Behavioral Malware Detection in Delay Tolerant Networks

Wei Peng, *Student Member, IEEE,* {Feng Li, Xukai Zou}, *Member, IEEE,* and Jie Wu, *Fellow, IEEE*

Abstract—The delay-tolerant-network (DTN) model is becoming a viable communication alternative to the traditional infrastructural model for modern mobile consumer electronics equipped with short-range communication technologies such as Bluetooth, NFC, and Wi-Fi Direct. Proximity malware is a class of malware that exploits the opportunistic contacts and distributed nature of DTNs for propagation. Behavioral characterization of malware is an effective alternative to pattern matching in detecting malware, especially when dealing with polymorphic or obfuscated malware. In this paper, we first propose a general behavioral characterization of proximity malware which based on Naive Bayesian model, which has been successfully applied in non-DTN settings such as filtering email spams and detecting botnets. We identify two unique challenges for extending Bayesian malware detection to DTNs ("insufficient evidence vs. evidence collection risk" and "filtering false evidence sequentially and distributedly"), and propose a simple yet effective method, *look-ahead*, to address the challenges. Furthermore, we propose two extensions to *look-ahead*, *dogmatic filtering* and *adaptive look-ahead*, to address the challenge of "malicious nodes sharing false evidence". Real mobile network traces are used to verify the effectiveness of the proposed methods.

Index Terms—delay-tolerant networks; proximity malware; behavioral malware characterization; Bayesian filtering

1 Introduction

The popularity of mobile consumer electronics, like laptop computers, PDAs, and more recently and prominently, smartphones, revives the delay-tolerant-network (DTN) model as an alternative to the traditional infrastructure model. The widespread adoption of these devices, coupled with strong economic incentives, induces a class of malware that specifically targets DTNs. We call this class of malware *proximity malware*.

An early example of proximity malware is the Symbian-based *Cabir* worm, which propagated as a Symbian Software Installation Script (.sis) package through the Bluetooth link between two spatially proximate devices [1]. A later example is the iOS-based *Ikee* worm, which exploited the default SSH password on jailbroken [2] iPhones to propagate through IP-based Wi-Fi connections [3]. Previous researches [4] quantify the threat of proximity malware attack and demonstrate the possibility of launching such an attack, which is confirmed by recent reports on hijacking hotel Wi-Fi hotspots for drive-by malware attacks [5]. With the adoption of new short-range communication technologies such as NFC [6] and Wi-Fi Direct [7] that facilitate spontaneous bulk data transfer between spatially proximate mobile devices, the threat of proximity malware is becoming more realistic and relevant than ever.

Proximity malware based on the DTN model brings unique security challenges that are not present in the infrastructure model. In the infrastructure model, the cellular carrier centrally monitors networks for abnormalities; moreover, the resource scarcity of individual nodes limits the rate of malware propagation. For example, the installation package in *Cabir* and the SSH session in *Ikee*, which were used for malware propagation, cannot be detected by the cellular carrier. However, such central monitoring and resource limits are absent in the DTN model. Proximity malware exploits the opportunistic contacts and distributed nature of DTNs for propagation.

A prerequisite to defending against proximity malware is to detect it. In this paper, we consider a general behavioral characterization of proximity malware. Behavioral characterization, in terms of system call and program flow, has been previously proposed as an effective alternative to pattern matching for malware detection [8, 9]. In our model, malware-infected nodes' behaviors are observed by others during their multiple opportunistic encounters: Individual observations may be *imperfect*, but abnormal behaviors of infected nodes are *identifiable* in the long-run. For example, a single suspicious Bluetooth connection or SSH session request *during one encounter* does not confirm a *Cabir* or *Ikee* infection, but repetitive suspicious requests *spanning multiple encounters* is a strong indication for malware infection. The imperfection of a single, local observation was previously in the context of distributed IDS against slowly propagating worms [10].

Instead of assuming a sophisticated malware con-

- *W. Peng and Dr. X. Zou are with the Department of Computer and Information Science, Indiana University-Purdue University Indianapolis, Indianapolis, IN, 46202.*
- *Dr. F. Li is with the Department of Computer, Information, and Technology, Indiana University-Purdue University Indianapolis, Indianapolis, IN, 46202.*
- *Dr. J. Wu is with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA, 19122.*

tainment capability, such as patching or self-healing [11, 12], we consider a simple "cut-off" strategy: If a node i suspects another node j of being infected with the malware, i simply ceases to connect with j in the future to avoid being infected by j. Our focus is on how individual nodes shall make such cut-off decisions against potentially malware-infected nodes, based on direct and indirect observations.

A comparable example from everyday experience is fire emergency. An early indication, like dark smoke, prompts two choices. One is to report fire emergency immediately; the other is to collect further evidence to make a better informed decision later. The first choice bears the cost of a false alarm, while the second choice risks missing the early window to contain the fire.

In the context of DTNs, we face a similar dilemma when trying to detect proximity malware: Hypersensitivity leads to false positives, while hypo-sensitivity leads to false negatives. In this paper, we present a simple, yet effective solution, *look-ahead*, which naturally reflects individual nodes' intrinsic risk inclinations against malware infection, to balance between these two extremes. Essentially, we extend the Naive Bayesian model, which has been applied in filtering email spams [13, 14, 15], detecting botnets [16], and designing IDSs [10, 17], and address two DTN-specific, malware-related, problems.

1. *Insufficient evidence vs. evidence collection risk.* In DTNs, evidence (such as Bluetooth connection or SSH session requests) is collected only when nodes come into contact. But contacting malware-infected nodes carries the risk of being infected. Thus, nodes must make decisions (such as whether to cut off other nodes and, if yes, when) *online* based on potentially insufficient evidence.

2. *Filtering false evidence sequentially and distributedly.* Sharing evidence among opportunistic acquaintances helps alleviating the aforementioned insufficient evidence problem; however, false evidence shared by malicious nodes (the *liars*) may negate the benefits of sharing. In DTNs, nodes must decide whether to accept received evidence sequentially and distributedly.

Our contributions are summarized below.

1. We present a general behavioral characterization of proximity malware, which captures the functional but imperfect nature in detecting proximity malware (Section 2).

2. Under the behavioral malware characterization, and with a simple cut-off malware containment strategy, we formulate the malware detection process as a distributed decision problem. We analyze the risk associated with the decision, and design a simple, yet effective, strategy, *look-ahead*, which naturally reflects individual nodes' intrinsic risk inclinations against malware infection. Look-ahead extends the Naive Bayesian model, and addresses the DTN-specific, malware-related, "insufficient evidence vs. evidence collection risk" problem (Section 3.1).

3. We consider the benefits of sharing assessments among nodes, and address challenges derived from the DTN model: liars (i.e., bad-mouthing and false-praising malicious nodes) and defectors (i.e., good nodes that have turned rogue due to malware infections). We present two alternative techniques, *dogmatic filtering* and *adaptive look-ahead*, that naturally extend look-ahead to consolidate evidence provided by others, while containing the negative effect of false evidence. A nice property of the proposed evidence consolidation methods is that the results will not worsen even if liars are the majority in the neighborhood (Section 3.2). Real contact traces are used to verify the effectiveness of the methods (Section 4).

2 MODEL

Consider a DTN consisting of n nodes. The neighbors of a node are the nodes it has (opportunistic) contact opportunities with.

