Open access • Proceedings Article • DOI:10.1109/MASS.2012.6502543

Collaborative assessment of functional reliability in wireless networks

Zi Feng, Konstantinos Pelechrinis, Srikanth V. Krishnamurthy, Ananthram Swami ...+2 more authors

Institutions: University of California, Riverside, University of Pittsburgh, United States Army Research Laboratory, University of California, Davis ...+1 more institutions

Published on: 08 Oct 2012 - Mobile Adhoc and Sensor Systems

Topics: Wireless network, Routing protocol and Flooding (computer networking)

Collaborative Assessment of Functional Reliability in Wireless Networks

Zi Feng*, Konstantinos Pelechrinis&, Srikanth Krishnamurthy*, Ananthram Swami†, Felix Wu‡, Munindar P. Singh#
*UC Riverside, &University of Pittsburgh, †US Army Research Labs, ‡UC Davis, # North Carolina State University
zfeng@cs.ucr.edu, *kpele@pitt.edu*, *krish@cs.ucr.edu*, *aswami@arl.army.mil*, *wu@cs.ucdavis.edu*, *singh@ncsu.edu*

***Abstract*—Nodes that are part of a multihop wireless network, typically deployed in mission critical settings, are expected to perform specific functions. Establishing a notion of reliability of the nodes with respect to each function (referred to as functional reliability or FR) is essential for efficient operations and management of the network. This is typically assessed based on evidence collected by nodes with regards to other nodes in the network. However, such evidence is often affected by factors such as channel induced effects and interference. In multihop contexts, unreliable intermediary relays may also influence evidence. We design a framework for collaborative assessment of the FR of nodes, with respect to different types of functions; our framework accounts for the above factors that influence evidence collection. Each node (say Chloe) in the network derives the FR of other nodes (say Jack) based on two types of evidence: (i) *direct* evidence, based on her direct transactions with each such node and (ii) *indirect* evidence, based on feedback received regarding Jack from others. Our framework is generic and is applicable in a variety of contexts. We also design a module that drastically reduces the overhead incurred in the propagation of indirect evidence at the expense of slightly increased uncertainty in the assessed FR values. We implement our framework on an indoor/outdoor wireless testbed. We show that with our framework, each node is able to determine the FR for every other node in the network with high accuracy. Our indirect evidence propagation module decreases the overhead by 37% compared to a simple flooding based evidence propagation, while the accuracy of the FR computations is decreased only by 8%. Finally, we examine the effect of different routing protocols on the accuracy of the assessed values.**

I. Introduction

In mission-critical deployments (e.g., tactical missions, disaster recovery) of multihop wireless networks, nodes are expected to perform specific functions (such as forward packets or respond to queries). The reliability of nodes in performing these functions, referred to as functional reliability or FR, is critical for the efficient operations and management of a network. Other nodes may rely on those nodes that are deemed reliable in performing a desired function. We defer a formal definition of FR to Section III. Roughly, the FR of a node with respect to a specific function is the reliability (or responsiveness) of that node in performing the function. A node may become functionally unreliable for various reasons; e.g., it may misbehave due to a low battery or being disconnected, misconfigured, or compromised.

Assessing a node's FR in a wireless network is challenging. First, links are lossy; second interference may cause unreliable operations or faulty observations. Finally, information is relayed by other users (nodes), who themselves may not be completely reliable. To our best knowledge, the dependencies between these factors and a user's FR have not been previously investigated.

We design a framework accounting for the above factors wherein nodes collaboratively assess the FR of every other node. Every node (say Chloe) maintains an *FR tuple* with respect to every other node (say Jack). Each element of the tuple corresponds to Jack's assessed FR with respect to a different functionality (e.g., routing/forwarding, responding to queries etc.). For instance, when Chloe wants to assess the end-to-end (e2e) FR of Jack (whether he is reliable in responding to a query), she accounts for the possibility that a transaction may fail due to wireless induced effects or due to an unreliable relay. In particular, Chloe builds evidence for Jack based on the *direct* transactions she has with him. This is referred to as direct evidence. Note that direct evidence does not mean that there exists a physical one-hop distance between Chloe and Jack. "Direct" here pertains to the fact that Chloe gathers this evidence solely based on her transactions with Jack. Based on a series of such transactions, Chloe applies the Maximum Likelihood Estimation (MLE) framework, to estimate a direct FR value for Jack. Here, she accounts both for the FR of each relay on the path to Jack in forwarding packets (forwarding FR), as well as the qualities of the links en route to Jack.

Chloe then combines the above *direct FR* for Jack, with feedback relating to Jack from other users (say Tony) using a *gossiping* scheme; this is referred to as *indirect* evidence. The direct FR is combined with this *indirect FR* using the Dempster-Shafer theory of evidence (DSTE). Indirect evidence is vital since Chloe may not sufficiently interact with Jack; in some extreme cases she may have no transactions at all and may have to rely on other nodes to assess Jack's FR.

The transactions between the FR assessment process and the different network functionalities have complex interdependencies. On the one hand, the assessed values can influence various network functionalities (e.g., relay node selection). On the other hand, the FR inference engine can itself be affected by the operations of various network protocols. For instance, different routing metrics, will result in the use of different paths; the choice of the path influences the evidence collected for FR assessment. In our work, we also experimentally assess the impact of various routing policies on FR assessment.

In brief, our main contributions are summarized below:

(a) We design a collaborative FR assessment framework that jointly considers the impact of the unique aspects of a wireless network (i.e., channel related effects and unresponsive relays). To our best knowledge, this is the first framework to jointly consider these factors.

(b) We incorporate a lightweight evidence propagation scheme in our framework, which intelligently filters duplicated evidence and reduces the message complexity of propagation from $O(N^2)$ to $O(N)$ (N is the number of nodes in the network). The reduction in message complexity comes at a price—increased uncertainty in the reliability computations.

(c) We implement and evaluate our scheme on our wireless indoor/outdoor 802.11 testbed. Our experiments show that each node infers the FR values for every other node in the network with high accuracy. Our lightweight evidence propagation scheme reduces the propagation overhead by 37% compared to a simple flooding based evidence propagation.

(d) We experimentally examine the impact of using different routing metrics on FR establishment.

Scope of our work: As implicitly alluded to earlier, reliability is typically function dependent. Jack may forward packets destined to Chloe. However, he may not reply to e2e queries for a specific application because the corresponding

application software (residing in his machine) is malfunctioning or he is restricted by policy (Chloe may be unaware of this). Our proposed framework is generic and can be used to assess the FR relating to various wireless network functional contexts. We showcase our framework by assessing e2e (response to queries) and forwarding FRs. However, the applicability of our framework is not limited to these contexts.

A limitation of our approach is that its assesses FR based on Boolean outcomes (e.g., *Did Jack respond to a query?*). It does not take into account possible subjectivity in assessing an outcome. Further, if for example, the question is *Did Jack provide the relevant information in response to a query?*, there may be a response that indicates that he only provided partial information. The extent of this partial information is not accounted for and the system just counts the observation to indicate a success or a failure. Taking into account subjectivity of observations and partially successful outcomes is beyond the scope of this work. We emphasize that our framework is designed to capture the average FR based on observations over sufficiently long periods. It does not address short-term trust variations. Likewise, we do not consider security aspects such as nodes that lie or collude.

