iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** V **Month of publication:** May 2015

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Savoir-faire malware tracking down in Delay Tolerant Networks

Raksha R S[1], Rakshitha P G[2], Shalini M S[3], Shwetha L[4], Mrs. Ayesha Taranum[5]
8th Sem CSE, Assistant Professor, CSE, GSSSIETW, Mysuru

Abstract — With the universal presence of short-range connectivity technologies (e.g., Bluetooth and, more recently, Wi-Fi Direct) in the consumer electronics market, the delay tolerant-network (DTN) model is becoming a viable alternative to the traditional infrastructural model. 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. In modern network the malware is one of the serious issues where it can be identified by many roles such as email spam, Denial of service and Trojan like viruses. DTN (Delay Tolerant Network) suffered from the above malware related problems. 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 spam's and detecting bonnets. We identify two unique challenges for extending Bayesian malware detection to DTNs ("insufficient evidence vs. evidence collection risk" and "filtering false evidence sequentially and distributed"), and propose a simple yet effective method, look-ahead, to address the challenges

Keywords—Delay-Tolerant Networks; Proximity Malware; behavioral malware characterization;dogmatic filtering,adaptive look-ahead

I. INTRODUCTION

Delay Tolerant Networking is a networking architecture that is designed to provide communications in the most unstable and stressed environments, where the network would normally be subject to frequent and long lasting disruptions and high bit error rates that could severely degrade normal communications. Delay-Tolerant Networks are limited end-to-end connectivity, due to mobility, power saving, or unreliable networks In other words, such networks consist of a few constantly available links to a set of nodes (backed by the cellular channel) and many intermittently available links between (potentially) all the nodes in the network (backed by the proximate channel and defined by the mobility of the nodes).DTNs overcome the problems associated with intermittent connectivity, long or Variable delay, asymmetric data rates, and high error rates by using store-and forward Message switching. The storage places (such as hard disk) can hold messages indefinitely. They are called persistent storage, as opposed to very short-term storage provided by memory chips and buffers. 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. 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. A later example is the iOS-based Ikee worm, which exploited the default SSH password on jailbroken iPhones to propagate through IP-based Wi-Fi connections. Previous researches 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. With the adoption of new short-range communication technologies such as NFC and Wi-Fi Direct that facilitate spontaneous bulk data transfer between spatially proximate mobile devices, the threat of proximity malware is becoming more realistic and relevant than ever. Mobile malware, however, has another opportunity for propagation. It can propagate through direct pair-wise communication mechanisms, such as Bluetooth or Wi-Fi, between devices in geographic proximity.

One way of defending against malware is to detect it based on behavioral characterization which is introduced in this paper. The behavioral characterization, with respect to system calls and program flow is projected as an efficacious alternative to pattern matching for detecting malware. 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 widespread adoption of the mobile devices, coupled with strong economic incentives, includes a class of class of malware that specifically targets DTNs.We call this class of malware as proximity malware. An early example of proximity malware is the Symbian based Cabir worm[1] A later example is the iOS-based Ikee worm, which exploited the default SSH password on jail-broken[2] I phone to propagate through IP-based Wi-Fi connections[3].Previous researches[4],[5] quantify the threat of proximity malware attack in NFC and WI-if direct[6][7].

Proximity malware based on DTN-model brings unique security challenges that are not present and also malware propagation cannot be detected by the cellular carrier in the traditional model. 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].Malware infected node behaviors are observed by others during their multiple opportunistic encounters: Individual observation may be imperfect ,but abnormal behavior of infected nodes are identifiable in the long run.

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 containment capability, such as patching or self-healing [11],[12] we consider a simple “cut - off” strategy. Our focus is on how individual nodes shall make such cut-off decisions against potentially malware-infected nodes based on direct and indirect observations. In the context of DTNs, we face a dilemma when trying to detect proximity malware: Hypersensitivity leads to false positives, while hyposensitivity leads to false negatives. Our solution, look ahead, reflects individual node’s intrinsic risk inclinations against malware infection, to balance between these two extremes.

Basically, a Naïve Bayesian Model is developed. Then Look Ahead is added for addressing the challenges such as ‘Insufficient Evidence and Evidence Collection Risk’. Moreover two extensions, namely Adaptive Look Ahead and Dogmatic Filtering are developed for addressing the challenges of Liars and Defectors. We summarize our contributions below:

We give a general behavioral characterization of proximity malware, which allows for functional but imperfect
Assessments on malware presence.