Proximity malware is a malicious program that disrupts the host node's normal function and has a chance of duplicating itself to other nodes during (opportunistic) contact opportunities between nodes in the DTN. When a duplication occurs, the other node is infected with the malware.

In our model, we assume that each node is capable of assessing the other party for suspicious actions after each encounter, resulting in a *binary* assessment. For example, a node can assess a Bluetooth connection or a SSH session for potential *Cabir* or *Ikee* infection. The watchdog components in previous works on malicious behavior detection in MANETs [18] and distributed reputation systems [19, 20] are other examples. A node is either evil or good, based on if it is or is not infected by the malware. The suspicious-action assessment is assumed to be an *imperfect* but *functional* indicator of malware infections: It may occasionally assess an evil node's actions as "non-suspicious" or a good node's actions as "suspicious", but most suspicious actions are correctly attributed to evil nodes. A previous work on distributed IDS presents an example for such imperfect but functional binary classifier on nodes' behaviors [10].

The functional assumption characterizes a malware-infected node by the assessments of its neighbors. If node i has N (pair-wise) encounters with its neighbors and s_N of them are assessed as suspicious by the neighbors, its *suspiciousness* S_i is defined as:

$$S_i = \lim_{N \to \infty} \frac{s_N}{N}. \tag{1}$$

By Equation (1), $S_i \in [0, 1]$. A number $L_e \in (0, 1)$ is chosen as the *line between good and evil*. L_e depends on the quality of a particular suspicious-action assessment and, if the assessment is a functional discriminant feature of the malware and the probabilistic distribution of the suspiciousness of both good and evil nodes are known, L_e can be chosen as the (Bayesian) decision boundary, which minimizes classification errors [21]. Node i is good if $S_i \leq L_e$, or evil if $S_i > L_e$: We draw a fine line between good and evil, and judge a node by its deeds.

Instead of assuming a sophisticated malware coping mechanism, such as patching or self-healing, we consider a simple and widely applicable malware containment strategy: Based on past assessments, a node i decides whether to refuse future connections ("cut off") with a neighbor j.

3 Design

In the following discussion, we investigate the decision process of a node i, which has k neighbors $\{n_1, n_2, \dots, n_k\}$, against a neighbor j; with no loss of generality, let j be n_1.

3.1 Household Watch

Consider the case in which i bases the cut-off decision against j *only* on i's own assessments on j. Since only direct assessments are involved, we call this model *household watch* (the naming will become more evident by the beginning of Section 3.2).

Let $\mathcal{A} = (a_1, a_2, \dots, a_A)$ be the assessment sequence (a_i is either 0 for "non-suspicious" or 1 for "suspicious") in chronological order, i.e., a_1 is the oldest assessment, and a_A is the newest one.

Bayes' theorem tells us:

$$P(S_j|\mathcal{A}) \propto P(\mathcal{A}|S_j) \times P(S_j). \tag{2}$$

$P(S_j)$ encodes our prior belief on j's suspiciousness S_j; $P(\mathcal{A}|S_j)$ is the likelihood of observing the assessment sequence $\mathcal{A}$ given S_j; $P(S_j|\mathcal{A})$ is the posterior probability, representing the plausibility of j having a suspiciousness of S_j given the observed assessment sequence $\mathcal{A}$. Since the evidence $P(\mathcal{A})$ does not involve S_j and serves as a normalization factor in the computation, we omit it and write the quantitative relationship in the less cluttered proportional form[1].

By Sections 1.1 and 1.2 of the supplementary document, we have:

$$P(S_j|\mathcal{A}) \propto S_j^{s_\mathcal{A}} (1 - S_j)^{|\mathcal{A}| - s_\mathcal{A}} \tag{3}$$

and:

$$\underset{S_j \in [0,1], \mathcal{A} \neq \emptyset}{\arg\max} \; P(S_j|\mathcal{A}) = \frac{s_\mathcal{A}}{|\mathcal{A}|}\,, \tag{4}$$

in which $s_\mathcal{A}$ is the number of suspicious assessments in $\mathcal{A}$.

Figure 1 shows the normalized posterior distributions $P(S_j|\mathcal{A})$ for assessment samples with different sizes, given by Equation 3. In each case, the ratio between suspicious and non-suspicious assessments is the same, i.e., 1:3; by Equation 4, $S_j = \frac{1}{1+3} = 0.25$ is the maximizer of $P(S_j|\mathcal{A})$, which is clearly shown in Figure 1. The distribution becomes sharper with a larger sample, which accords to the intuition of the increasing certainty on the suspiciousness S_j.

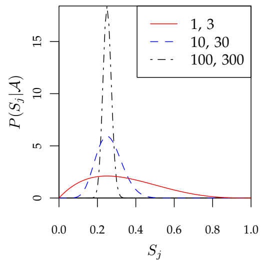


Fig. 1: The normalized posterior distribution $P(S_j|\mathcal{A})$ for assessment samples with different sizes. The two numbers for each line in the legend show the number of suspicious and non-suspicious assessments, respectively. In each case, the ratio between suspicious and non-suspicious assessments is $1:3$. All distributions have a maximal value at $S_j = \frac{1}{1+3} = 0.25$. However, the distribution becomes shaper with a larger sample, which corresponds to a sense of increasing certainty regarding the suspiciousness S_j.

The uncertainty over j's suspiciousness S_j (and, hence, the risk of losing a good neighbor) holds i back from cutting j off immediately, based on insufficient evidence. In the following discussion, we consider two alternative approaches, *distribution* and *maximizer*, to handle the insufficient-evidence problem, based on Equations (3) and (4), respectively.

In the *distribution* approach, i considers the whole posterior suspiciousness distribution (Equation (3)) in making the cut-off decision against j. From i's perspective, after observing an assessment sequence $\mathcal{A}$, the probability $P_g(\mathcal{A})$ that j is good is:

$$P_g(\mathcal{A}) = \int_0^{L_e} P(S_j|\mathcal{A})\, \mathrm{d}S_j; \tag{5}$$

the probability $P_e(\mathcal{A})$ that j is evil is:

$$P_e(\mathcal{A}) = 1 - P_g(\mathcal{A}) = \int_{L_e}^1 P(S_j|\mathcal{A})\, \mathrm{d}S_j. \tag{6}$$

Let $\mathcal{C} = (\int_0^1 S_j^{s_\mathcal{A}} (1 - S_j)^{|\mathcal{A}| - s_\mathcal{A}}\, \mathrm{d}S_j)^{-1}$ be the (probability) normalization factor in Equation 3; we have:

$$P_g(\mathcal{A}) = \mathcal{C} \int_0^{L_e} S_j^{s_\mathcal{A}} (1 - S_j)^{|\mathcal{A}| - s_\mathcal{A}}\, \mathrm{d}S_j \tag{7}$$

and

$$P_e(\mathcal{A}) = \mathcal{C} \int_{L_e}^1 S_j^{s_\mathcal{A}} (1 - S_j)^{|\mathcal{A}| - s_\mathcal{A}}\, \mathrm{d}S_j. \tag{8}$$

When $P_g(\mathcal{A}) \geq P_e(\mathcal{A})$, the evidence collected so far (i.e., $\mathcal{A}$) is favorable to j. However, when $P_g(\mathcal{A}) < P_e(\mathcal{A})$, the evidence is unfavorable to j and suggests that j might be an evil node. i needs to *decide whether to cut j off.*

The structure of the behavioral malware characterization model (specifically, a single threshold L_e is used to distinguish the nature of a node) gives rise to a subtlety concerning i's prejudice against j in the distribution approach. By Section 1.2 of the supplementary document,if i makes no presumption on j's suspiciousness, when no assessment has been made yet (i.e., $\mathcal{A} = \emptyset$), $P(S_j|\mathcal{A}) = 1$. If $L_E \neq 0.5$, by Equations (5) and (6), either $P_g(\mathcal{A}) < P_e(\mathcal{A})$ (if $L_E < 0.5$) or $P_g(\mathcal{A}) > P_e(\mathcal{A})$ (if

[1] When we use proportional form in this paper, we have implicitly done the same thing.