Organization: The paper is organized as follows. Section II discusses related work and provide background for our framework. Section III describes our FR establishment scheme. Section IV presents our lightweight evidence propagation mechanism. Section V presents our implementation and the evaluations of our scheme and Section VI our conclusions.

II. Related Work and Background

Related Work: Wireless multihop networks require users to perform specific functions for required network operations. There exists work in the literature to determine whether or not nodes are performing their functions in a non-cooperative setting. Specifically, reputation systems to evaluate, and incentive-based mechanisms to encourage cooperation and functional compliance, have been studied. While our work is similar in spirit to reputation systems, we believe we are the first to account for wireless effects and the impact of other unreliable nodes while estimating the functional reliability of nodes.

Reputation systems: Marti *et al.* [1] propose a scheme for identifying reputable nodes with respect to the routing functionality. They propose *watchdogs* that identify nodes that drop packets based on promiscuous observations and a *pathrater* that avoids paths with such misbehaving nodes. CONFIDANT [2] [3] seeks to identify the routing reliability of nodes. The architecture is similar to that of [1]; a monitoring system is used along with reputation and path selection mechanisms (no details are provided on how the reputation of a node is updated in time). The above schemes focus only on the routing/forwarding functionality. Moreover, they do not account for loss of information due to channel induced effects. Michiardi *et al.* [4] design CORE, which is the first work to define functional reputation. A node might have different reputation values for different network functionalities. Without getting into the details on every possible functionality, the authors present a general scheme that makes use of observations from the users of the network to estimate the functional reliability of a user. What is missing from the above scheme however, is that it does not account for the effect of wireless induced factors or interference while assessing reputation. *In summary, none of the above studies account for the impact of the unique factors that exist in a wireless network, on the estimation process.* To our best knowledge, we are the first to account for wireless induced factors and the network functional context while assessing the FR of a node.

Trust Assessment: *Trust* assessment is loosely connected to our work. Probst *et al.* [5] propose local *trust* computations based only on neighbors' past behaviors. They do not consider aggregation of trust values and their scheme is specific to the topology and density of the network. Velloso *et al.* [6] present an approach which combines local measurements with aggregated trust values computing a weighed trust. However, they do not provide a method to efficiently propagate these values in the network. The interested reader can find a detailed study on trust management in [7]. In contrast, our work is focused on the assessment of the FR of a node (not trustworthiness) in a wireless network, taking into account wireless induced factors.

Incentive-based mechanisms: Buttyan and Hubaux's [8] scheme provides incentives for users to cooperate and forward packets for other users; however, it does not provide a rating mechanism for the users. Users need to pay credits in order to get their packets forwarded. Relays can accumulate credits for future use; a node that does not have enough credits cannot use the network services itself. SPRITE [9] also uses credits to provide incentives to selfish users to cooperate; however, it does not require any tamper-proof hardware as is the case with [8]. Our work is on assessing the functional reliability of nodes and does not design methods toward ensuring compliance of non-cooperative nodes.

Dempster Shafer Theory of Evidence: DSTE is a generalization of the Bayesian inference theory. Based on evidence from one or more observations (possibly by different entities called *sensors*) of a system, DSTE estimates the system's state.

Let us assume that Θ, is the set of all possible states of the system and H (hypothesis) is a subset of Θ. Every sensor that reports evidence is described by a *Basic Probability Assignment (bpa)*, m, representing a *"measure of belief committed exactly at (each) H"* [10]:

$$m : 2^{\Theta} \rightarrow [0, 1] \quad (1) \qquad m(\emptyset) = 0 \quad (2)$$

$$m(H) \geq 0, \forall H \subseteq \Theta \quad (3) \qquad \sum_{H \subseteq \Theta} m(H) = 1 \quad (4)$$

Defining the belief (Bel) and plausibility (Pl) of H as:

$$Bel(H) = \sum_{B \subseteq H} m(B), \;\; Pl(H) = \sum_{B \cap H \neq \emptyset} m(B), \quad (5)$$

the true belief on H lies within the interval $[Bel(H),\;\; Pl(H)]$.

In the case of multiple sensors reporting independent evidence for the system's state, the DSTE rule of combination (also known as **orthogonal product** $\oplus$) can be used. In particular, let's assume that we have two sources of independent evidence with assigned bpas m_1 and m_2, respectively. Then, these two sources of evidence can be combined to form a single source of evidence with bpa, $m_{12}(\Theta)$ for hypothesis Θ:

$$m_{12}(\Theta) = m_1 \oplus m_2 = \frac{\sum_{B \cap C = H} m_1(B) m_2(C)}{\sum_{B \cap C \neq \emptyset} m_1(B) m_2(C)} \quad (6)$$

Intuitively, since the two sources of evidence are independent, the product of the corresponding bpas for the two hypotheses (e.g., B and C) gives the belief value on their intersection. As a result, Eq. 6 provides the portion of the total belief committed to hypothesis H from both sources of evidence. The numerator computes the belief on H, since B

and C are constrained to the pair of sets whose intersection is H, while the denominator computes the total belief ($B \cap C \neq \oslash$). More details on DSTE can be found in [10] and [11].

III. Assessing Functional Reliability

We now describe our FR assessment framework. Formally, the reliability or responsiveness of a node with respect to a function (or operation) is the likelihood that it will perform the function. For instance, if Chloe seeks some information from Jack, Jack's FR (from Chloe's perspective) reflects the likelihood that he will respond to that query. In a different context, if Chloe relies on Jack to forward her packets to Bob, Jack's FR captures the likelihood that he will relay her traffic toward the destination. As will be clear in the following, Jack is associated with a tuple of FR values, whose elements embody the likelihood that Jack reliably performs a corresponding network function.

Our approach in brief: FR (as defined above) is assessed based on a node's own transactions with a peer[1] and responsiveness information (relating to the same peer) obtained from other nodes. We refer to the former as *direct evidence* and the latter as *indirect evidence*. The details of the transactions depend on the specific function considered. Irrespective of the specifics of a function, given a series of observations (of whether or not the peer performed the function), we use MLE to determine the probability that the peer is reliable with regards to the particular function. For instance, Chloe establishes direct evidence on e2e transactions with regards to Jack based on the success or failure of her transactions with him. We take into account the "forwarding reliability" of relay nodes[2] en route the peer (say a vector T), and the qualities of the links on the path used for the transaction (say a vector Q). These factors capture the possibilities that a transaction may fail not because the end peer did not respond, but because of link failures or an intermediate node being unreliable with respect to forwarding traffic. Note that the cardinality of T is the number of relays on the route and that of Q is the number of links on the route. We apply MLE to determine the probability that a peer is reliable in responding to the e2e queries, given a series of observations and the vectors T and Q, associated with each transaction attempt.

The FR established based on direct evidence is next updated based on indirect evidence, i.e., through **gossiping** with other nodes. We incorporate a degree of uncertainty in the computed values as explained later. For simplicity, the term reliability or responsiveness (FR) refers to e2e reliability unless explicitly specified. We discuss the applicability of our framework in other contexts in Section III-D.

