Advanced Adaptive Gossiping Using 2-Hop Neighborhood Information

Boto Bako, Frank Kargl, Elmar Schoch, Michael Weber
Institute of Media Informatics
Ulm University, Germany
Email: {boto.bako|frank.kargl|elmar.schoch|michael.weber}@uni-ulm.de

Abstract—Efficient information dissemination is one of the challenging tasks in most ad-hoc network application domains, be it in wireless sensor networks, mobile ad-hoc networks, or vehicular ad-hoc networks. It is obvious, that flooding has a high communication overhead, leading to channel congestion and packet collisions. Therefore, more efficient dissemination mechanisms were investigated and proposed by the research community.

One class of such algorithms is gossiping, where each node forwards a message with a certain probability. The main challenge in gossiping is the proper determination of the forwarding probability, i.e., on the one hand this value has to be chosen high enough to assure a high delivery ratio, on the other hand it has to be as low as possible to reduce communication overhead. At the same time this probability has to be adapted to dynamic network conditions, like changing node density.

In this work an advanced scheme for adaptive gossiping is proposed, capable to adjust the dissemination probability dynamically and in a distributed manner. The gossiping probability is determined at a node for every message that has to be forwarded based on two hop neighborhood information. That way, the proposed protocol is capable of working in arbitrary network topologies and densities, enabling more efficient data dissemination compared to existing approaches.

I. INTRODUCTION

In many dynamic network scenarios like Wireless Sensor Networks (WSNs) or Vehicular Ad-Hoc Networks (VANETs), researchers realized that advanced information dissemination strategies are more useful than traditional packet routing. One common case is that a message needs to be broadcast to all nodes in a network to fulfill application specific services or to enable some basic mechanisms for routing. Wireless Sensor Networks use it to disseminate TAG-type queries [1] or to broadcast control messages. Flooding as the simplest form of broadcast is used in many MANET routing protocols like DSR [2], CGGC [3] is an example of a VANET protocol using flooding for area forwarding.

Although flooding works well for a small number of nodes, channel congestion and packet collisions become major problems as network density increases. This is known as the “broadcast storm problem” [4]. In WSNs this also involves high energy consumption, drastically shortening battery durability. To address these problems, more efficient message dissemination schemes have been designed.

All these mechanisms can be categorized in either deterministic or probabilistic approaches. For a more detailed classification of broadcast mechanisms see [5] and [6]. The basic idea common to all approaches is that not all nodes need to retransmit a message. Depending on the network topology, a small subset might be sufficient, reducing the number of transmitted messages drastically while still reaching a high percentage of nodes. To evaluate and compare those broadcast protocols, mainly three aspects need to be evaluated:

- **Delivery Ratio**
  expressing the reliability of the protocol
- **Communication Overhead**
  indicating the efficiency of the approach
- **Robustness**
  indication how good a protocol copes with node failures and packet losses

Deterministic approaches explicitly select a small subset of neighbors as forwarding nodes which are sufficient to reach the same destinations as all nodes together. An example is multipoint relaying, introduced by Qayyum et al. [7]. As finding an optimal subset (i.e., with minimal size) is NP-hard, heuristics are used to find not necessarily optimal but still sufficient relaying nodes. Another example for finding such a subset of nodes is clustering. Here, forwarding nodes are called cluster-heads and gateways. An overview of clustering algorithms is given in [8].

A common problem of deterministic approaches is the lack of robustness. If a designated forwarding node fails to forward a message, e.g., due to wireless losses, the overall reliability can decrease significantly. Additionally, in energy aware networks like WSNs, the deterministic selection of the same forwarding nodes leads to quick battery consumption of these designated nodes.

Probabilistic broadcast approaches (in this work used with gossip as a synonym) don’t have these disadvantages. Because all nodes can potentially forward a message this dissemination method is more robust to link losses, node failures and so on. Also the energy consumption is distributed quite well over all nodes.