$L_E > 0.5$). In other words, while i makes no presumption on j's suspiciousness, i may nevertheless be prejudiced against j by the distribution approach's decision rule.

This leads to a discussion on whether such prejudices are warranted. The choice of L_e depends on the assessment mechanism itself and, as mentioned previously, if the probabilistic distributions of suspiciousness of both good and evil nodes are known, can be determined by minimizing Bayesian decision errors. If $L_e > 0.5$, the assessment mechanism is biased towards false positive (good nodes' actions being assessed as suspicious); if $L_e < 0.5$, the assessment mechanism is biased towards false negative (evil nodes' actions being assessed as non-suspicious). However, before any assessment is made, i has no clue about the true nature of j. A bias in the assessment mechanism should not affect the i's neutrality on j's nature before the first assessment is made. Thus, we stipulate that the comparison between $P_g(\mathcal{A})$ and $P_e(\mathcal{A})$ should be made only when $\mathcal{A} \neq \emptyset$.

Alternatively, in the *maximizer* approach, i uses the suspiciousness distribution's maximizer (Equation (4)) when making the cut-off decision against j. The justification for the maximizer approach is that the suspicious distribution's maximizer is the *single most probable* estimation of j's suspicisousness given the evidence. The maximizer approach precludes the prejudice problem, because the maximizer is undefined when $\mathcal{A} = \emptyset$. Similar to the distribution approach, i compares evidence that is both favorable and unfavorable to j. Evidence $\mathcal{A}$ is favorable to j if $s_{\mathcal{A}}/|\mathcal{A}| \leq L_e$ and is unfavorable to j if $s_{\mathcal{A}}/|\mathcal{A}| > L_e$. The maximizer approach significantly reduces the computation cost, in comparison with the distribution approach, while partially discarding information contained in the suspiciousness distribution derivable from the evidence collected so far.

Whichever approach is taken, the cut-off decision problem has an *asymmetric* structure in the sense that cutting j off will immediately terminate the decision process (i.e., i will cease connecting with j; no further evidence will be collected), while the opposite decision will not. Thus, we only need to consider the decision problem when i considers cutting j off due to unfavorable evidence against j.

The cut-off decision is made based on the risk estimation of such a decision. The key insight is that i shall estimate the cut-off decision's risk by *looking ahead*.

More specifically, given the current assessment sequence $\mathcal{A} = (a_1, \ldots, a_A)$, the next assessment a_{A+1} (which has not been taken yet) might be either 0 (non-suspicious) or 1 (suspicious). Let $\mathcal{A}' = (\mathcal{A}, a_{A+1})$.

If $a_{A+1} = 1$, by Section 1.3 of the supplementary document, either $P_g(\mathcal{A}') < P_g(\mathcal{A}) < P_e(\mathcal{A}) < P_e(\mathcal{A}')$ (the distribution approach) or $s_{\mathcal{A}'}/|\mathcal{A}'| = (1+s_{\mathcal{A}})/(1+|\mathcal{A}|) > s_{\mathcal{A}}/|\mathcal{A}| > L_e$ (the maximizer approach): The evidence against j becomes more unfavorable.

However, if $a_{A+1} = 0$, the evidence might become either favorable or unfavorable to j. If the evidence is still unfavorable toward j, we say that i's decision of cutting j off is *one-step-ahead robust*. If the cut-off decision is one-step-ahead robust, i is certain that exposing itself to the potential danger of infection by collecting *one further assessment* on j will not change the outlook that j is evil.

Similarly, i can look *multiple* steps ahead. In fact, the number of steps i is willing to look ahead is a *parameter* of the decision process rather than a *result* of it. This parameter shows i's willingness to be exposed to a higher infection risk in exchange for a higher certainty about the nature of j and a lower risk of cutting off a good neighbor; in other words, it reflects i's *intrinsic* risk inclination against malware infection.

Definition 1 (Look-ahead λ): The *look-ahead* λ is the number of steps i is willing to look ahead before making a cut-off decision.

We can make a similar decision-robustness definition for look-ahead λ.

Definition 2 (λ-robustness): At a particular point in i's cut-off decision process against j (with assessment sequence $\mathcal{A} = (a_1, \ldots, a_A)$), i's decision of cutting j off is said to be *λ-step-ahead robust*, or simply *λ-robust*, if 1) the current evidence $\mathcal{A}$ is unfavorable toward j; 2) even if the next λ assessments $(a_{A+1}, \ldots, a_{A+\lambda})$) all turn out to be non-suspicious (i.e., 0), the evidence against j is still unfavorable.

Given the look-ahead λ, the proposed malware containment strategy is *to cut j off if the cut-off decision is λ-robust, and not to cut j off otherwise.*

In Section 2 of the supplementary document, we discuss how to adapt the look-ahead λ to individual nodes' intrinsic risk inclinations against the malware.

3.2 Neighborhood Watch

Besides using i's own assessments, i may incorporate other neighbors' assessments in the cut-off decision against j. This extension to the evidence collection process is inspired by the real-life neighborhood (crime) watch program, which encourages residents to report suspicious criminal activities in their neighborhood. Similarly, i shares assessments on j with its neighbors, and receives their assessments on j in return.

In the neighborhood-watch model, the malicious nodes that are able to transmit malware (we will see next that there may be malicious nodes whose objective is other than transmitting malware) are assumed to be *consistent over space and time*. These are common assumptions in distributed trust management systems (summarized in Section 5), which incorporate neighboring nodes' opinions in estimating a local trust value.

By being consistent over space, we mean that evil nodes' suspicious actions are observable to all their neighbors, rather than only a few. If this is not the case, the evidence provided by neighbors, even if truthful, will contradict local evidence and, hence, cause confusions:

Nodes shall discard received evidence and fall back to the household watch model.

By being consistent over time, we mean that evil nodes can not play strategies to fool the assessment mechanism. This is equivalent to the functional assumption in characterizing the nature of nodes by suspiciousness (Equation 1). The case in which the evil nodes can circumvent the suspiciousness characterization (such as by first accumulating good assessments, and then launch an attack through a short burst of concentrated suspicious actions) calls for game-theoretic analysis and design, and is beyond the scope of this paper. Instead, we propose a behavioral characterization of proximity malware; further game-theoretic analysis and design could base on this foundation.

3.2.1 Challenges

Two cases complicate the neighborhood watch model: *liars* and *defectors*.