A. FR representation

If one were to have a strict notion of FR, it should be represented by a binary variable Z; Z is 0 if the node is unreliable and, 1 otherwise. However, in reality there is an uncertainty associated with FR and thus, we denote Z to be the likelihood or probability that the node is responsive (with respect to a function) and hence, $Z \in [0, 1]$.

However, this single *crisp* value does not capture the *degree* of uncertainty with regards to the peer entity under discussion. To account for this, the actual value is considered to lie within $I = [a, b] \subseteq [0, 1]$. The interval signifies the uncertainty associated with the determination of the probability; its width captures uncertainty that we have in our estimation. However, in some parts of the paper, we will reduce this interval I to a single point value r, through a function h for clarity and tractability. Specifically, the function returns the mean value of the interval I i.e., $r = h(I) = \frac{a+b}{2}$. One can easily use other functions such as $min\{a, b\}$ or $max\{a, b\}$ instead.

We assume that nodes either have a priori perceptions of initial FR levels with respect to other nodes (as an example, a resource rich node may initially be deemed completely reliable), or that every node is reliable or unreliable with an equal likelihood (i.e., each node associates an FR value in the interval $[0.5-\epsilon, 0.5+\epsilon]$ for all other nodes). These (initial) values dynamically evolve as entities interact. If a node is responsive, it should eventually be deemed reliable with a low uncertainty.

B. Updating FR values based on direct evidence

The first source of evidence for a node's (say Chloe's) view of the FR of a peer (say Jack) originates from Chloe's direct transactions with Jack. The outcomes of these transactions via a wireless network, depend on 3 factors: **(i)** the *forwarding FR* of intermediate nodes that are responsible for relaying the transaction data, **(ii)** the wireless link qualities on the route R from Chloe to Jack, **(iii)** Jack's reliability (which Chloe wants to estimate).

In order to perform her estimation, Chloe monitors the outcome of k consecutive direct transactions with Jack. These observations form a sample set, indexed by j. For each transaction i Chloe records the outcome, e_i, the probability that the communication path meets the requirements of the application, Q_i, and the forwarding FR of the path, T_i. For a successful transaction, we have $e_i = 1$; otherwise $e_i = 0$. Q_i depends on the specifics of the e2e transaction considered. When only the delivery of the transaction packets is required (e.g., no delay constraints), Q_i is the delivery probability on the route R_i, followed for the transaction i; this is estimated based on the link quality q_l of each of the intermediate links l of the route from Chloe to Jack. Here, q_l is essentially the Packet Delivery Ratio (PDR) on link l and it can be calculated by having neighbor nodes exchange probe packets[3] [12]. Section III-D examines transactions with different requirements and their mapping onto Q_i. T_i is calculated based on the forwarding reliability intervals I_j of the intermediate nodes j that comprise the route R_i. In particular:

$$Q_i = \prod_{l \in R_i} q_l, \; T_i = \prod_{j \in R_i} h(I_j) \tag{7}$$

The above equations assume the independence (i) of the quality of the links on a route and (ii) of the forwarding FR of the intermediate relay nodes. In practice, there may be correlations. First, the projected interference (which affects the quality of the links) on consecutive links may not be independent. Second, the forwarding reliability of the intermediate relays may depend on evidence from common sources causing the independence assumption to not hold. We make the independence assumption due to the complexity in modeling correlations; however, our evaluations suggest that in spite of these assumptions, our models work well in practice.

Let us assume that Chloe associates with Jack a reliability value of p_i during her i^{th} transaction with him. Then, it is easy

[1] We will use the terms peer and node interchangeably in the rest of the paper.

[2] As explained later, "forwarding FR" is assessed using a different set of observations but using the same statistical framework.

[3] We assume that this function(exchanging probe packets) is reliable; however, our framework could be used to assess the reliability of this function as well.

to see that the i^{th} transaction is a Bernoulli trial X, with a probability of success $p_i \cdot Q_i \cdot T_i$. Thus, the pdf of X is:

$$f_i(X = e_i) = (p_i \cdot Q_i \cdot T_i)^{e_i} \cdot (1 - p_i \cdot Q_i \cdot T_i)^{1-e_i} \quad (8)$$

We use the MLE method [13] to update the estimate of the FR of Jack, p, based on the current trial and the previous $k-1$ trials. Then, Chloe's view of Jack's FR is the solution to the optimization problem:

$$\max_{p_j} \quad \frac{1}{k} \cdot \sum_{i=1}^{k} log(f_i(e_i|p_j)) \quad (9)$$

$$p_j \in [\widehat{p}, 1] \quad (10)$$

where p_j is the FR estimate based on sample set j. Given $\overrightarrow{e_j}$, Jack's FR cannot be smaller than the percentage of successful transactions in $\overrightarrow{e_j}$. When $\overrightarrow{e_j} = \oslash$, $\widehat{p}$ captures the non-zero probability that all Chloe's transactions with Jack in the sample window fail due to wireless induced failures or unreliable intermediaries. Considering that Jack is 100% reliable, this probability is equal to $x = \prod_{i=1}^{k}(1 - Q_i \cdot T_i)$. If $x = 0$, then all the transactions failed due to Jack and hence, $p = 0$. As x increases, the minimum FR of Jack increases as well. Even if all transactions failed due to bad links or non-reliable relays (i.e., $x = 1$), Jack cannot be deemed 100% reliable. In fact here, Chloe does not know anything about Jack, which implies that $p = 0.5$. As we see, there is a dependence between p and x (i.e., $p = f(x)$ for some function $f()$). Assuming, for simplicity, a linear relation between x and p we can calculate the minimum FR of Jack in the average case to be:

$$\widehat{p} = \begin{cases} (\sum_{i=1}^{k} e_i)/k & \text{if } \overrightarrow{e_j} \neq \oslash \\ (\prod_{i=1}^{k}(1 - Q_i \cdot T_i))/2 & \text{if } \overrightarrow{e_j} = \oslash \end{cases} \quad (11)$$

p_j cannot be smaller than $\widehat{p}$, the min FR of Jack as per Chloe's view. Note here that, **the optimization problem Eqs. (9)–(10) always has a solution** since the objective function is continuous, and is constrained on a closed and bounded set.

Considering one sample set j and solving the MLE problem provides Chloe with a single point estimate $\widehat{p_j}$. In order to compute the uncertainty on the FR value, she uses m consecutive sample sets, i.e., a *sliding window* of samples. In particular, if the first sample set consists of the observations indexed by $\{1, 2, \ldots k\}$, the second sample set consists of the observations $\{2, 3, \ldots, k+1\}$, and so on. Using the estimates computed from MLE for each of the above sets, Chloe computes the average estimator $\overline{\widehat{p}}$ and its standard deviation $\widetilde{p_{sd}}$. Then for the real FR value p^*, the following approximations hold:

$$p^* \in [p^*_{min}, p^*_{max}] \quad (12)$$

$$p^*_{min} = \max\{0, \overline{\widehat{p}} - \frac{\widetilde{p_{sd}}}{2}\} \quad (13)$$

$$p^*_{max} = \min\{\overline{\widehat{p}} + \frac{\widetilde{p_{sd}}}{2}, 1\} \quad (14)$$

One could have used a wider interval (e.g., equal to two or three standard deviations). However, we want to keep the uncertainty lower, by possibly trading some level of accuracy. Eqs. (12)-(14) define the FR interval I with respect to Jack from the perspective of Chloe, based on the direct evidence.