Under the behavioral malware characterization, and with a simple cut-off malware containment strategy, we formulate the malware detection process as a decision problem. We analyze the risk associated with the decision and design a simple yet effective malware containment strategy, look ahead, which is distributed by nature and reflects an individual node’s intrinsic trade-off between staying connected with other nodes and staying safe from malware (Section III-A).

We consider the benefits of sharing assessments among directly connected nodes and address the challenges derived from the DTN model in the presence of liars (i.e., malicious nodes sharing false assessments) and defectors (i.e., good nodes that have turned malicious due to malware infection) (Section III-B). Real mobile network traces are used to verify our analysis and design.

II. 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 node during (opportunistic) contact opportunities between nodes in the DTN. When 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 by the assessment. For example, a node can assess a Bluetooth connection or an SSH session for potential Caber 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 are “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 node’s behaviors [10].

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

$$Si = \lim_{N\to\infty} \frac{sN}{N} \quad (1)$$

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

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

III. PROBLEM FORMULATION

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 piece of malicious program that disrupts the host node's normal function and has a chance of duplicating itself to other nodes during (opportunistic) inter-nodal communication in the DTN. 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. 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

$$Si = \lim_{N\to\infty} S_{N/N}$$

IV. DESIGN

In the following discussion we investigate the decision process of a node i, which has k neighbors {n1, n2… nk}, against a neighbor j, with no loss of generality, let j be n1.

A. *Homely 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 evidently the beginning of Section 3.2).

Let A = (a1, a2. . . aA) be the assessment sequence (aiis either 0 for "non-suspicious" or 1 for "suspicious") in chronological order, i.e., a1 is the oldest assessment, and A is the newest one.

Baye's theorem tells us:

$$P(Sj|A) \propto P(A|Sj) \times P(Sj) \quad (2)$$

P (Sj) encodes our prior belief on j's suspiciousness Sj; P (A|Sj) is the likelihood of observing the assessment sequence A given Sj; P (Sj |A) is the posterior probability, representing the plausibility of j having a suspiciousness of Sj given the observed assessment sequence A. Since the evidence P (A) does not involve Sj and serves as normalization factor in the computation, we omit it and write the quantitative relationship in the less cluttered proportional form1.

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

$$P(Sj|A) \propto SsAj\,(1-Sj)|A|-sA \quad (3)$$

And:

$$\arg\max_{Sj\in[0,1],A6=\emptyset} P(Sj|A) = sA/|A| \quad (4)$$

In which sA is the number of suspicious assessments in A.

Figure 1 shows the normalized posterior distributions

P (Sj |A) for assessment samples with different sizes, given by Equation 3. In each case, the ratio between suspicious and no suspicious assessments is the same ,i.e., 1:3; by Equation 4, So = 1/(1+3) = 0.25 is the maximize of P(Sj |A), which is clearly shown in Fig 1. The distribution becomes sharper with a larger sample, which accords to the intuition of the increasing certainty on the suspiciousness Sj .

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

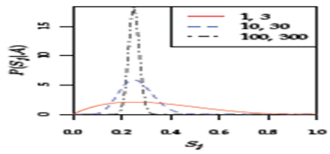


Fig. 1: The normalized posterior distribution P(Sj |A) for assessment samples with different sizes.

The two numbers for each line in the legend show the number o f suspicious and non-suspicious assessments, respectively. In each case, the ratio between suspicious and nonsuspicious assessments is 1 : 3. All distributions have a maximal value at Sj = 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 Sj.

The uncertainty over j's suspiciousness Sj (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 consider the whole Posterior suspiciousness distribution (Equation (3)) in making the cut-off decision against j. From i's perspective, after observing an assessment sequence A, the probability Pg(A) that j is good is:

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

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

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

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:

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

and

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

When Pg(A) ≥ Pe(A), the evidence collected so far (i.e., A) is favorable to j. However, when Pg(A) < Pe(A), the Evidence is unfavorable to j and suggests that j might be an evil node. i need to *decide whether to cut* j *off.*

The structure of the behavioral malware characterization model (specifically, a single threshold Le 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., A = ∅), P(Sj |A) = 1.