Liars are those evil nodes who confuse other nodes by sharing false assessments. A false assessment is either a *false praise* or a *false accusation*. False praises understate evil nodes' suspiciousness, while false accusations exaggerate good nodes' suspiciousness. Furthermore, a liar can fake assessments on nodes that it has never met with. To hide their true nature, liars may do no evil other than lying, and, therefore, have low suspiciousness.

Defectors are those nodes that change their nature due to malware infections. They start out as good nodes and faithfully share assessments with their neighbors; however, due to malware infections, they become evil. Their behaviors after the infection are under the control of the malware.

These complications call for *evidence consolidation*. Two extremal, but naive, evidence-consolidation strategies are 1) to trust no one and 2) to trust everyone. The former degenerates to the household-watch model with the twist of the defectors (defectors change their nature and hence their behavioral pattern); the latter leads to confusions among good nodes.

3.2.2 Evidence

For a pair of neighboring nodes i and j, let $\mathcal{N}_i$ and $\mathcal{N}_j$ be the neighbors of i and j, respectively. At each encounter, i shares with j its assessments on the neighbor set $\mathcal{N}_i - \{j\}$, and j shares with i its assessments on the neighbor set $\mathcal{N}_j - \{i\}$.

Since the cut-off decision only needs to be made against a neighbor, i only considers the assessments of its own neighbors $\mathcal{N}_i \cap (\mathcal{N}_j - \{i\})$ from the evidence provided by j. Without superimposed trust relationships among the nodes in the model, i and j only share *their own* assessments, instead of forwarding the ones provided by their neighbors.

3.2.3 Evidence Aging

The presence of defectors breaks the assumption when we characterize a node's nature by suspiciousness in Equation 1. A defector starts as a good node but turns evil due to malware infections; the assessments collected before the defector's change of nature, even truthful, are misleading.

To alleviate the problem of outdated assessments, old assessments are discarded in a process called *evidence aging*. Each assessment is associated with a timestamp. Only assessments with timestamps less than a specific *aging window* T_E from now are included in the cut-off decision.

To see that the aging window T_E alleviates the defector problem, consider a node that is infected at time T. Without evidence aging, all evidence before T mounts to testify that the node is good; if the amount of this prior evidence is large, it may take a long time for its neighbors to find out about the change in its nature. In comparison, with evidence aging, at time $T + T_E$, all prior evidence expires and only those assessments after the infection are considered, which collectively testify against the node.

However, in practice, the choice of the aging window T_E depends on the context. While a small T_E may speed up the detection of defectors by reducing the impact of stale information, T_E must be large enough to accommodate enough assessments to make a sound cut-off decision. If T_E is too small, a node will not have enough assessments to make a λ-robust cut-off decision.

3.2.4 Evidence Consolidation

We propose two alternative methods, *dogmatic filtering* and *adaptive look-ahead*, for consolidating evidence provided by other nodes, while containing the negative impact of liars. For exposition, we consider a scenario in which node i uses the assessments within the evidence aging window $[T - T_E, T]$ provided by i's neighbors (other than one of the neighbors, say, j) in making the cut-off decision against j.

The following observation inspires our solution: Given enough assessments, i is more likely to correctly estimate j's suspiciousness than otherwise. Consider a simple numerical illustration. If j has in total 4 suspicious actions and 12 non-suspicious actions assessed by its neighbors, its (true) suspiciousness is $\frac{4}{4+12} = 0.25$. If i has made 4 out of the $4 + 12 = 16$ assessments, by the space-consistency assumption, i is equally likely to obtain *any* sub-sequence of the 16 assessment sequence. The total possibilities of i making x $(0 \leq x \leq 4)$ suspicious assessments and $4 - x$ non-suspicious assessments are $\binom{4}{x}\binom{12}{4-x}$; a straightforward calculation shows that the number is maximized when $x = 1$. In other words, i is more likely to estimate j to be $\frac{1}{1+3} = 0.25$, which agrees with the true suspiciousness, compared to otherwise.

In general, suppose j has been assessed n times by its neighbors, and s of them are suspicious. Its suspicisousness, by definition, is $\frac{s}{n}$. If n' $(0 < n' \leq n)$ of the assessments are from i and s' $(s - (n - n') \leq s' \leq \min(s, n'))$ of them are suspicious (thus, from i's perspective, j's suspiciousness is $\frac{x'}{n'}$), s' is more likely to be either $\lfloor \frac{s}{n} n' \rfloor$ or $\lceil \frac{s}{n} n' \rceil$ (i.e., i's estimation of j's suspiciousness agrees with the true suspiciousness) than otherwise, since, as in the previous numerical example, $\binom{s}{s'}\binom{n-s}{n'-s'}$ is maximized when $\frac{s'}{s} \approx \frac{n'-s'}{n-s}$ for a given n'.

The implications are:

- Given enough assessments, honest nodes are likely to obtain a close estimation of a node's suspiciousness (suppose they have not cut the node off yet), even if they only use their own assessments.
- The liars have to share a significant amount of false evidence to sway the public's opinion on a node's suspiciousness.
- The most susceptible victims of liars are the nodes that have little evidence.

Dogmatic filtering *Dogmatic filtering* is based on the observation that one's own assessments are truthful and, therefore, can be used to bootstrap the evidence consolidation process. A node shall only accept evidence that will not sway its current opinion too much. We call this observation the *dogmatic principle*.

Our interpretation of the dogmatic principle depends on the following generalization of Definition 2.

Definition 3 (λ-robust judgment): Let $\mathcal{A}$ be the suspicious-action assessments that i has on j. We say that i's *judgment on j's nature is λ-robust (or $(-\lambda)$-robust) based on $\mathcal{A}$*, if 1) the evidence $\mathcal{A}$ is favorable (or unfavorable) toward j, 2) the evidence remains so *even if* the next λ assessments are all suspicious (or non-suspicious), and 3) the evidence becomes unfavorable (or favorable) toward j if the next $\lambda+1$ assessments are all suspicious (or non-suspicious).

As a special case, if a judgment is not even 1-robust (or (-1)-robust), we say that the judgment is 0-robust or not robust at all.

λ-robust judgment reflects i's certainty of its judgment on j's nature (based on the evidence collected so far). The λ-robust cut-off decision against j (Definition 2) is equivalent to the $(-\lambda)$-robust judgment on the (evil) nature of j. The sign of λ in Definition 3 represents j's nature: A negative number represents evilness, and a positive number represents goodness.

i's cut-off decision against j works as follows with dogmatic filtering.

- i will not consider cutting j off until i has at least one assessment on j.
- After its first encounter with j and with its own assessments $\mathcal{A}$ with the evidence aging window $[T-T_E, T]$, i considers whether or not to take another neighbor k's alleged assessments on j within the same window $\mathcal{B}$ when i and k meet.
- Suppose that the judgment on j's nature is $\lambda_{\mathcal{A}}$-robust and $\lambda_{(\mathcal{A}+\mathcal{B})}$-robust, based on $\mathcal{A}$ and $(\mathcal{A}+\mathcal{B})$, respectively. i will take $\mathcal{B}$ only if $\lambda_{\mathcal{A}} \neq 0$ and $\frac{|\lambda_{\mathcal{A}} - \lambda_{(\mathcal{A}+\mathcal{B})}|}{|\lambda_{\mathcal{A}}|} \leq \delta$; $\delta > 0$ and is called the *dogmatism*.
- i makes a λ-robust cut-off decision against j, based on either $\mathcal{A}$ or $(\mathcal{A}+\mathcal{B})$, depending on whether $\mathcal{B}$ has passed the dogmatism test.