The use of a sliding window results in a subset of the samples being common across windows. Thus, the estimates $\widehat{p_j}$ are *biased* by the samples that are common across the windows. To obtain unbiased estimates, one would need to use non-overlapping windows. However, in such a case, the updates are performed less frequently and one runs into the problem of the evidence becoming *stale*. Our evaluations show that the sliding window works well in practice.

C. Combining indirect evidence

Chloe can update her direct view of Jack's FR via feedback from other entities (say Tony) in the network. These entities are the *gossipers*. Using the DSTE, Chloe can combine the obtained feedback to derive an **aggregated** FR value for Jack. The use of indirect evidence is vital; Chloe may have conducted only a few or no transactions with Jack. In such cases, indirect evidence helps her assess Jack's FR. Our trust propagation technique helps address the challenge that indirect evidence may be unreliable.

As mentioned earlier, DSTE can be used to infer the likelihood of a *system* of being in a particular state based on a set of possibly contradicting pieces of evidence. Here, there are two states in our "virtual" system; θ_1, Jack is reliable, and θ_2, he is unreliable. Without loss of generality, we assume that we have two independent sources of evidence; the interval I_d derived from Chloe's direct observations on Jack, and the interval, I_g, that Chloe obtains from a gossiper, Tony. More than two sources of evidence can be aggregated sequentially in pairs.

Directly performing the aggregation on the intervals I_* is hard. Thus, we perform two separate aggregations; one on the lower bounds of the intervals, and one on the upper bounds. Each aggregation will yield an interval in which the real value lies. Thus, Chloe will end up with an interval for the lower bound for Jack's FR, and another interval for the upper bound. However, as we show later, for our system these intervals are reduced to a single value.

A sketch of the aggregation process: Assume that $I_d = [a_1, b_1]$, $I_g = [a_2, b_2]$ and consider the aggregation on the lower bound of the FR interval. First, we define the bpa functions (recall section II), m, associated with each source of evidence. The powerset $2^\Theta = \{\emptyset, \{\theta_1\}, \{\theta_2\}, \{\theta_1, \theta_2\}\}$. Note that the elements of the powerset are the different hypotheses H, as introduced in Section II. For the bpa of the direct observations we have:

$$m_d^{min}(\emptyset) = 0 \quad (15) \qquad m_d^{min}(\{\theta_1\}) = a_1 \quad (16)$$

$$m_d^{min}(\{\theta_2\}) = 1 - a_1 \quad (17) \qquad m_d^{min}(\{\theta_1, \theta_2\}) = 0 \quad (18)$$

For the bpa of the gossip-based/indirect evidence we have:

$$m_g^{min}(\emptyset) = 0 \quad (19) \qquad m_g^{min}(\{\theta_1\}) = a_2 \quad (20)$$

$$m_g^{min}(\{\theta_2\}) = 1 - a_2 \quad (21) \qquad m_g^{min}(\{\theta_1, \theta_2\}) = 0 \quad (22)$$

We assume that the available pieces of evidence lead to a ***probabilistic binary decision*** (i.e., a node is functionally reliable or not). In other words, there is no uncertainty or ambiguity with regard to the state, i.e., the probability that $\theta_1 \cap \theta_2$ is 0. This results in Eqs. (18) and (22) and these are key for proving Lemma 1. As discussed in Section II, the bpa m_d^{min} expresses the measure of belief committed on each hypothesis from the direct evidence with regards to the minimum FR value of a node. Since the direct FR of Jack as per Chloe is given by the interval I_d, the belief committed on the hypothesis that a node is responsive (hypothesis θ_1), with respect to its minimum FR, is $m_d^{min} = \inf\{I_d\} = a_1$ (Eq. (16)). Given, that θ_1 and θ_2 are complementary we get Eq. (17). The above steps apply to m_g^{min} as well. (Below, we write $m(*)$ instead of $m(\{*\})$.)

Lemma 1: With the bpas defined in Eqs. (15)–(22), the aggregated intervals are reduced to a single value.

Proof: Using the rule of combination (see Eq. 6), we first compute the aggregated bpa, m_{agg}^{min}. It is easy to see that: $m_{agg}^{min}(\emptyset) = m_{agg}^{min}(\theta_1, \theta_2) = 0$. In addition we have:

$$m_{agg}^{min}(\theta_i) = \frac{2 \cdot m_d^{min}(\theta_i) \cdot m_g^{min}(\theta_i)}{K}, \ for \ i = 1, 2 \qquad (23)$$

where $K = 2 \cdot m_d^{min}(\theta_1) \cdot m_g^{min}(\theta_1) + 2 \cdot m_d^{min}(\theta_2) \cdot m_g^{min}(\theta_2)$.

Using the above aggregated bpa and the definitions of belief and plausibility (Eq. (5)), we have:

$$Bel^{min}(\theta_1) = Pl^{min}(\theta_1) = m_{agg}^{min}(\theta_1) = \\ = \frac{a_1 \cdot a_2}{a_1 \cdot a_2 + (1 - a_1) \cdot (1 - a_2)} \qquad (24)$$

This concludes the proof for the lower bound of the FR. Similar steps can be followed for the upper bound. ■

Thus, after the aggregation process, Chloe's updated estimate of Jack's FR is the interval:

$$T_{\{Chloe,Jack\}} = \Big[\frac{a_1 \cdot a_2}{a_1 \cdot a_2 + (1 - a_1) \cdot (1 - a_2)}, \\ \frac{b_1 \cdot b_2}{b_1 \cdot b_2 + (1 - b_1) \cdot (1 - b_2)} \Big] \qquad (25)$$

D. Our framework in different contexts

For ease of discussion, we have so far assumed a scenario where Chloe estimates the e2e FR of Jack with a simple transaction type without any QoS requirements. Our framework, however, is independent of the context as long as the observations are Boolean outcomes. To illustrate this, we consider three contexts next.

e2e FR: The first scenario is a case where Jack is expected to perform a function to satisfy an end-to-end requirement. In the simplest case (as considered in our narrative), the desired function is to just respond to a query. The only metric of interest is the delivery of transaction packets e2e; in this case, the effect of the wireless medium that is of interest is simply the PDR. However, one can easily envision applications that have other requirements. For example, there may be a requirement that the response is received within a prespecified delay. In such a case, one will have to hypothesize about the *timeliness* of Jack's response. The constraints will be the delays imposed by retransmissions and queuing on the wireless medium and the likelihood of the packets being delayed by unresponsive relays. As a second example, if a query requests a video clip, one can impose a requirement on the quality of the clip. Then one needs to compute the likelihood that the degradation was caused by channel induced failures or packet drops by relays as opposed to Jack sending a poor quality clip. The examples here are not exhaustive; however, if one can compute the likelihood of a transaction not meeting the requirements due to wireless effects or packet drops/delays by relays, one can apply our framework to provide an assessment of Jack's FR.