If LE 6= 0.5, by Equations (5) and (6),either Pg(A) < Pe(A) (if LE < 0.5) or Pg(A) > Pe(A) (if LE > 0.5). In other words, while i make 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 Le depends on the assessment mechanism itself and, as mentioned previously, if the probabilistic distributions of suspiciousness of both good and devil nodes are known, can be determined by minimizing Bayesian decision errors. If Le > 0.5, the assessment mechanism is biased towards false

positive(good nodes' actions being assessed as suspicious); if Le <0.5, the assessment mechanism is biased towards false negative (evil nodes' actions being assessed as no 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 Pg(A) and Pe(A) should be made only when A 6=∅.Alternatively, in the *maximize* approach, i uses the suspiciousness distribution's maximize (Equation (4))when making the cut-off decision against j. The justification for the maximize approach is that the suspicious for distribution's maximize is the *single most probable* estimation of j's suspiciousness given the evidence. The maximize approach precludes the prejudice problem, because the maximize is undefined when A = ∅. Similar to the distribution approach, i compares evidence that is both favorable and unfavorable to j. Evidence A is favorable to j if sA/|A| ≤ Le and is unfavorable to j if sA/|A| > Lethe maximize 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 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 consider cutting j off due to unfavorable evidence against j.The cut-off decision is made based on the risk estimation of such 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 A = (a1, . . . , aA), the next assessment aA+1(which has not been taken yet) might be either 0 (no suspicious) or 1 (suspicious). Let A′ = (A, aA+1).If aA+1 = 1, by Section 1.3 of the supplementary document, either Pg(A′) < Pg(A) < Pe(A) < Pe(A′) (the distribution approach) or sA′/|A′|= (1+sA)/(1+|A|) >sA/|A| > Le (the maximize approach): The evidence against j becomes more unfavorable.

However, if aA+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 I 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 good neighbor; in other words, it reflects i's*intrinsic* risk inclination against malware infection.

1) *Definition 1 (Look-ahead):*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.
2) *Definition 2 (λ-robustness):* At a particular pointing i's cut-off decision process against j (with assessment sequence A = (a1, . . . , aA)), i's decision of cutting j off is said to be λ-*step-ahead robust*, or simply λ-*robust*, if 1) the current evidence A is unfavorable toward j; 2) even if the next λ assessments (aA+1, . . . , aA+λ)) 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.

B. Locality 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 processes 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. These are common are assumptions in distributed trust management systems.(summarized in Section 5), which incorporate neighboring nodes' opinions in estimating a local trust value. In this model the malicious node which can transmit the malware is said to be consistent over space and time. By being consistent over space, we mean that evil nodes suspicious action 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 tithe household watch model. By being consistent over time, we mean that evil nodes cannot 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, ands beyond the scope of this paper. Instead, we propose behavioral characterization of proximity malware; further game-theoretic analysis and design could base on this foundation.

1) *CHALLENGES:* THERE ARE TWO CASES WHICH 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 false praise or a false accusation. False praises understate evil nodes' suspiciousness, while false accusations exaggerate good nodes' suspiciousness. Furthermore, alias 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 extremely, 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.

2) *Evidence:* For a pair of neighboring nodes i and j, let Ni and Nj bathe neighbors of i and j, respectively. At each encounter shares with j its assessments on the neighbor set Ni −{j}, and j shares with i its assessments on the neighbor set Nj − {i}.

Since the cut-off decision only needs to be made against a neighbor, i only considers the assessments of its own neighbors Ni ∩ (Nj − {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. 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* TE from now are included in the cut-off decision.To see that the aging window TE 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 + TE, all prior evidence expires and only those assessments after the infection are considered, which collectively testify against the node.

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

3) *Evidence Consolidations:* 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 − TE, T] provided by i's neighbors(other than one of the neighbors, say, j) in making the cut-off decision against j.

The implications are:

a) 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.

b) The liars have to share a significant amount of false evidence to sway the public's opinion on a node's suspiciousness.

c) 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. 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 B provided by a liar k must have high λB (albeit of the wrong sign) to be effective in confusing. The liar's strategy will not work because i will refuse to take B when λ|A| is small with dogmatic filtering, while λA and λB should be of different signs when λA is large (because by then, i should have a close estimation of j's true suspiciousness, and hence, λA is of the correct sign). The evidence filtering works even when the liars are the majority among i's neighbors.

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

V. SIMULATION

A. Data Sets

Design of our project can be verify using two real mobile traces. Bargain and MIT Entity.