With dogmatic filtering, i is very conservative when its certainty about j's nature is still low (i.e., $\lambda_{\mathcal{A}}$ is small). At this early stage, i will accept the evidence provided by j only if the evidence would not significantly change its certainty on j's nature. In particular, if $\lambda \leq 1$, i will *never* accept a piece of evidence that would change its judgment on j's nature because $|\lambda_{\mathcal{A}} - \lambda_{(\mathcal{A}+\mathcal{B})}| > |\lambda_{\mathcal{A}}|$ if $\mathcal{A}$ and $(\mathcal{A}+\mathcal{B})$ are of different signs.

Dogmatic filtering significantly contains the impact of liars on i while still allowing a change of certainty (on j's nature) comparable to its own. The aforementioned observation that the liars have to fabricate a significant amount of false evidence to confuse honest nodes means that the evidence $\mathcal{B}$ provided by a liar k must have a high $\lambda_{\mathcal{B}}$ (albeit of the wrong sign) to be effective in confusing i. The liar's strategy will not work because i will refuse to take $\mathcal{B}$ when $|\lambda_{\mathcal{A}}|$ is small with dogmatic filtering, while $\lambda_{\mathcal{A}}$ and $\lambda_{\mathcal{B}}$ should be of different signs when $\lambda_{\mathcal{A}}$ is large (because by then, i should have a close estimation of j's true suspiciousness, and hence, $\lambda_{\mathcal{A}}$ is of the correct sign). The evidence filtering works even when the liars are the majority among i's neighbors.

Adaptive look-ahead *Adaptive lookahead* takes a different approach towards evidence consolidation. Instead of deciding whether to use the evidence provided by others *directly* in the cut-off decision, adaptive lookhead *indirectly* uses the evidence by adapting the *steps to look ahead* to the diversity of opinion.

Adaptive look-ahead works as follows.

- Suppose that at a particular moment, the distribution maximizer derived from the assessments (within the evidence aging window) on j (Equation (4)) made by i is s_0; similarly, the distribution maximizer derived from the assessments (within the evidence aging window) on j that i *received* from its neighbors is $s_1, s_2, \ldots, s_n$.
- i computes the following *ego-centric variance* σ_i as a metric on the diversity of opinions (from its own assessments):

$$\sigma_i = \frac{\sqrt{\sum_{i=1}^{n}(s_i - s_0)^2}}{n}. \tag{9}$$

- Let the *maximal* ego-centric variance up to (and including) now be σ_i^* (thus, we have $\sigma_i \leq \sigma_i^*$). i makes its cut-off decision against j *if the decision is $f(\sigma_i, \sigma_i^*, \lambda)$-robust*, where $f(\cdot,\cdot,\cdot)$ is a three-parameter integer function ranging from 0 to λ, which we call the *adaptive-lookahead function*. A particular instantiation is the *linear* adaptive lookahead function.

$$f(\sigma_i, \sigma_i^*, \lambda) = \lceil \lambda \frac{\sigma_i}{\sigma_i^*} \rceil. \tag{10}$$

The idea of adaptive look-ahead is to adapt the risk inclination, embodied in the λ-robust cut-off decision in Definition 2, to the diversity of public opinions, embodied in the ego-centric variance in Equation (9). The dogmatism principle underlies the use of the ego-centric variance: The agreement of the public's opinions

TABLE 1: Dataset statistics.

	nodes	entries	time span	avg. interval
Haggle	41	$112,295$	15 days	12 secs
MIT reality	96	$114,046$	490 days	371 secs

TABLE 2: Neighbor nature and cut-off decision combination.

	...gets cut off.	...stays connected.
An evil neighbor...	True positive.	False negative.
A good neighbor...	False positive.	True negative.

with that of i is an indication that i is approaching the true suspiciousness; thus, to expedite the detection of evil nodes (and hence reduce the risk of infection from further contact), i reduces the steps to look ahead in making the cut-off decision.

Because the value of the adaptive-lookahead function is no greater than 1, the worst that liars can do is to degenerate i's cut-off decision to a λ-robust one. Also, since i has a chance of estimating a close-to-true suspiciousness than otherwise, liars' false opinions are likely to be different from that of i, and good nodes' opinions are likely to agree with that of i. Thus, i will be more proactive if good nodes make up the majority of its neighborhood and less so if the liars are the majority.

4 Simulation

4.1 Datasets

We verify our design with two real mobile network traces: Haggle [22] and MIT reality [23].

The raw datasets are rich in information, some of which is irrelevant to our study, e.g., call logs and cell tower IDs in MIT reality. Therefore, we remove the irrelevant fields and retain the node IDs and time-stamps for each pair-wise node encounter. Since the Haggle dataset has only $22,459$ entries spanning over 3 days, we repeat it another 4 times to make it into a dataset with $112,295$ entries spanning over 15 days, and thus make it comparable to the MIT reality dataset in quantity. Some statistics of the processed datasets are summarized in Table 1.

4.2 Setup

Without loss of generality, we choose $L_e = 0.5$ to be the line between good and evil. For each dataset, we randomly pick 10% of the nodes to be the evil nodes and assign them with suspiciousness greater than 0.5; the rest of the nodes are good nodes and are assigned suspiciousness less than 0.5.

For a particular pairwise encounter, a uniform random number is generated for each node; a node receives a "suspicious" assessment (by the other node) if the random number is greater than its suspiciousness and receives a "non-suspicious" assessment otherwise. Thus, each assessment is binary, while the frequency of "suspicious" assessments for a particular node reflects its suspiciousness in the long term.

4.3 Performance Metric

The performance comparison is based on two metrics: *detection rate* and *false positive rate*. The categories of the "neighbor's nature" and "cut-off decision" combinations are shown in Table 2. For each combination, we sum up all the decisions made by *good* nodes (evil nodes' cut-off decisions are irrelevant) and obtain four counts: TP (true positives), FN (false negatives), TN (true negatives), and FP (false positives). The detection rate DR is defined as:

$$DR = \frac{TP}{TP + FN} \times 100\%,$$

and the false positive rate FPR is defined as:

$$FPR = \frac{FP}{FP + TN} \times 100\%.$$

A high detection rate and a low false positive rate are desirable. When a balance must be stricken between the two, one might be emphasized over the other, depending on the context.