Forwarding FR: One may envision the forwarding FR (alluded to in our earlier discussion) to be independent of the e2e FR. Due to intermittent link qualities, interference, poor battery state, or because of compromise Jack may not forward traffic as expected. Jack's neighbors can monitor (perhaps promiscuously) his activities [1] with respect to forwarding packets and obtain direct evidence $\vec{e}$, which will lead to their assessments of his forwarding FR. An important difference with the e2e FR is that Chloe may not have a direct link to Jack and thus, no opportunity to observe Jack's forwarding behavior. Thus, she has no direct evidence on Jack. Our framework can still be applied by combining indirect evidence from Jack's neighbors. We evaluate our framework in terms of its effectiveness in assessing the forwarding FR of nodes in Section V.

Gossiping FR: In our framework, Chloe updates the FR of Jack using indirect evidence from Tony. We have thus far implicitly assumed that all nodes do provide such indirect evidence. For a variety of reasons discussed earlier (e.e., poor battery, loss of connectivity) Tony may not however, provide timely or accurate information with regards to Jack. Furthermore, the accuracy of the evidence from Tony may depend on factors such as his distance from Jack and Chloe (the greater the distance the less accurate the evidence). The probability of Tony providing timely/accurate evidence refers to the gossiping FR of Tony as per Chloe. Modeling this leads to additional complexities (finding the likelihood that timely and accurate evidence is received from Tony). In our evaluations, we assume that nodes are reliable with regards to the gossiping function and we evaluate our framework on e2e and forwarding FR. Determining the gossiping FR is cumbersome in terms of obtaining the required evidence. Nevertheless, once the evidence is in place, it is straightforward to infer the gossiping FR. We will consider this in the future.

IV. Lightweight evidence propagation

Implicit in our FR assessment scheme was the use of a propagation protocol for distributing indirect evidence. The indirect evidence that Chloe obtained from Tony with respect to Jack was simply Tony's *direct* FR assessment of Jack. A simple approach for evidence propagation is a flooding scheme; Tony propagates his assessed direct FR values with respect to all other nodes, to everyone. With this, every node (say, Chloe again) will have a global view of the direct relationships, and thus, she can use DSTE's orthogonal product to compute the aggregated FR on each of her peer network users. While the scheme provides simplicity and accuracy the associated overhead is large; the number of messages that need to be transmitted is $O(N^2)$, where N is the number of nodes in the network.

We therefore design a new mechanism for propagating the assessed FR values (evidence) with the objective of reducing the communication overhead incurred. If each node only communicates with its direct physical neighbors, the overhead can be reduced drastically. As will be evident, the number of messages that need to be transmitted with such an approach is linear with respect to the number of users in the network, i.e., $O(N)$.

Double Counting of Evidence: The use of local broadcasts results in a challenge that we have to address. Each node propagates evidence only to its physical neighbors. These neighbors then *fuse* or *combine* this evidence and propagate it to other nodes. Let us assume that Chloe gets observations with regards to Jack from two of her neighbors, say, Jill and Jane. However, this evidence may have originated at a single node, say, Tony. Thus, Chloe will have to ensure that she does not "double" count this evidence when computing an FR value for Jack (since the originator is Tony for both pieces of evidence). Below we present three operators that are essential for our scheme for filtering such duplicate evidence.

Indirect relaying of evidence (as will be the case here) could also result in a decrease in the accuracy of the computation of FR at each node. Our experimental evaluations presented in Section V demonstrate, however, that this impact is low.

A. Path FR operators

1) Fusion $\odot$: For simplicity, let us consider the 3-hop physical topology of Fig. 1. The extension to an n-hop case is trivial. Node C, wants to update X's assessed FR, using information gathered along the shown path. The steps followed are:

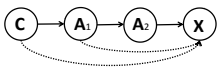


Fig. 1. Fusion on path P: $P(X) = C(X) \odot A_1(X) \odot A_2(X)$.

Fig. 2. Select operator on two dependent paths.

Step 1: A_1 updates X's FR value through DSTE's rule of combination. The two sources of evidence combined are: (i) X's direct FR as per A_1 (direct evidence) and, (ii) X's direct FR as per A_2.

Step 2: C updates X's FR in a similar manner using: (i) its own direct FR for X and, (ii) A_1's updated FR for X (as computed in Step 1).

Note that there are two crucial features of the fusion operator. At every step, (i) the sources of the pieces of evidence that are being combined are independent. Thus, the only requirement for using the DSTE's orthogonal product is fulfilled. (ii) Only new information is being added, which guarantees that there is *no double counting of evidence*.

2) Path Aggregation $\otimes$: In the majority of the cases, there are multiple physical paths from source node C to node X. In general, each of these paths will result in a different assessed FR interval for X. C should be able to aggregate the different FR intervals with respect to X derived on the basis of the different paths. For this we will use the path aggregation operator $\otimes$.

Let us assume that (a) we have k **independent** paths (i.e., they do not share any intermediate common nodes) from node C to node X and (b) using path P_i, C derives the interval $[a_i, b_i]$ to reflect the FR of node X. Then, the path aggregated FR interval for X, from C's perspective, is computed as

$$P(X) = \otimes_{i=1}^{k} P_i(X) = [\min_{j\in\{1,2,\ldots,k\}}\{a_j\}, \max_{j\in\{1,2,\ldots,k\}}\{b_j\}] \quad (26)$$

The path aggregation operator computes the global min (max) of the individually estimated lower (upper) bounds from each considered independent path. Thus, the computed interval is likely to be large and hence the uncertainty on the computed FR value will be larger than what is computed with the basic approach where all information is strictly accounted for. One could more carefully try to combine evidence from the multiple paths but the processing complexity will be higher. We choose lower complexity in lieu of lower uncertainty with the objective of keeping the process lightweight. Our experimental evaluations show that this results in a small increase in inaccuracy.

In topologies similar to the one in Fig. 2, if one were to apply the fusion operation on the two paths leading to node X and then aggregate the FR intervals, double counting of evidence will occur. This is because the paths are not independent (they share common intermediate nodes); the evidence from nodes A_5 and A_6 will be counted twice. We adopt a variant of the select operator $\langle S\rangle$ [14] that chooses the stronger of two paths (trivially extended to multiple paths).

Select makes use of the fusion operator, to compute the FR interval on the furthest common node Y (A_6 in our example) along the two different paths ($P_i = C - A_1 - A_2 - A_5 - A_6$ and $P_j = C - A_3 - A_4 - A_5 - A_6$ as in Fig. 2). Let us assume that these intervals are: $P_i(Y) = [a_i, b_i]$ and $P_j(Y) = [a_j, b_j]$. Then, the select operator picks path P as follows:

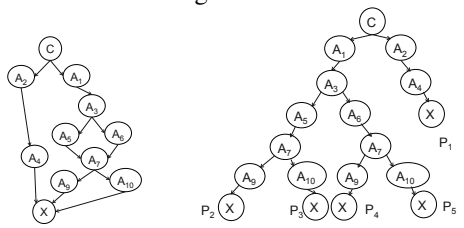
$$P = P_i(C,Y)\langle S\rangle P_j(C,Y) = \begin{cases} P_i, & \text{if } a_i > a_j, \\ P_j, & a_i < a_j. \end{cases} \quad (27)$$

The select operator is *optimistic*, in the sense that it chooses the path that leads to the *maxmin* FR value on the common intermediate node. Note that, if $a_i = a_j$, P is randomly selected.