Information is rich in raw data sets, some of which is irrelevant to our study, for example, call logs and cell tower IDs in MIT entity. So the irrelevant fields and retain the node IDs and time-stamps for each pair wise node encounter should be removed. Since the Bargain data set has only 11,230 entries spanning over three days, we repeat it another four times to make it into a data set with 56,148 entries spanning over 15 days, and thus make it comparable to the MIT entity data set in quantity. Some statistics of the processed data sets are summarized in Table 1.

TABLE 1
Data Set Statistics

	NODES	ENTRIES	TIME-SPAN	AVG.INTERVAL
BARGAIN	20	56148	8 DAYS	6 SECS
MIT Entity	48	57023	245 DAYS	185 SECS

B. SETUP

We choose Le=0.25 to be the line between Corrupt and Acceptable, without loss of generality. For each data set, we randomly pick 5 percent of the nodes to be the evil nodes and assign them with suspiciousness greater than 0.25; the rest of the nodes are good nodes and are assigned suspiciousness less than 0.25. For a particular pair wise 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.

C. Performance Metric

Comparison of performance is based on two metrics: Hit-on rate and spurious absolute rate. The categories of the "Acquaintance outlook" and "check accord" combinations are shown in Table 2. For each combination, we sum up all the decisions made by good nodes (evil nodes' check 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

$$DR = TP \div (TP+FN) * 100\%$$

and the false positive rate FPR is defined as

FPR=FP÷(FP+TN)*100%

A high Acquaintance outlook and low check accord are desirable. When a balance must be stricken between the two, one might be emphasized over the other, depending on the context.

Table 2
Acquaintance outlook and Check accord Combination

	…..gets cut-off	…. Stay connected
Corrupt acquaintance	True positive	False negative
Acceptable acquaintance	False negative	True positive

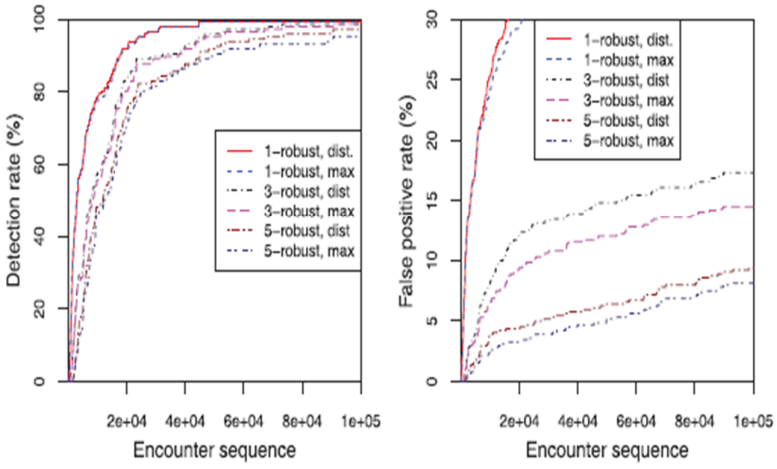
D. Results

1) View fore: Dispersal versus Distend: We compare the two alternative approaches, Dispersal versus Distend, to the view-fore strategy (see Section 3.1).The results are shown in Fig. 2.

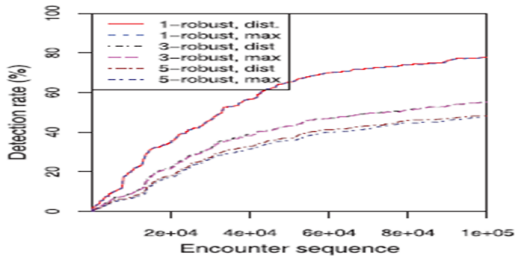
The view-fore parameter α reflects a node's infection risk inclination. In both Bargain (see Figs. 2a and 2b) and MIT entity (see Figs. 2c), the α robust cut-off strategy with a larger α corresponds to a higher Hit-on rate (in the early stage for Bargain and MIT entity) and a significantly lower spurious absolute rate (for both data sets). In Bargain, the eventual Hit-on rates for all three view-fore parameters are close to 100 percent. The difference in the eventual detection rate between

Bargain and MIT entity is attributed to the different contact patterns in these data sets: The contact pattern in Bargain is more homogeneous than that in MIT entity, 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 entity. Thus, the hit-on rate is more sensitive to the change of α in MIT entity than in Bargain.

In both data sets, the Hit-on rate and spurious absolute rate are comparable for the dispersal and distend approach, with the dispersal approach having a slightly higher hit-on rate and spurious rate. The small difference in performance, coupled with the significant reduction in computation view-fore(integration for the dispersal approach versus arithmetic operations for the distend approach), make the distended approach with a moderate α as the preferred view forest strategy. In the following sections, we show results for the distend approach with α= 3.