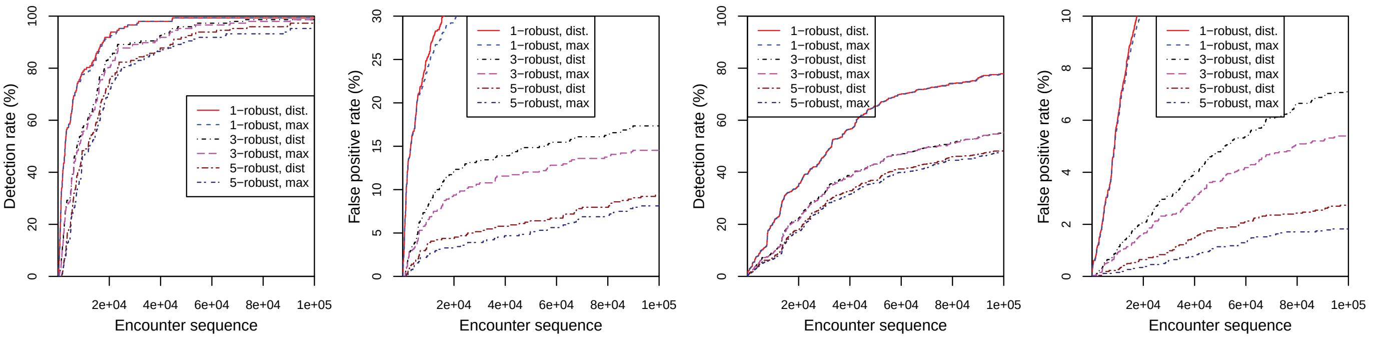
4.4 Results

4.4.1 Look-ahead: distribution vs maximizer

We compare the two alternative approaches, distribution and maximizer, to the look-ahead strategy (Section 3.1). The results are shown in Figure 2.

The look-ahead parameter λ reflects a node's intrinsic (infection) risk inclination. In both Haggle (Figures 2a and 2b) and MIT reality (Figures 2c and 2d), the λ-robust cut-off strategy with a larger λ corresponds to a higher detection rate (in the early stage for Haggle and throughout for MIT reality) and a significantly lower false positive rate (for both datasets). In Haggle, the eventual detection rates for all three look-ahead parameters are close to 100%. The difference in the eventual detection rate between Haggle and MIT reality is attributed to the different contact patterns in these datasets: The contact pattern in Haggle is more homogeneous than that in MIT reality, in the sense that the variation of the interval between encounters is significantly higher and a few nodes contribute most of the assessments in MIT reality. Thus, the detection rate is more sensitive to the change of λ in MIT reality than in Haggle.

In both datasets, the detection rate and false positive rate are comparable for the distribution and maximizer approach, with the distribution approach having a slightly higher detection rate and false positive rate. The small difference in performance, coupled with the significant reduction in computation overhead (integration for the distribution approach versus arithmetic operations for the maximizer approach), make the maximizer approach with a moderate λ as the preferred look-ahead strategy. In the following sections, we show results for the maximizer approach with $\lambda = 3$.



(a) Haggle.

(b) Haggle.

(c) MIT reality.

(d) MIT reality.

Fig. 2: Performance comparison between the λ-robust cut-off strategy with the distribution (dist) and maximizer (max) evidence weighing approaches; $\lambda = 1, 3,$ and 5.

4.4.2 Look-ahead

We compare Bayesian-based strategies with, and without, the look-ahead extension (i.e., λ-robust cut-off decision) under the household-watch model (i.e., no evidence exchange). The vanilla Bayesian strategy does not look ahead and proceeds with cutting-off once the evidence becomes unfavorable to the neighbor. It can be seen as a degenerated λ-robust cut-off strategy with $\lambda = 0$. The results are shown in Figure 3.

In Figure 3, the vanilla Bayesian strategy has the highest detection and false positive rate. Both rates drop with an increasing look-ahead parameter. However, the false detection rate drops much faster than the detection rate. Indeed, for Haggle, the 1-robust and the vanilla Bayesian strategies have almost the same detection rate after $30,000$ encounters, but there is a 30% difference in the false positive rate. The difference in detection rate is more pronounced for MIT reality, but the reduction in false positive rate far outweighs that of detection rate. For the risk-taking nodes, sacrificing a little detection rate for a large reduction in false positive rate is desirable: the look-ahead parameter λ provides an effective mechanism to tune for a desirable balance.

The results confirm the intuition that leads to the look-ahead extension to the vanilla Bayesian strategy: Being conservative in making cut-off decisions (by looking ahead) pays off by retaining utility without sacrificing much security.

4.4.3 Evidence consolidation

We also evaluate the benefits of sharing assessments among nodes, and the effect of the proposed evidence consolidation strategies in minimizing the negative impact of liars on the shared evidence's quality. We compare the dogmatic filtering (with dogmatism of 0.0001, 0.01, and 1, respectively) and adaptive look-ahead evidence consolidation methods with two other (naive) evidence consolidation methods: 1) taking *no* indirect evidence, i.e., look-ahead with no evidence consolidation, and 2) taking *all* indirect evidence without filtering.

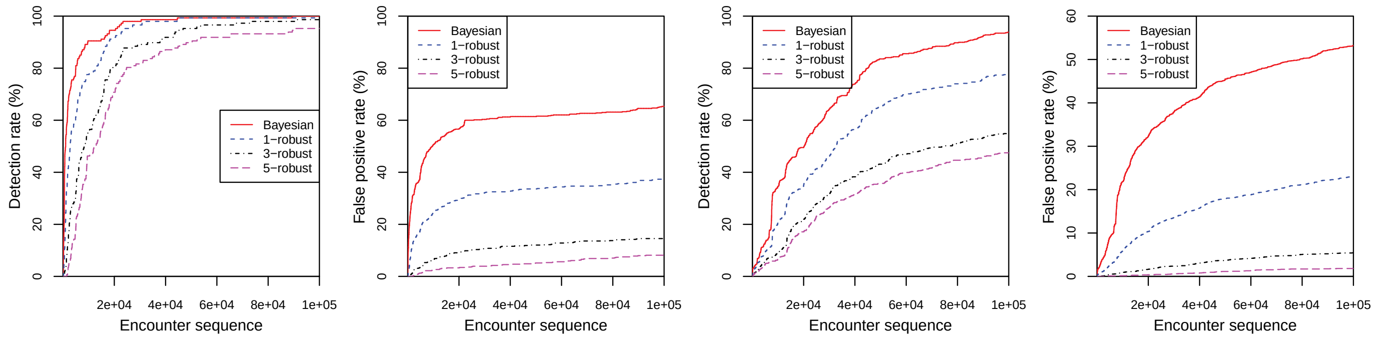
In our study, 10% of the evil nodes play the dual roles of evil-doers and liars. There are many possible liar strategies. Based on our observations in Section 3.2.4, we adopt an *exaggerated false praise/accusation* liar strategy. More specifically, a liar (falsely) accuses good nodes of suspicious actions and (falsely) praises other evil nodes for non-suspicious actions. Besides, to exert a significant influence on the public opinion, they exaggerate the false praises/accusations by 10 times (since they are only 10% of the whole population). The results on the performance of various evidence consolidation strategies under this setting are shown in Figure 4.

Figure 4 clearly shows the negative impact of liars on malware detection if evidence is not filtered: Under the influence of liars, the naive "all" strategy has a low detection rate and a high false positive rate. This calls for a non-trivial evidence consolidation strategy to deal with the liars.

Both dogmatic filtering and adaptive look-ahead show significant increases in detection rate and modest increases in false positive rate over the baseline 3-robust lookahead strategy with no evidence filtering. Together with Figure 3, the results indicate that the 3-robust look-ahead, with either dogmatic filtering or adaptive lookahead, is comparable in detection rate and, even in the presence of liars, shows a significantly lower false positive rate in comparison with both the Bayesian and 1-robust strategies.