B. Tree construction and evidence propagation

Next, we present our tree based, lightweight evidence propagation protocol. The goal is to identify a set of intermediate nodes that will provide the indirect evidence in order to update the assessed FR value on a specific network entity. By only considering a subset of nodes in the network to provide indirect evidence, we may reduce the accuracy of the assessments; however, it helps overcome problems arising from the duplication or double counting of evidence while reducing the overhead incurred in FR propagation. Our scheme is based on the physical network topology. In brief, all the independent paths toward the target node are identified and the FR for a node is updated only via these paths, utilizing the operators presented above.

Toy example: Let us consider the physical topology presented in Fig. 3(a). Node C wants to update X's FR value. To achieve this, C needs to assimilate the knowledge obtained from the nodes along the path toward X. However, blindly aggregating the FR values reported by the intermediate nodes can lead to double counting of evidence.



(a) Physical topo (b) FR propagation tree.

Fig. 3. Lightweight Propagation.

In order to construct the tree, we first identify all the different paths that lead to node X. In the scenario under consideration we have five paths. Among these, P_1 is the only path that does not share a common node with any other path i.e., P_1 is *independent* of all the other paths. Using the fusion operator we can estimate X's FR through P_1 to be $P_1(X)$, where $P_1(X) = C(X) \odot A_2(X) \odot A_4(X)$.

The remaining four paths are not independent and therefore we need to eliminate the dependencies. Starting bottom up (i.e., from the target node X to the source C), P_2 and P_4 have A_9 as a common node, while P_3 and P_5 both include A_{10}. For each of the above pairs of paths we apply the select operator. In particular, we have $P_{2,4} = P_2(C, A_9)\langle S\rangle P_4(C, A_9)$ and $P_{3,5} = P_3(C, A_{10})\langle S\rangle P_5(C, A_{10})$. Now we have reduced the number of paths from four to two. These two paths however, are still *dependent* (node A_7 is common to both of them; A_1 and A_3 are also common on both paths, but A_7 is the deepest match). Thus, we apply the select operator again on the two resulting paths and we have: $P_{2,3,4,5} = P_{2,4}(C, A_7)\langle S\rangle P_{3,5}(C, A_7)$.

The final step is to combine the evidence obtained from the two independent paths, P_1 and $P_{2,3,4,5}$ to form an FR value for X, using the path aggregation operator. In other words, C computes $P(X) = P_1(X) \otimes P_{2,3,4,5}(X)$.

Generalizing our algorithm: The first step toward updating the assessed FR on a network entity is to construct a logical tree that gathers all the possible physical paths from the source to the target node X. The root of the tree is the source node, C, while a leaf of the tree corresponds to the target X. The 1^{st} level of the tree (*children* of root C) includes the physical neighbors of root C. Recursively, the children of the nodes of the i^{th} level (forming the $(i+1)^{st}$ level) include the neighbors of the nodes residing at this level. We continue until we cannot further update the tree paths and we keep only the paths that end at node X. The procedure requires nodes to indicate the *chain of evidence* in their local broadcasts; in other words, they announce the path via which the evidence was propagated. This allows a node (say node C) to determine the topology. Clearly, in the worst case, a piece of evidence has $O(N)$ associated node identities in the chain. For moderate sized networks, we expect that this will not result in much overhead (assuming no more than 32 bits if IP addresses are used as identifiers). Using hashes of addresses could further decrease this overhead.

Next, we parse the tree and perform the following 3 steps:

1) Step 1 - Identify independent paths: If all nodes A_i, $i \in \{1, 2, ..., n\}$, belonging to a path P_i do not belong to any other path P_j, then P_i is independent of any other path. Thus, the FR of the leaf X along P_i is estimated using the fusion operator, $P_i(X) = C(X) \odot_{i=1}^{n} A_i$.

2) Step 2 - Prune path dependencies: If two paths P_m and P_n, share common nodes, we use the select operator to eliminate the dependencies. As discussed, we first identify the *deepest* matching node (e.g., node f). Note that the common nodes can appear at different tree levels across the different paths; however, they will appear in the same order. This is easy to verify since the tree is based on the physical network topology.

After identifying f we retain path P, where $P = P_m(C, f)\langle S \rangle P_n(C, f)$. This process continues until we remove all dependencies. At the end of this step all paths that are still under consideration, are independent.

3) Step 3 - Aggregate path FR: Using the path aggregation operator, we aggregate the FR values along all the z (independent) paths identified at the end of Step 2. In particular, the FR of node X is updated to be: $P(X) = \otimes_{i=1}^{z} P_i(X)$.

Message complexity: The tree-based algorithm reduces the communication overhead compared to the flooding approach. It can be shown formally that the message complexity of our scheme is $O(N)$, where N is the number of nodes in the network. Similarly, the time-complexity of tree-construction is also O(N). We omit the proofs due to space limitations.

Discussion: With lightweight propagation, due to either link failures or poor forwarding FR, indirect evidence from some of the paths may be lost. However, we find in our experiments that this does not significantly affect the accuracy in FR assessment since in most cases, evidence is collected along the most reliable paths. Finally note here that if a node on a path does not have any evidence relative to a node (say, A_7 does not have any evidence relating to X), it may simply forward the evidence from its predecessor or use a value of 0.5 relating to the FR of X. The latter simply indicates that from A_7's perspective, the events that X is functionally reliable *or* unreliable are equally likely. We use the latter approach in our experimental studies.

Finally, we point out that the mechanics of the lightweight evidence propagation is not new. While the mechanics of the flood based approach is similar to link state routing update propagation, that of the lightweight approach is similar to the propagation of routing updates in distance vector routing [15]. The novelty of the approach is in filtering duplicate evidence.

V. Implementation and evaluation of our system

We now present the implementation and experimental evaluations of our framework. We implement our scheme with both (a) flooding based direct evidence propagation and (b) our lightweight evidence propagation.

Protocol implementation and experimental setup: Our implementation is on our 42-node wireless testbed, which consists of both indoor and outdoor links as detailed elsewhere [16].

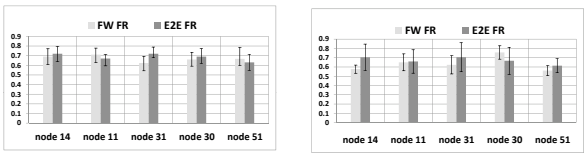
Our measurements span many wireless links and routes of different lengths, and packet delivery ratios (PDR). We experiment with the 802.11g mode. Our framework is implemented using the Click toolkit [17]. By default, we use ETX routing [12] and the ETX metric is used to estimate link qualities.

Ground truth: We preconfigure each node's forwarding FR and the likelihood of its responding to e2e queries (e2e FR); this defines the *ground truth* in terms of the actual long term behavior of the node. Each node also has an *initial FR value* for both forwarding and e2e queries with regards to every other node. We set this to be 0.5 with an uncertainty of 0 for both operations. With time, we expect that with our framework, FR values evolve from this initial state, based on both direct and indirect evidence; the FR values at the end of an observation period is the *assessed FR* at the end of the period. Our objective is to see how the assessed FR compares with the ground truth.