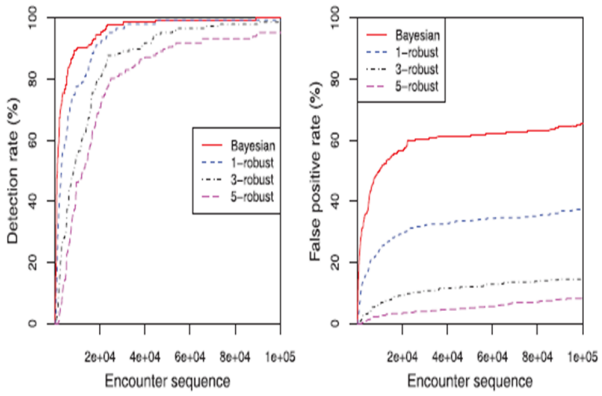


(a) Bartain (b) Bartain

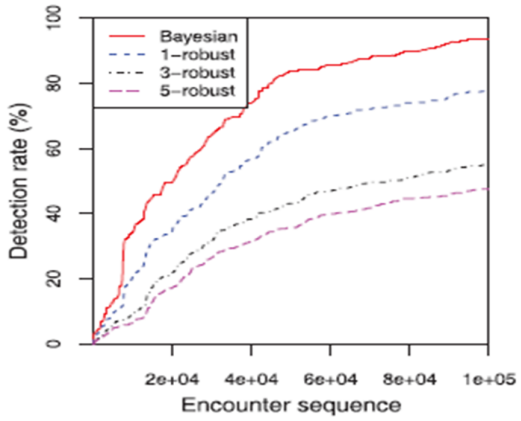


(c) MIT entity

Fig. 2. Performance comparison between the α robust strategy with the Dispersal and Distend evidence weighing approaches; α= 1; 3; and 5.



(a) Bartain (b) Bartain



(c) MIT entity

Fig. 3. Performance comparison between the vanilla Bayesian (degenerated 0-robust) cut-off strategy and the 3-robust view-fore cut-off strategy.

B. View fore

We compare Bayesian-based strategies with, and without, the view-fore extension (i.e. α 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 unfavourable to the neighbour. It can be seen as a degenerated α robust cut-off strategy with α=0. The results are shown in fig 3.

In Fig. 3, the vanilla Bayesian strategy has the highest Hit-on rate and spurious absolute rate. Both rates drop with an increasing view-fore parameter. However, the spurious absolute rate drops much faster than the hit-on rate. Indeed, for Bargain, the 1-robust and the vanilla Bayesian strategies have almost the same hit-on rate after 30,000 encounters, but there is a 30 percent difference in the spurious absolute rate. The difference in hit-on rate is more pronounced for MIT entity, but the reduction in spurious absolute rate far outweighs that of hit-on rate. For the risk-taking nodes, sacrificing a little hit on rate for a large reduction in spurious absolute rate is desirable: the view-fore parameter _ provides an effective mechanism to tune for a desirable balance.

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

C. Evidence Consolidation

We examine the uses of sharing assessments among nodes, and the effect of the proposed evidence consolidation strategies in lowering the unacceptable impact of liars on the shared evidence's quality. We compare the dogmatic filtering (with dogmatism of 0.0002, 0.02, and 2, respectively) and adaptive view-fore evidence consolidation methods with two naive evidence consolidation methods: 1) taking no indirect evidence, i.e., view fore with no evidence consolidation, and 2) taking all indirect evidence without filtering.

In this paper, 10 percent of the evil nodes play the two roles of bad-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 percent of the whole population. Under the influence of liars, the naive "all" strategy has a low hit-on rate and a high spurious absolute 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 hit-on rate and0 increases in spurious absolute rate over the baseline 3-robust view-fore strategy with no evidence filtering. Together with Fig. 3, the results indicate that the 3-robust view-fore, with either dogmatic filtering or adaptive view-fore, is comparable in hit-on rate and, even in the presence of liars, shows a significantly lower spurious absolute rate in comparison with both the Bayesian and 1-robust strategies. In Fig. 4, the eventual detection rates converge to almost 100 percent for Bargain but diverge for MIT entity. The convergence in hit-on rate is expected for a homogeneous data set like Bargain, 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 entity shows that this data set 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 α robust judgment alone. In this case, evidence consolidation helps the latter nodes in collecting enough evidence to make a α robust decision.