Probabilistic approaches work with retransmission probabilities determining if a node forwards a message or not. In static gossiping ([9]–[11] e.g.) all nodes have the same forwarding probability. All these approaches work only if the network characteristics are static, homogeneous, and known in advance. Otherwise they result in a low delivery ratio or a high number of redundant messages.
To address these problems, recent work proposes adaptive gossiping. Most approaches use simple heuristics to independently determine the probability at each node, e.g., considering number of neighbors or number of overheard messages. But as we will show in the next section, these methods still have problems in random topologies and as a result the overall performance might not be satisfying.

II. RELATED WORK

Haas et al. [10] introduced the so called two-threshold scheme, an improvement for static gossiping based on neighbor count. A node forwards a message with probability \( p_1 \) if it has more than \( n \) neighbors. If the number of neighbors of a node drops below this threshold \( n \) then messages are forwarded with a higher probability \( p_2 \). The obvious advantage of this improvement is that in regions of the network with sparse connectivity messages are prevented to die out because the forwarding probability is higher than in dense regions.

[10] also describes a second improvement which tries to determine if a message is "dying out". Assuming a node has \( n \) neighbors and the gossiping probability is \( p \) then this node should receive every message about \( p \cdot n \) times from its neighbors. If this node receives a message significantly fewer, the node will forward the message unless it has not already done so.

In [4], Ni et al. introduced the Counter-Based Scheme. Whenever a node receives a new message, it sets a randomly chosen timeout. During the timeout period a counter is incremented for every duplicate message received. After the timeout has expired, the message is only forwarded if the counter is still below a certain threshold value.

Although all these adaptations improve the broadcast performance, they still face problems in random network topologies. For example, if a node has a very large number of neighbors, this results in a small forwarding probability in all of these schemes. Despite this, there could e.g. still be an isolated neighbor which can only receive the message from this node. An example of such a situation is shown in Figure 1 (example taken from [12]).

![Fig. 1. Sample topology where static gossiping fails](image)

When node A sends a message, all nodes in its neighborhood receive it. In this example scenario only node E should forward it with the probability of 1 since E is the only node that can propagate the message to node G. If the gossiping probability is only based on the neighbors count, node E will be assigned a low probability since it has many neighbors. So the broadcast message will "die out" with a high probability and never reach G and all later nodes. If the part of the network connected only via G is very large, the overall delivery ratio will drop dramatically. Such situations can occur quite regularly in dynamic networks of a certain density.

Recently, the Smart Gossip protocol which addresses this problem was introduced by Kyasanur et al. [12]. The protocol assumes a static network with one single message source. Every node in the network uses neighborhood information from overheard messages to build a dependency graph. I.e., each node has a parent, sibling, and child set. Parents are nodes where a node receives new messages from, siblings receive messages from the same parents and a node delivers messages to child nodes.

Depending of the number of parents, every node calculates the probability by which its parents should forward a message and informs its parents about this probability. A parent sets its forwarding probability to the maximum of all child probabilities. If a node has only one parent, the forwarding probability will be automatically set to 1. If a node has many parents, the probability will be comparatively small, but still large enough to ensure that the node will most likely receive the message at least once.

Smart Gossip solves the problem of adapting the forwarding probability dynamically, but in some cases there are still disadvantages: the described parent child relationship is dynamic, depending on which node sends or forwards a message. For ensuring to build up a stable directed graph, the authors make the assumption that there is only one message originator in the whole network. This assumption may be sufficient in a few scenarios, but in this work we want to evaluate the protocols for a broader range of applications.

In a recent work, we proposed an extension of the Smart Gossip protocol in [13]. The main objective of that work is to overcome the problem of multiple senders. Position information is used to build up the parent-child-relationship and prove that with this extension the protocol can be applied in high dynamic networks with multiple message originators. Additionally, node mobility is considered and good results are achieved. We will later use this and the original Smart Gossip protocol to compare the performance of the protocol proposed in this work.