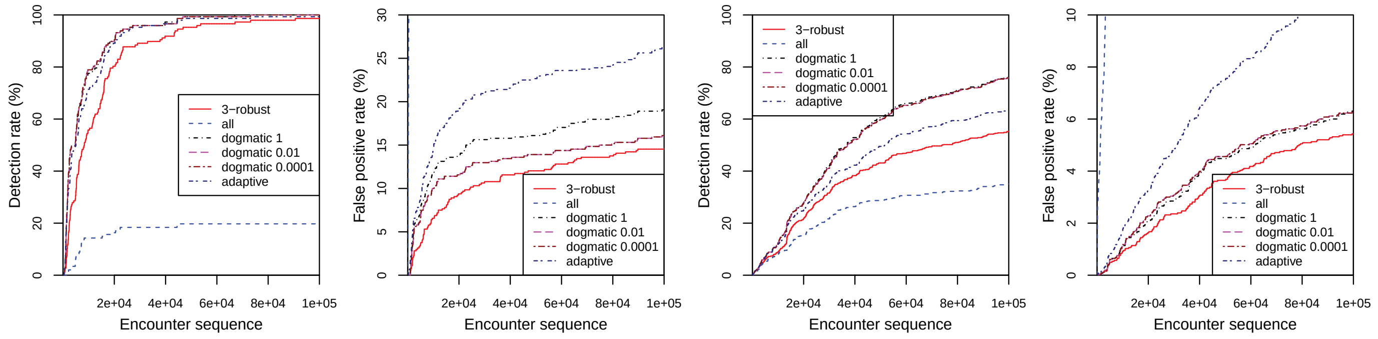
In Figure 4, the eventual detection rates converge to almost 100% for Haggle but diverge for MIT reality. The convergence in detection rate is expected for a homogeneous dataset like Haggle, in which most nodes are well-connected and are able to collect enough evidence to eventually make a sound cut-off decision. In this case, evidence consolidation helps to expedite the decision-making process without driving the false positive rate up too much. A closer look at MIT reality indicates that this dataset is highly heterogeneous: A few well-connected nodes contribute most of the assessments, and leave the other less well-connected nodes with insufficient evidence to make a λ-robust judgment alone. In this case, evidence consolidation helps the latter nodes in collecting enough evidence to make a λ-robust decision.

Two of the dogmatic filtering strategies (with a dog-



(a) Haggle. (b) Haggle. (c) MIT reality. (d) MIT reality.

Fig. 3: Performance comparison between the vanilla Bayesian (degenerated 0-robust) cut-off strategy and the 3-robust look-ahead cut-off strategy.



(a) Haggle. (b) Haggle. (c) MIT reality. (d) MIT reality.

Fig. 4: Performance impact of various evidence consolidation methods on the look-ahead cut-off strategy. *all*: naive strategy without filtering (Section 3.2); *dogmatic* δ: dogmatic filtering with dogmatism δ (Section 3.2.4); *adaptive*: adaptive lookahead (Section 3.2.4).

matism of 0.01 and 0.0001) show almost the same performance, with the other dogmatic filtering strategy (with a dogmatism of 1) show a slight difference in comparison with other strategies. In both datasets, the adaptive look-ahead strategy shows an inferior performance in comparison to the three variations of the dogmatic filtering strategy. However, it automatically (i.e., with no parameter to tune) achieves superior detection rate over both Bayesian and 3-robust strategies in the presence of liars.

5 Related Work

Proximity malware and mitigation schemes. Su et al. collected Bluetooth traces and demonstrated that malware could effectively propagate via Bluetooth with simulations [24]. Yan et al. developed a Bluetooth malware model [25]. Bose and Shin showed that Bluetooth can enhance malware propagation rate over SMS/MMS [26]. Cheng et al. analyzed malware propagation through proximity channels in social networks [27]. Akritidis et al. quantified the threat of proximity malware in wide-area wireless networks [4]. Li et al. discussed optimal malware signature distribution in heterogeneous, resource-constrained mobile networks [28]. In traditional, non-DTN, networks, Kolbitsch et al. [8] and Bayer et al. [9] proposed to detect malware with learned behavioral model, in terms of system call and program flow. We extend the Naive Bayesian model, which has been applied in filtering email spams [13, 14, 15] , detecting botnets [16], and designing IDSs [10, 17], and address DTN-specific, malware-related, problems. In the context of detecting slowly propagating Internet worm, Dash et al. presented a distributed IDS architecture of local/global detector that resembles the neighborhood-watch model, with the assumption of attested/honest evidence, i.e., without liars [10].

Mobile network models and traces. In mobile networks, one cost-effective way to route packets is via the short-range channels of intermittently connected smartphones [29, 30, 31]. While early work in mobile networks used a variety of simplistic random i.i.d. models, such as random waypoint, recent findings [32] show that these models may not be realistic. Moreover, many recent studies [33], based on real mobile traces, revealed that a node's mobility shows certain social network properties. Two real mobile network traces were used in our study.

Reputation and trust in networking systems. In the neighborhood watch model, suspiciousness, defined in Equation (1), can be seen as nodes' reputation; to cut a node off is to decide that the node is not trustworthy. Thus, our work can be viewed from the perspective of reputation/trust systems. Three schools of thoughts emerge from previous studies. The first one uses a central authority, which by convention is called the trusted third party. In the second school, one global trust value is drawn and published for each node, based on other nodes' opinions of it; eigenTrust [34] is an example. The last school of thoughts includes the trust management systems that allow each node to have its own view of other nodes [35, 36]. Our work differs from previous trust management work in addressing two DTN-specific, malware-related, trust management problems: 1) insufficient evidence vs. evidence collection risk and 2) sequential and distributed online evidence filtering.

6 Concluding Remarks

Behavioral characterization of malware is an effective alternative to pattern matching in detecting malware, especially when dealing with polymorphic or obfuscated malware. Naive Bayesian model has been successfully applied in non-DTN settings, such as filtering email spams and detecting botnets. We propose a general behavioral characterization of DTN-based proximity malware. We present *look-ahead*, along with *dogmatic filtering* and *adaptive look-ahead*, to address two unique challenging in extending Bayesian filtering to DTNs: "insufficient evidence vs. evidence collection risk" and "filtering false evidence sequentially and distributedly". In prospect, extension of the behavioral characterization of proximity malware to account for strategic malware detection evasion with game theory is a challenging yet interesting future work.

References

[1] Trend Micro Inc. (2004) SYMBOS_CABIR.A. [Online]. Available: http://goo.gl/aHcES

[2] [Online]. Available: http://goo.gl/iqk7

[3] Trend Micro Inc. (2009) IOS_IKEE.A. [Online]. Available: http://goo.gl/z0j56

[4] P. Akritidis, W. Chin, V. Lam, S. Sidiroglou, and K. Anagnostakis, "Proximity breeds danger: emerging threats in metro-area wireless networks," in *Proc. USENIX Security*, 2007.

[5] A. Lee. (2012) FBI warns: New malware threat targets travelers, infects via hotel Wi-Fi. [Online]. Available: http://goo.gl/D8vNU

[6] NFC Forum. About NFC. [Online]. Available: http://goo.gl/zSJqb

[7] Wi-Fi Alliance. Wi-Fi Direct. [Online]. Available: http://goo.gl/fZuyE

[8] C. Kolbitsch, P. Comparetti, C. Kruegel, E. Kirda, X. Zhou, and X. Wang, "Effective and efficient malware detection at the end host," in *Proc. USENIX Security*, 2009.

