in "International Journal of Computer Science and Mobile Computing", vol. 3, pp. 46–53, 2014.

[6] Z. Haas, J. Halpern, and L. Li, "Gossiping-based ad hoc routing," in In proceedings of IEEE INFOCOM 200 The 21st Annual Joint Conference of the IEEE Computer and Communications Societies 2002 New York USA, 2002.

[7] Z. Haas, J. Halpern, and L. Li, "Gossiping-based ad hoc routing," in "IEEE/ACM Transactions on Networking (TON)", vol. 14, pp. 479–491, 2006.

[8] D. Kifle, G. Gianini, and M. Libsie, "Improving gossiping performance by means of local topology information," in *In Proceedings of the 7th International Conference on Management of Computational and Collective Intelligence in Digital EcoSystems - MEDES 2015*, pp. 142–147, 2015.

[9] D. Kifle, G. Gianini, and M. Libsie, "Local topology aware probabilistic routing," in *Proceedings of the 14th ACM International Symposium on QoS and Security for Wireless and Mobile Networks*, Q2SWinet'18, (New York, NY, USA), pp. 70–76, ACM, 2018.

[10] Z. Shi and H.Shen, "Adaptive gossip-based routing algorithm," *In Proceedings of 23rd IEEE International Performance, Computing and Communications Conference, IPCCC*, 2004.

[11] Q. Zhang and D. P. Agrawa, "Performance evaluation of leveled probabilistic in manets and wireless sensor networks," *Transactions of the Society for Modeling and Simulation International (SCS)*, vol. 8, no. 8, pp. 533–546, 2005.

[12] F. Mohammed, Y. Sufian, T. Shasikala, D. Sanjeev, and C. Michael, "Performance analysis of the gossip-based ad hoc routing using received signal strength of aodv," in *International Journal of Engineering and Technology*, vol. 6, 2016.

[13] J. Cartigny and D. Simplot, "Broder node retransmission based probabilistic broadcast protocols in ad hoc networks," in *In HICSS3: Proceeding of the 36th Annual Hawii International Conference on System Science (HICSS03), USA*, p. 303, 2003.

[14] J. Cartigny and D. Simplot, "Broder node retransmission based probabilistic broadcast protocols in ad hoc networks," in *"Telecommunication Systems"*, 2003.

[15] Q. Zhang and D. Agrawal, "Dynamic probabilistic broadcasting in manets," in *Journal of Parallel Distribution Computing*, vol. 65, pp. 220–233, 2005.

[16] A. Mohammed, M. Ould-Khaoua, and L. M. Mackenzie, "An efficient counter-based broadcast scheme for mobile ad hoc networks," in *In Formal Methods and Stochastic Models for Performance Evaluation, Springer Berlin / Heidelberg, 2007*, vol. 4748, pp. 275–283, 2007.

[17] A. Mohammed, M. Ould-Khaoua, and L. M. Mackenzie, "An adjusted counter-based broadcast scheme for mobile ad hoc networks," in *Computer Modeling and Simulation, International Conference*, vol. 0, pp. 441–446, 2008.

[18] Z. H. J. Halpern and L. Li, "Gossiping-based ad-hoc routing," in *IEEE/ACM Transactions on Networking (TON)*, vol. 14, pp. 479–491, 2006.