III. THE ADVANCED ADAPTIVE GOSSIPING PROTOCOL

In this section the proposed Advanced Adaptive Gossiping protocol (AAG) is described which is based on a heuristic reduction of the forwarding probability depending on the number of one- and two-hop neighbors.

A. Protocol description

Many ad-hoc network applications require neighborhood information realized through so called hello messages or beacons. The proposed protocol makes use of such information to calculate the gossip probability. Therefore, each node periodically sends a beacon containing its neighbor table. This way all nodes know their 2-hop neighborhood.

The basic idea behind the proposed protocol is the following: a node X receiving a message from node S determines its neighbors which according to the 2-hop neighborhood table could not receive that message from S – these are called X’s child nodes. Next, X determines all common neighbors...
between $S$ and each child node. These nodes are potentially forwarders for messages sent from $S$ to that child and are called *parents*. If the number of parents of a child is known, then a very efficient forwarding probability can be computed for that child. Finally, $X$ computes the forwarding probability for each child and forwards a message with the maximum of all calculated probability values. The higher the number of parent nodes, the lower can the forwarding probability be. In the case when $X$ is the single parent of a child node, it forwards the message with a probability of 1.0 since that child depends completely on $X$ for messages from $S$.

In the following a more formal definition of the algorithm is given. Let $N$ be the set of all nodes in the network and $X \in N$ be a forwarding node, which receives a message $M$ from a sender $S \in N$. $S$ is a 1-hop neighbor of $X$ ($S \in \text{neighbor}(X)$). The set of nodes that received the message $M$ is called $M_r$ and is equal to the neighbors of $S$. $X$ first determines its neighbors which have not received the message:

$$\text{child}(X) = \{ n \in N \mid n \in \text{neighbor}(X) \land n \notin M_r \}$$

Having the child set of node $X$, for each child $c_i \in \text{child}(X)$ all nodes are determined which are possible forwarders of message $M$:

$$\text{parent}(c_i) = \{ p \in N \mid p \in \text{neighbor}(c_i) \land p \in M_r \}$$

Based on the number of parents ($\#\text{parent}(c_i)$) of a child node, the probability $p_i$ for child $c_i$ is calculated as:

$$p_i = \text{prob}(\#\text{parent}(c_i), \delta)$$

where $\delta$ is the average diameter of the network (more precisely of the broadcast area) measured as hop count. The diameter is important since with each forward the gossip probabilities are multiplied and thus they become lower and lower. Therefore, if this parameter is not considered, broadcast messages tend to “die out” in larger networks. Subsection III-B describes this function in detail.

If $\#\text{parent}(c_i) = 1$ a gossip probability of 1 is assigned since node $X$ is the only parent. Finally, node $X$ forwards the message with a probability of $p = \max(p_1, ..., p_n)$.

This calculation is made at every node which receives a new broadcast message. From the 2-hop neighborhood information very accurate forwarding probabilities can be determined, as we consider the number of nodes which can potentially forward a message and adapt the forwarding probability dynamically based on the current topology. Before going into details of the probability determination function $\text{prob}(\#\text{parent}(c_i), \delta)$, the mode of operation of the proposed adaptive gossip protocol is explained with the help of an illustrative example.

Assume a network topology – or a sub-network – as shown in Figure 2. All nodes know their two hop neighbors because they are constantly gathering this information e.g. through beacon messages. Let node 1 be the initiator of a broadcast message, which can be received by the nodes 2 and 3. Node 2 determines its child nodes, only node 4 in this case, since node 3 has already received the message from 1. Next, node 2 searches for parents of node 4 – nodes which can forward the message to it. There are two such parent nodes: node 2 itself and node 3. Based on this information, node 2 computes the gossip probability with a parent count value of 2. Node 3 performs the same steps as node 2. Its result differs as it has nodes 5 and 6 as additional child nodes. Since it considers itself to be the only node which can forward to them this message, its gossip probability is set to 1. After that node 4, 5 and 6 receive the message from node 3 and perform the same steps. Duplicate messages are simply ignored.