[9] U. Bayer, P. Comparetti, C. Hlauschek, C. Kruegel, and E. Kirda, "Scalable, behavior-based malware clustering," in *Proc. IEEE NDSS*, 2009.

[10] D. Dash, B. Kveton, J. Agosta, E. Schooler, J. Chandrashekar, A. Bachrach, and A. Newman, "When gossip is good: Distributed probabilistic inference for detection of slow network intrusions," in *Proc. AAAI*, 2006.

[11] G. Zyba, G. Voelker, M. Liljenstam, A. Méhes, and P. Johansson, "Defending mobile phones from proximity malware," in *Proc. IEEE INFOCOM*, 2009.

[12] F. Li, Y. Yang, and J. Wu, "CPMC: an efficient proximity malware coping scheme in smartphone-based mobile networks," in *Proc. IEEE INFOCOM*, 2010.

[13] I. Androutsopoulos, J. Koutsias, K. Chandrinos, and C. Spyropoulos, "An experimental comparison of naive bayesian and keyword-based anti-spam filtering with personal e-mail messages," in *Proc. ACM SIGIR*, 2000.

[14] P. Graham. Better Bayesian filtering. [Online]. Available: http://goo.gl/AgHkB

[15] J. Zdziarski, *Ending spam: Bayesian content filtering and the art of statistical language classification*. No Starch Press, 2005.

[16] R. Villamarín-Salomón and J. Brustoloni, "Bayesian bot detection based on DNS traffic similarity," in *Proc. ACM SAC*.

[17] J. Agosta, C. Diuk-Wasser, J. Chandrashekar, and C. Livadas, "An adaptive anomaly detector for worm detection," in *Proc. USENIX SysML*, 2007.

[18] S. Marti, T. Giuli, K. Lai, M. Baker *et al.*, "Mitigating routing misbehavior in mobile ad hoc networks," in *Proc. ACM MobiCom*, 2000.

[19] P. Michiardi and R. Molva, "Core: a collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks," in *IFIP TC6/TC11 Sixth Joint Working Conference on Communications and Multimedia Security*, 2002, p. 107.

[20] S. Buchegger and J. Le Boudee, "Self-policing mobile ad hoc networks by reputation systems," *IEEE Comm. Mag.*, vol. 43, no. 7, pp. 101–107, 2005.

[21] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. Wiley-Interscience, Nov. 2001.

[22] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "CRAWDAD data set cambridge/haggle (v. 2006-09-15)," http://goo.gl/RJrKN, Sep. 2006.

[23] N. Eagle and A. Pentland, "CRAWDAD data set MIT/reality (v. 2005-07-01)," http://goo.gl/V3YKc, Jul. 2005.

[24] J. Su, K. Chan, A. Miklas, K. Po, A. Akhavan, S. Saroiu, E. de Lara, and A. Goel, "A preliminary investigation of worm infections in a bluetooth environment," in *Proc. ACM WORM*, 2006.

[25] G. Yan, H. Flores, L. Cuellar, N. Hengartner, S. Eidenbenz, and V. Vu, "Bluetooth worm propagation: mobility pattern matters!" in *Proc. ACM ASIACCS*, 2007.

[26] A. Bose and K. Shin, "On mobile viruses exploiting messaging and bluetooth services," in *Proc. IEEE SecureComm*, 2006.

[27] S. Cheng, W. Ao, P. Chen, and K. Chen, "On modeling malware propagation in generalized social networks," *IEEE Comm. Lett.*, vol. 15, no. 1, pp. 25–27, 2011.

[28] Y. Li, P. Hui, L. Su, D. Jin, and L. Zeng, "An optimal distributed malware defense system for mobile networks with heterogeneous devices," in *Proc. IEEE SECON*, 2011.

[29] A. Vahdat and D. Becker, "Epidemic routing for partially-connected ad hoc networks," Duke Univ., Tech. Rep., 2002.

[30] J. Burgess, B. Gallagher, D. Jensen, and B. Levine, "MaxProp: routing for vehicle-based disruption-tolerant networks," in *Proc. IEEE INFOCOM*, 2006.

[31] V. Erramilli, M. Crovella, A. Chaintreau, and C. Diot, "Delegation forwarding," in *Proc. ACM MobiHoc*, 2008.

[32] W. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, "Modeling time-variant user mobility in wireless mobile networks," in *Proc. IEEE INFOCOM*, 2007.

[33] E. Daly and M. Haahr, "Social network analysis for information flow in disconnected delay-tolerant MANETs," *IEEE TMC*, vol. 8, no. 5, pp. 606–621, 2009.

[34] S. Kamvar, M. Schlosser, and H. Garcia-Molina, "The eigentrust algorithm for reputation management in P2P networks." in *Proc. ACM WWW*, 2003.

[35] S. Buchegger and J. Boudec, "Performance analysis of the CONFIDANT protocol," in *Proc. ACM MobiHoc*, 2002.

[36] A. Srinivasan, J. Teitelbaum, and J. Wu, "DRBTS: Distributed reputation-based beacon trust system," in *Proc. IEEE DASC*, 2006.

Wei Peng is a PhD student with the Department of Computer and Information Science of Indiana University-Purdue University Indianapolis. His co-advisors are Dr. Feng Li and Dr. Xukai Zou. He has worked on problems on delay-tolerant networks, security, privacy, and social networks. His research vision is to explore human factors in computing and networking and, in turn, make them more human friendly.

Feng Li received his Ph.D. in Computer Science from Florida Atlantic University in Aug. 2009. His Ph.D. advisor is Dr. Jie Wu. He joined the Department of Computer, Information, and Leadership Technology at Indiana University-Purdue University Indianapolis (IUPUI) as an assistant professor in Aug. 2009. His research interests include the areas of wireless networks and mobile computing, security, and trust management. He has published more than 30 papers in conferences and journals.

Xukai Zou is a faculty member with the Department of Computer and Information Sciences at Indiana University-Purdue University Indianapolis. He completed his PhD Degree in Computer Science from University of Nebraska-Lincoln. His current research focus is Applied Cryptography, Network Security, and Communication Networks. His research has been supported by NSF, the Department of Veterans Affairs and Industry such as Cisco.

Jie Wu is chair and professor in the Department of Computer and Information Sciences at Temple University. Prior to joining Temple University, he was a program director at the National Science Foundation. His research interests include wireless networks, mobile computing, routing protocols, fault-tolerant computing, and interconnection networks. Dr. Wu publications include over 550 papers in scholarly journals and conference proceedings. Additionally, he has served on several editorial boards, including IEEE Transactions on Computers and Journal of Parallel and Distributed Computing. Dr. Wu was general co-chair for IEEE MASS 2006, IEEE IPDPS2008 and DCOSS 2009 and is the program co-chair for IEEE INFOCOMM 2011. He served as an IEEE Computer Society distinguished visitor. Currently, he is the chair for the IEEE Technical Committee on Distributed Processing (TCDP), an ACM distinguished speaker and a Fellow of the IEEE. Dr. Wu is the recipient of 2011 China Computer Federation (CCF) Overseas Outstanding Achievement Award.