Two of the dogmatic filtering strategies (with a dogmatism of 0.02 and 0.0002) show almost the same performance, with the other dogmatic filtering strategy show a slight difference in comparison with other strategies. In both data sets, the adaptive view-fore 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 α robust strategies in the presence of liars.

VI. RELATED WORK

There are several common malware detection method currently in practice is pattern matching, which is a supervised data matching technique. The existing pattern matching suffers from the following drawbacks [2]

Processing overhead the lack of generality,

High false positive rate in one round of analysis make it unsuitable for DTN applications in real-time.

A. Proximity malware and existing prevention schemes

A number of studies demonstrate the severe threat of proximity malware propagation. Su et al. collected Bluetooth scanner traces and used simulations to show that malware can effectively propagate via Bluetooth [5]. Yan et al. developed a Bluetooth malware model [6]. Bose and Shin showed that malware that uses both SMS/MMS and Bluetooth can propagate faster than by messaging alone [7]. In mobile networks, one cost-effective way to route packets is via short range communication capabilities of intermittently connected smart phones [8],[9], [10].Moreover, many recent studies [11], [12], [13], based on real mobile traces, revealed that nodes' mobility showed certain social network properties.

B. Trust evaluation schemes

We base our design on the observation that trust evaluations can link past experiences with future predictions. Various frameworks [14] have been designed to model trust relationships. Three schools of thoughts emerge from studies. Su et al. [24] collected Bluetooth traces and demonstrated that malware could effectively propagate via Bluetooth with simulations. Yan et al. [25] developed a Bluetooth malware model. Bose and Shin [26] showed that Bluetooth can enhance malware propagation rate over SMS/MMS. Cheng et al. [27] analyzed malware propagation through proximity channels in social networks. Akritidis et al. [4] quantified the threat of proximity malware in wide-area wireless networks. Li et al. [28] discussed optimal malware signature distribution in heterogeneous, resource-constrained mobile networks. 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 smart phones [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 (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 versus evidence collection risk and
2) Sequential and distributed online evidence filtering.

VII. CONCLUSION

We give a general behavioral characterization of proximity malware, which allows for functional but imperfect assessments on malware presence. Under the behavioral malware characterization, and with a simple cut-off malware containment strategy, we formulate the malware detection process as a decision problem.

We analyze the risk associated with the decision and design a simple yet effective malware containment strategy, look ahead, which is distributed by nature and reflects an individual node's intrinsic trade-off between staying connected with other nodes and staying safe from malware. We consider the benefits of sharing assessments among directly connected nodes and address the challenges derived from the DTN model in the presence of liars (i.e., malicious nodes sharing false assessments) and defectors (i.e., good nodes that have turned malicious due to malware infection). The template will number citations consecutively within

brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..." Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the reference list. Use letters for table footnotes. Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols. For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