![Fig. 2. A sample topology](image)

### B. Determination of the gossip probability

So far, the AAG protocol was described in detail, now we take a closer look at the probability determination function $\text{prob}(\#\text{parent}(c_i), \delta)$. This is the same function as used in the Smart Gossip protocol, introduced by Kyasanur et al. in [12]. The authors assume that an application can specify its reliability requirement as an average reception percentage $\tau_{arb}$. For example an application can determine that it requires 99% reception rate. Since the actual reception rate becomes smaller with every forwarding step, it is translated into a so called per hop reception probability $\tau_{rel}$ which considers also the network diameter $\delta$ to ensure the required receptions rate also in large networks. $\tau_{rel}$ is thus the probability a node uses to forward a message. This probability can be determined by the following equation:

$$(\tau_{rel})^\delta = \tau_{arb}$$

where $\delta$ represents the diameter of the network.

If a child has only one parent then the message has to be forwarded at least with the probability $\tau_{rel}$ (in the proposed protocol the probability is set to 1 in this case). However, if a node has more parents then it is sufficient to forward with a lower probability $p_{\text{forward}}$. This probability has to be set to a value which ensures that a child receives the messages with at least the probability $\tau_{rel}$ from its parents. Knowing the number of parents ($K = \#\text{parent}(c_i)$), the forward probability can be determined with the following equation:

$$(1 - p_{\text{forward}})^K < (1 - \tau_{rel})$$

Using this function for probability calculation, our evaluation shows that we can achieve both a high delivery ratio and a good efficiency due to a low forwarding rate. In the next subsection an optimization of the protocol is described which lowers the forwarding ratio even further.
C. Network density based enhancement

While performing simulations for this work, we discovered that the network density influences the considered broadcast schemes. In an ideal case, the broadcast scheme should reduce the forwarding probability drastically in very dense networks, whereas in very sparse networks it should fall back to pure flooding. As we will show in the evaluation section, the Smart Gossip protocol [12] fails to manage this adaptation. I.e., with growing node density the forwarding probability is not lowered properly, resulting in a very high forwarding ratio.

Our protocol performs much better: the forwarding ratio decreases constantly with growing density. Nevertheless, a significant improvement is still possible which is described here. We recognized in our simulations that despite a lower application reception probability requirement (\(\tau_{\text{arp}}\)) the reception rate remained much higher. Especially in dense networks \(\tau_{\text{arp}}\) can be set to a much smaller value without a drastic impact on the delivery ratio. We therefore suggest to introduce an additional reduction factor \(\text{red}\) that depends on the network density. This reduces the overhead while still meeting the application requirements.

\[
\text{red}(D) = \frac{0.5}{1 + e^{-a(D-b)}},
\]

where \(a\) and \(b\) are two parameters determined empirically through simulations. The result of this function in relation to varying connectivity \(D\) is shown in Figure 3. As shown, \(\text{red}\) is almost zero for low connectivity, whereas the reduction value rises up to 0.5 for higher connectivity. Applying this reduction value, the \(\tau_{\text{arp}}\) value is replaced by: \(\tau'_{\text{arp}} = \tau_{\text{arp}} - \text{red}\).

The exact performance gain of this enhancement as well as an extensive comparison with the other probabilistic broadcast protocols is shown in the evaluation section.

IV. PROTOCOL EVALUATION

For the evaluation, we used the Java based network simulator JiST/SWANS [14] to compare the performance of the Advanced Adaptive Gossipping protocol (AAG) proposed in this work with the Smart Gossip protocol from [12] and the 

Position based Gossip (PbG) protocol [13]. As Link-/MAC-layer IEEE 802.11b is used, wireless transmission range is set to 280 meters. For each simulation setup 100 simulation runs are done and the results averaged.