Functions examined: Each node (Chloe) randomly picks a target (Jack) and sends *ICMP* queries; these queries form the basis for the direct e2e observations. To decide on the success/failure of an e2e transaction, we send 10 ICMP_ECHO_REQUEST messages and we expect $x\%$ these to successfully result in ICMP_ECHO_REPLY messages. We disable link layer retransmissions and we pick $x\%$ to be the minimum delivery probability among all the links of the route. To determine the success/failure of *forwarding operations*, we configure the sender to be in the promiscuous mode to overhear forwarded packets. If the PDR of the link between the forwarder and the sender is $y\%$, we expect the sender to overhear at least $y\%$ of the ICMP_ECHO_REQUEST messages ***delivered*** to the forwarder. Each experiment runs for 3000 seconds in which each node makes on average 10–15 observations for every other node.

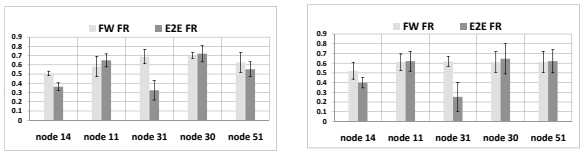
Protocol Details: With flood based propagation, the direct evidence of a node is broadcast every 10 seconds (this forms indirect evidence for other nodes). Each node appends any new information and re-broadcasts a received broadcast. With the lightweight propagation scheme, each node *locally* broadcasts its *aggregated* FR estimates for other nodes, every 10 seconds. These broadcasts include the chain of evidence, which allows each node receiving them to locally recreate the network topology. When a node receives such local information, she updates its aggregated FR (using DSTE) for each of its peers (indirect evidence aggregation) and re-broadcasts the new information. For both schemes, direct evidence is computed using MLE with a sliding window of eight observations.

Accuracy in reliable and unreliable settings: First, we examine the accuracy of the estimation process when using (i) the flood-based evidence propagation and (ii) our lightweight protocol. Initially, we preconfigure all nodes to be responsive (forwarding and e2e FR values are 1 and their uncertainty is 0); this represents the ground truth in terms of FR. Fig. 4 shows representative assessed FR values and their uncertainty for 5



(a) Flood-based propagation (b) Lightweight scheme

Fig. 4. FR assessment under benign settings (Preconfigured FR is 1 for all nodes).



(a) Flood-based propagation (b) Lightweight scheme

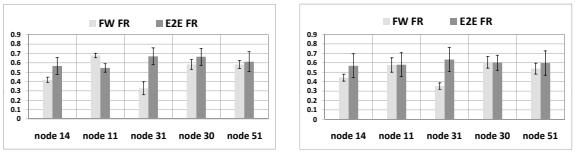
Fig. 6. The assessed e2e FR for unreliable nodes (Nodes 14 and 31 have a preconfigured e2e FR of '0'. Other nodes are responsive).

nodes at the end of our experiment (later we present statistics from a large set of trials). To annotate, the bars corresponding to node 14 indicate that, the mean FR (computed over all nodes in the network) on this node is 0.7, the maximum of the mean FR values from among all these FR values is 0.8 and the minimum is about 0.6. The uncertainty on the estimated values for each individual assessment is typically < 10 % of the mean FR value and these are not plotted to ensure clarity. These results suggest that with both schemes the average assessed FR values are sufficiently *close* to the ground truth. Since there are uncertainties that influence the computed FR (wireless effects, varying FR reports from gossipers), the average value almost never converges to the ground truth within the experiment duration. We see that the accuracy is typically lower with the lightweight protocol since, with the latter fewer observations are combined to form indirect evidence in the gossiping phase. Although it is possible that with an increased number of samples, the accuracy can sometimes decrease (rather than increase), we do not observe this to be the case here.

Our results with the preconfigured forwarding FR values of nodes 14 and 31 set to 0 with an uncertainty of 0 (ground truth) are in Fig. 5. We observe that the assessed "average" e2e FR is lower as compared with that in the "reliable" settings scenario. For example, the inferred average e2e FR of node 11 is approximately 20% lower for both schemes. Many transactions fail due to the forwarding unreliability. Unfortunately, the estimation engines (slightly) penalize the end node as well, due to the uncertainty in ascertaining the reason that caused the failure.

Finally, we preconfigure the e2e FR of nodes 14 and 31 to 0 (i.e., they do not respond to ECHO_REQUESTS) and restore their forwarding responsiveness to '1' (ground truth). The results with our framework are presented in Fig. 6. We see that the average e2e FR of these nodes is significantly lower as compared to the other nodes (e.g., node 31 exhibits an approximately 60% lower e2e FR as compared with node 30 for both schemes). This value still is about 0.2, due to the small number of transactions. It is also influenced by the initial FR value of 0.5. These factors result in increased uncertainty in the assessment process, which reduces accuracy.

FR evolution with different initial values: As alluded to above, the initial FR value that bootstraps the assessment process can affect the estimated FR value since this is used in the aggregation. To examine the impact of this parameter,



(a) Flood-based propagation (b) Lightweight scheme

Fig. 5. Non-responsive relays can affect the e2e FR assessment (Nodes 14 and 31 have a preconfigured forwarding FR of 0. All other FR values are set to 1).

we experiment with different initial FR values. In particular, we examine the average FR for node 31 (over the observation period) with 3 different initial FR values, 0.2, 0.5 and 1 (all with uncertainty 0). Fig. 7 presents our results, for 3 different scenarios; node 31 is (i) a responsive node, (ii) an unreliable relay and (iii) an unreliable node with respect to e2e queries. The values depicted are the estimated average FR values (the average computed on the perception of the mean responsiveness of 31 by all other nodes) after 3000 seconds. It is evident, that when the initial value is close to the actual value, the estimation within the considered time is much more accurate. For instance, when node 31 does not respond to e2e queries (e2e FR is 0), when the initial FR is 0.2, the assessed value is approximately 0.18, while with an initial FR of 0.5 (respectively, 1) the estimated values are larger, 0.31 (respectively, 0.34). With an increase in the number of observations the effect of the initial FR values decreases and the assessed values come closer to the actual preconfigured FR values (shown next). However, *strict* convergence is not achieved since there is always some degree of uncertainty with regards to whether or not other factors (e.g., wireless effects) contributed to transaction/operation failures.

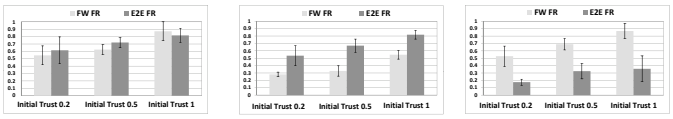
Accuracy vs number of observations: The number of e2e transactions between users affects the accuracy of estimation of both the forwarding FR as well as the e2e FR. Considering the same set up as above, and preconfiguring nodes 14 and 31 to be non-responsive relays we run our experiments for a larger period (≈ 10000 seconds), enough to perform up to 40 transactions pairwise. Fig. 8 presents the estimated FR values with 20 and 40 pairwise transactions. As one might expect, with more transactions (and thus, more observations), the assessed FR values are closer to the *actual* preconfigured ones for both responsive and unreliable nodes.

Overhead comparison of the flood based and lightweight propagation protocols: We compare the two propagation protocols in terms of the induced overhead; we also look at the accuracy achieved (in terms of the *distance* between the assessed and the preconfigured FR values i.e., the ground truth). We see from Table I that as expected flooding results in smaller uncertainty. However, the mean distances from the ground truth are very similar with both schemes. It is also evident that the lightweight propagation results in about a 37% decrease in the induced overhead. We believe that this is a significant reduction, at the expense of a slightly higher inaccuracy and uncertainty.

	Traffic Load(Bytes)	Distance	Uncertainty
Flooding	12387191	0.0945	0.121
Lightweight	7751196	0.1168	0.222

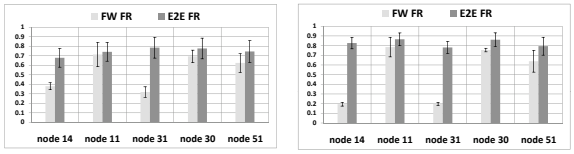
TABLE I
COMPARING FLOOD BASED AND LIGHTWEIGHT PROPAGATION.

Transactions between routing and FR establishment/propagation protocols: Next we want to study the impact of different routing protocols on the evolution of the FR values using our lightweight protocol. As observed in our first set of experiments, the presence of non-responsive



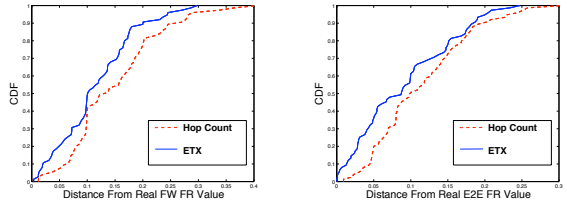
(a) Benign node (b) Non-responsive relay (c) e2e non-responsiveness

Fig. 7. Higher accuracy is achieved when the initial FR is *closer* to the preconfigured FR.



(a) 20 transactions (b) 40 transactions

Fig. 8. More observations lead to higher accuracy (Nodes 14 and 31 are non-responsive relays. Other nodes are responsive).



(a) Forwarding FR (b) e2e FR

Fig. 9. CDF of the distance between the assessed and real FR.

relays can affect the establishment of the e2e FR values when ETX routing is used. The routes do not account for the forwarding FR of nodes and hence, the presence of *bad* relays on a path can cause transactions to fail. This consequently results in a reduction in the accuracy of the assessed e2e FR.

We perform a large set of experiments where each node is preconfigured with randomly chosen FR tuples (for forwarding and e2e FR). We ensure that these FR values are evenly spread across $[0, 1]$. We use 2 different routing metrics to find routes, minimum hop count and ETX. Running 10 repetitions of our experiments for 5000 seconds each, we obtain the results in Fig. 9. These figures depict the CDF of the distance between the preconfigured (*real*) and the assessed average FR values for the nodes. We observe that in all scenarios, minimum hop distance performs the worst in terms of accuracy due to long unreliable links that contribute to high uncertainty. A routing metric that not only accounts for the link qualities (e.g., ETX), but also for the forwarding responsiveness of the relays can further improve the assessment accuracy. Designing such a metric is beyond the scope of our study and is left for future work.

On the hardness of convergence: We observe from Fig. 9 that in the best case, almost 40% of our assessments differ by at least 0.1 from the ground truth in terms of both the forwarding and e2e FR. It is really hard, if not impossible, to achieve *strict convergence* to the real FR values. There are several reasons that contribute to this hardness. As seen earlier, the forwarding FR affects the e2e FR (Fig. 5). Failures due to forwarding attackers, will influence the assessed e2e FR. As our experiments indicate (omitted due to space constraints) the same happens when the transactions fail due to wireless induced effects. In addition, as discussed earlier, the initial value affects the assessment as well. Finally, gossiping adds uncertainty and decreases the accuracy of the assessed FR. So even for a completely (e2e) responsive node the assessment engine cannot converge to the value of 1. We believe that this level of accuracy however, is sufficient in most cases when nodes make coarse grained assessments to hypothesize about the reliability of peers.

VI. Conclusions

We design a framework for collaborative FR assessment in wireless networks. Unlike in prior work, we account for wireless induced factors and the reliability of intermediary relays. The framework accounts for both direct transactions between nodes and indirect feedback obtained from gossipers about other nodes in the network. It consists of a lightweight evidence propagation scheme that carefully filters out duplicate evidence. Our evaluations on an indoor/outdoor wireless testbed show that each node is able to estimate the FR values for other nodes with a sufficiently high accuracy.

Acknowledgements

Research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-09-2-0053 and partially supported by the University of Pittsburgh Central Research Development Fund. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

References

[1] S. Marti, T. Giuli, K. Lai, and M. Baker. Mitigating routing misbehavior in mobile ad hoc networks. In *MobiCOM*, 2000.

[2] S. Buchegger and J.-Y. Le Boudec. Nodes Bearing Grudges: Towards Routing Security, Fairness and Robustness in Mobile Ad Hoc Networks. In *Proc. Euromicro Wkshp. Parallel, Distributed and Network-based Processing*, 2002.

[3] S. Buchegger and J.-Y. Le Boudec. Performance Analysis of CONFIDANT protocol: Cooperation of nodes - fairness and dynamic ad-hoc networks. In *ACM MobiHOC*, 2002.

[4] P. Michiardi and R. Molva. CORE: A Collaborative Reputation Mechanism to enforce node cooperation in Mobile Ad Hoc Networks. In *European Wireless Conference*, 2002.

[5] Matthew J. Probst and Sneha Kumar Kasera. Statistical Trust Establishment in Wireless Sensor Networks. In *Proc. 13th International Conference on Parallel and Distributed Systems*, 2007.

[6] Pedro B. Velloso, Rafael P. Laufer, Daniel de O. Cunha, Otto Carlos M. B. Duarte, and Guy Pujolle. Trust management in mobile ad hoc networks using a scalable maturity-based model. In *IEEE Trans. Netw. Service Manage., Vol. 7, No.3*, 2010.

[7] Kannan Govindan and Prasant Mohapatra. Trust Computations and Trust Dynamics in Mobile Adhoc Networks: A Survey. In *Communications Surveys Tutorials, IEEE, Vol. PP, Issue 99*, 2011.

[8] L. Buttyan and J.P. Hubaux. Enforcing Service Availability in Mobile Ad Hoc WANs. In *ACM MobiHOC*, 2000.

[9] S. Zhong, J. Chen, and Y. R. Yang. Sprite: A Simple, Cheat-Proof, Credit-Based System for Mobile Ad-Hoc Networks. In *IEEE INFOCOM*, 2003.

[10] G. Shafer. *A Mathematical Theory of Evidence*. Princeton Univ. Press, Princeton, NJ, 1976.

[11] A. P. Dempster. A generalization of Bayesian inference. In *Journal of the Royal Statistical Society, Series B, Vol. 30, pp. 205247*, 1968.

[12] D. S. J. De Couto, D. Aguayo, J. Bicket, and R. Morris. A High Throughput Path Metric for MultiHop Wireless Routing. In *ACM MOBICOM*, 2003.

[13] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, ISBN 0-13-345711-7.

[14] C.-W. Hang, Y. Wang, and M. P. Singh. Operators for propagating trust and their evaluation in social networks. In *Proc. AAMAS*, 2009.

[15] C. Perkins. *Ad hoc Networking*. Addison Wesley Professional Series, 2001.

[16] UCR Wireless Testbed. http://networks.cs.ucr.edu/testbed.

[17] Click Modular Router. http://read.cs.ucla.edu/click/.