REFERENCES

[1] Trend Micro Inc. SYMBOS_CABIR.A., http://goo.gl/aHcES, 2004.
[2] http://goo.gl/iqk7, 2013.
[3] Trend Micro Inc. IOS_IKEE.A., http://goo.gl/z0j56, 2009.
[4] P. Akritidis, W. Chin, V. Lam, S. Sidiroglou, and K. Anagnostakis, "Proximity Breeds Danger: Emerging Threats in Metro-Area Wireless Networks," Proc. 16th USENIX Security Symp., 2007.
[5] A. Lee, "FBI Warns: New Malware Threat Targets Travelers, Infects via Hotel Wi-Fi," http://goo.gl/D8vNU, 2012.
[6] NFC Forum. about NFC, http://goo.gl/zSJqb, 2013.
[7] Wi-Fi Alliance. Wi-Fi Direct, http://goo.gl/fZuyE. 2013.
[8] C. Kolbitsch, P. Comparetti, C. Kruegel, E. Kirda, X. Zhou, and X. Wang, "Effective and Efficient Malware Detection at the End Host," Proc. 18th Conf. USENIX Security Symp., 2009.
[9] U. Bayer, P. Comparetti, C. Hlauschek, C. Kruegel, and E. Kirda, "Scalable, Behavior-Based Malware Clustering," Proc. 16th Ann. Network and Distributed System Security Symp. (NDSS), 2009. PENG ET AL.: BEHAVIORAL MALWARE DETECTION IN DELAY TOLERANT NETWORKS 61
[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," Proc. 21st Nat'l Conf. Artificial Intelligence (AAAI), 2006.
[11] G. Zyba, G. Voelker, M. Liljenstam, A. Me´hes, and P. Johansson, "Defending Mobile Phones from Proximity Malware," Proc. IEEE INFOCOM, 2009.
[12] F. Li, Y. Yang, and J. Wu, "CPMC: An Efficient Proximity Malware Coping Scheme in Smartphone-Based Mobile Networks," 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," Proc. 23rd Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR), 2000.
[14] P. Graham, "Better Bayesian Filtering," http://goo.gl/AgHkB, 2013.
[15] J. Zdziarski, Ending Spam: Bayesian Content Filtering and the Art of Statistical Language Classification. No Starch Press, 2005.
[16] R. Villamarı´n-Salomo´n and J. Brustoloni, "Bayesian Bot Detection Based on DNS Traffic Similarity," Proc. ACMymp. Applied Computing (SAC), 2013.
[17] J. Agosta, C. Diuk-Wasser, J. Chandrashekar, and C. Livadas, "An Adaptive Anomaly Detector for Worm Detection," Proc. Second USENIX Workshop Tackling Computer Systems Problems with Machine Learning Techniques (SYSML), 2007.
[18] S. Marti et al., "Mitigating Routing Misbehavior in Mobile Ad Hoc Networks," Proc. ACM MobiCom, 2000.
[19] P. Michiardi and R. Molva, "Core: A Collaborative Reputation Mechanism to Enforce Node Cooperation in Mobile Ad Hoc Networks," Proc. IFIP TC6/TC11 Sixth Joint Working Conf. Comm. and Multimedia Security, p. 107, 2002.
[20] S. Buchegger and J. Le Boudee, "Self-Policing Mobile Ad Hoc Networks by Reputation Systems," IEEE Comm. Magazine, vol. 43, no. 7, pp. 101-107, July 2005.
[21] R.O. Duda, P.E. Hart, and D.G. Stork, Pattern Classification, second 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, Sept. 2006.
[23] N. Eagle and A. Pentland, "CRAWDAD Data Set MIT/Reality (v. 2005-07-01)," http://goo.gl/V3YKc, July 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," Proc. Fourth ACM Workshop Recurring Malcode (WORM), 2006.
[25] G. Yan, H. Flores, L. Cuellar, N. Hengartner, S. Eidenbenz, and V. Vu, "Bluetooth Worm Propagation: Mobility Pattern Matters!," Proc. Secon ACM Symp. Information, Computer and Comm. Security (ASIACCS), 2007.
[26] A. Bose and K. Shin, "On Mobile Viruses Exploiting Messaging and Bluetooth Services," Proc. SecureComm and Workshop, 2006.
[27] S. Cheng, W. Ao, P. Chen, and K. Chen, "On Modeling Malware Propagation in Generalized Social Networks," IEEE Comm. Letters, vol. 15, no. 1, pp. 25-27, Jan. 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," Proc. IEEE Eighth Ann. Comm. Soc. Conf. Sensor, Mesh and Ad Hoc Comm. and Networks (SECON), 2011.
[29] A. Vahdat and D. Becker, "Epidemic Routing for Partially- Connected Ad Hoc Networks," technical report, Duke Univ., 2002.
[30] J. Burgess, B. Gallagher, D. Jensen, and B. Levine, "MaxProp: Routing for Vehicle-Based Disruption-Tolerant Networks," Proc. IEEE INFOCOM, 2006.
[31] V. Erramilli, M. Crovella, A. Chaintreau, and C. Diot, "Delegation Forwarding," Proc. ACM MobiHoc, 2008.
[32] W. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, "Modeling Time-Variant User Mobility in Wireless Mobile Networks," Proc. IEEE INFOCOM, 2007.
[33] E. Daly and M. Haahr, "Social Network Analysis for Information Flow in Disconnected Delay-Tolerant MANETs," IEEE Trans. Mobile Computing, vol. 8, no. 5, pp. 606-621, May 2009.

[34] S. Kamvar, M. Schlosser, and H. Garcia-Molina, "The Eigentrust Algorithm for Reputation Management in P2P Networks," Proc. ACM 12th Int'l Conf. World Wide Web (WWW), 2003.

[35] S. Buchegger and J. Boudec, "Performance Analysis of the CONFIDANT Protocol," Proc. ACM MobiHoc, 2002.