The simulation setups are divided into three parts: in the first setup we use similar simulation parameters as in [12]. We evaluate the performance – reliability and efficiency – of these protocols for one message originator and varying node densities. In the second setup we analyze the impact of multiple message originators onto these protocols. Thus, the performance is evaluated in a scenario with high network traffic. In the last setup, node mobility is considered. Since the Smart Gossip protocol wasn’t designed to cope with mobile nodes we present the results only for the proposed and the position based gossip protocols and show their applicability in such mobile environments.

![Fig. 4. Average forwarding rate with 50 to 300 nodes in a 1000 × 1000 area, one sender](image-url)
Forwarding ratio alone is insufficient to measure the performance of a protocol. For being able to compare these protocols, the delivery ratio is also needed. In [13] a new metric was introduced which combines the measured values – forwarding and reception ratios – and enables thus a better comparison between the protocols. This combined metric is defined in the following way:

\[
\text{Efficiency Rate} = \frac{\text{Reception Rate}}{\text{Forwarding Rate}}
\]

<table>
<thead>
<tr>
<th>Nodes</th>
<th>SMART GOSSIP</th>
<th>PbG</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>99.97%</td>
<td>88.15%</td>
</tr>
<tr>
<td>100</td>
<td>99.99%</td>
<td>99.09%</td>
</tr>
<tr>
<td>150</td>
<td>99.99%</td>
<td>99.93%</td>
</tr>
<tr>
<td>200</td>
<td>99.99%</td>
<td>99.93%</td>
</tr>
<tr>
<td>300</td>
<td>99.99%</td>
<td>99.99%</td>
</tr>
</tbody>
</table>

**TABLE I**

**SMALL Gossip and PbG: average reception, forwarding and efficiency rates with 50 to 300 nodes in a 1000 x 1000 area, one sender**

<table>
<thead>
<tr>
<th>Nodes</th>
<th>AAG Log</th>
<th>AAG</th>
<th>SMART GOSSIP</th>
<th>PbG</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>99.71%</td>
<td>87.31%</td>
<td>1.36</td>
<td>98.08%</td>
</tr>
<tr>
<td>100</td>
<td>99.99%</td>
<td>99.09%</td>
<td>1.46</td>
<td>99.99%</td>
</tr>
<tr>
<td>150</td>
<td>99.99%</td>
<td>99.93%</td>
<td>1.44</td>
<td>99.99%</td>
</tr>
<tr>
<td>200</td>
<td>99.99%</td>
<td>99.93%</td>
<td>1.52</td>
<td>99.99%</td>
</tr>
<tr>
<td>300</td>
<td>99.99%</td>
<td>99.99%</td>
<td>1.43</td>
<td>99.99%</td>
</tr>
</tbody>
</table>

**TABLE II**

**AAG Log and AAG: average reception, forwarding and efficiency rates with 50 to 300 nodes in a 1000 x 1000 area, one sender**

With higher reception rate and lower forwarding rate the efficiency rate grows, i.e. the overall performance of the protocol is better. Table I and II depict the forwarding and reception rates of the evaluated protocols together with their efficiency values. As shown, Smart Gossip achieves the highest reception rate. But the other protocols are only marginally lower and regarding the forwarding ratio, they clearly outperform the Smart Gossip protocol. The proposed AAG protocol outperforms both Smart Gossip and PbG considerably regarding forwarding rate and efficiency rate. Considering AAG with the logistic improvement, the efficiency metric is about 3.7 times higher than the efficiency of Smart Gossip in a network with 300 nodes. But even AAG without the logistic improvement has a much better performance than the other two protocols. Altogether, AAG has the best overall performance when regarding reception rate per forwarding rate.

In the next simulation setup we analyze the protocols with multiple senders. Therefore, every node can potentially initiate a broadcast message every second. This way the messages are initiated in parallel and results therefore in much higher network traffic as in the last setup. The total number of broadcast messages is set to 150 as in the last simulation setup. These simulation results confirm our assumption that Smart Gossip fails in a scenario with multiple senders. The protocol’s reception rate drops drastically as shown in Figure 5. This is due to the fact that Smart Gossip builds up a directed graph which is used to compute the forwarding probability. In the case of multiple message originators this graph can’t be built up correctly. For a better comparison we also included flooding into these results. As shown in the figure, the reception rate (which is also the forwarding rate in this case) is quite good in low density networks but with increasing density the reception rate also drops notably.

On the other hand, AAG achieves the performance of flooding in low densities regarding the reception rate and outperforms it notably in higher densities as shown in Figure 6. At the same time the forwarding ratio is very low. AAG with the logistic function modification achieves a lower forwarding ratio in higher densities, but the delivery ratio also drops slightly. The reason is that without the modification the protocol has more redundancy, and therefore single packet losses don’t have such an impact. In this case it can’t be said which variant is more effective, it depends on application specific requirements. The reception rate of the PbG protocol is on average between the two variants of AAG. However, its forwarding rate is very high compared to AAG. Thus, the overall performance of AAG is clearly superior to the other protocols also in this scenario, underlined by the detailed results shown in Table III and IV.

In the last simulation configuration node mobility is considered. Therefore, 50 nodes are randomly placed onto a 1000 x 1000 meters field and are moving with varying speeds between 5 and 30 meters per second. All other simulation parameters
are the same as in the last setup. As shown in Table V and VI, the proposed protocol is almost not affected by node mobility. Only a slight decrease of the reception and forwarding rate can be stated. AAG with the logistic function modification is in this scenario more effective since its reception rate is almost the same but at the same time with a significantly lower forwarding rate. PbG can cope as well with node mobility but its performance is clearly beneath of AAG: for all node speeds its reception rate is lower than of AAG, at the same time having a higher forwarding rate.

<table>
<thead>
<tr>
<th>Node speed (m/s)</th>
<th>Reception</th>
<th>Forwarding</th>
<th>Efficiency</th>
<th>Reception</th>
<th>Forwarding</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>59.14%</td>
<td>25.89%</td>
<td>70.1%</td>
<td>49.97%</td>
<td>15.83%</td>
<td>62.1%</td>
</tr>
</tbody>
</table>

TABLE V
AAG LOG AND AAG: AVERAGE RECEPTION, FORWARDING AND EFFICIENCY RATES WITH 50 NODES IN A 1000 × 1000 AREA, MULTIPLE SENDERS, VARYING NODE SPEEDS

<table>
<thead>
<tr>
<th>Node speed (m/s)</th>
<th>Reception</th>
<th>Forwarding</th>
<th>Efficiency</th>
<th>Reception</th>
<th>Forwarding</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>61.13%</td>
<td>28.67%</td>
<td>69.96%</td>
<td>51.76%</td>
<td>16.64%</td>
<td>63.14%</td>
</tr>
</tbody>
</table>

TABLE VI
PbG: AVERAGE RECEPTION, FORWARDING AND EFFICIENCY RATES WITH 50 NODES IN A 1000 × 1000 AREA, MULTIPLE SENDERS, VARYING NODE SPEEDS

V. SUMMARY AND OUTLOOK

In this work a new adaptive probabilistic broadcast protocol is proposed that makes use of two-hop neighborhood information to dynamically calculate adequate forwarding probabilities. As the evaluation section shows, this protocol is well suited for a wide range of application scenarios; it delivers a good performance in static as well as in mobile networks with varying node densities and traffic volume. Its superiority to existing probabilistic approaches is proved by extensive simulations. In summary, the proposed protocol is highly adaptive and delivers a very good performance, measured in reception rate per delivery ratio.

Nevertheless, there are further improvements possible that are subject to future work. One such improvement would be the implementation of a mechanism to handle packet losses. Moreover, further improvements depending on the scenario the protocol is designed for are possible. For example in vehicular ad-hoc networks, position information are available which could be used to compute even a more precise probabilistic value for forwarding.

REFERENCES


978-1-4244-2324-8/08/$25.00 © 2008 IEEE.