[36] A. Srinivasan, J. Teitelbaum, and J. Wu, "DRBTS: Distributed Reputation-Based Beacon Trust System," Proc. IEEE Second Int'l Symp. Dependable, Autonomic and Secure Computing (DASC), 2006.

[37] G. Eason, B. Noble, and I.N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955. (references)

[38] Mohaisen, Aziz, Omar Alrawi, and M.Larson. AMAL: Highfidelity, behavior-based automated malware analysis and classification. Verisign Labs, Tech. Rep, 2013..

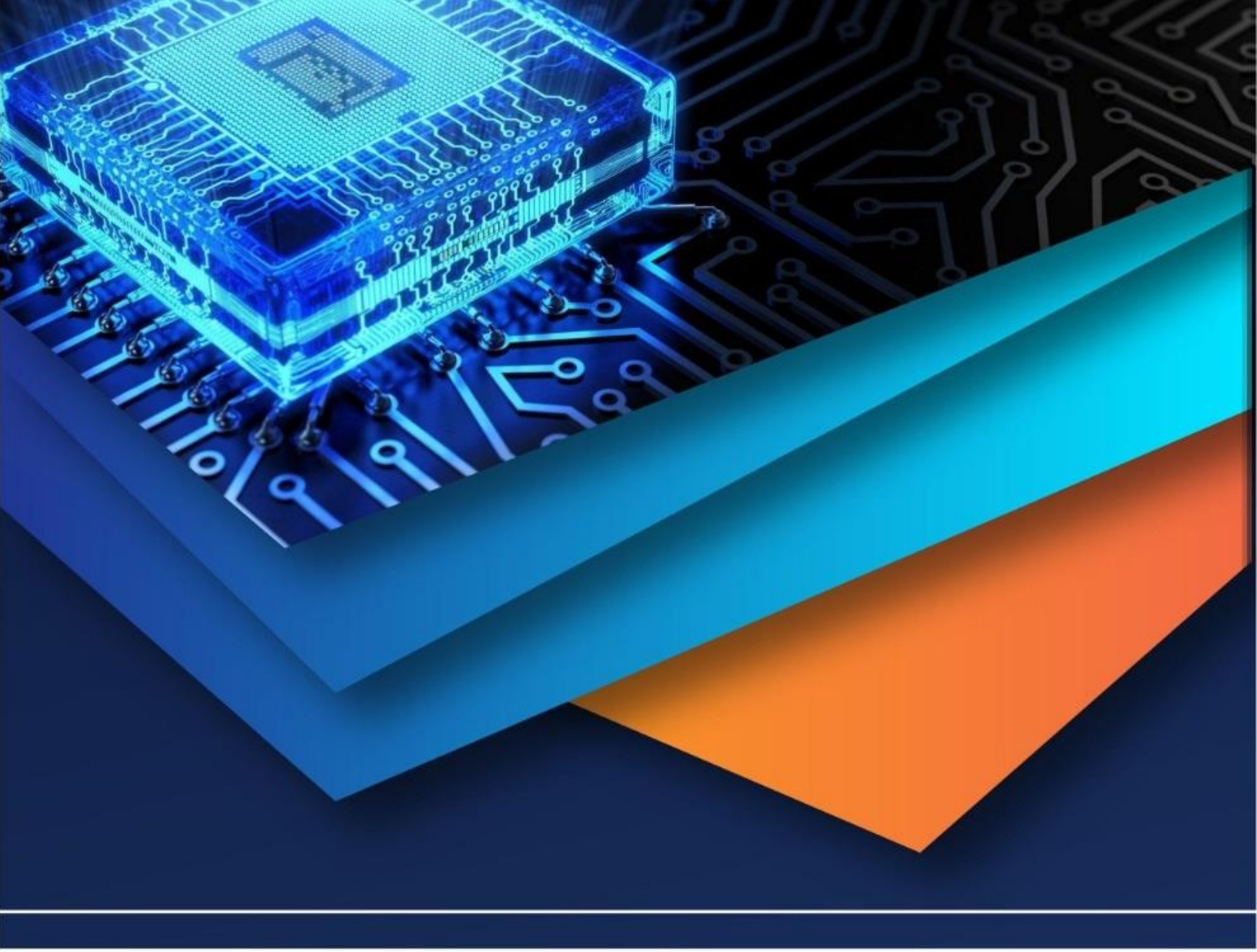
[39] I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.

[40] K. Elissa, "Title of paper if known," unpublished.

[41] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[42] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[43] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY