MALWARE DETECTION AND REMOVAL IN DELAY TOERANT NETWORKS BASED ON ITS BEHAVIOR

[1] MR. ABHINAV DUMBRE, [2] MR. DURGESH KHALANE,
[3]MR. AKSHAY CHANNAWAR

[1] Dept. of Computer Engineering, Sinhgad Institutes of Technology, Lonavala
[2] Dept. of Computer Engineering, Sinhgad Institutes of Technology, Lonavala
[3] Dept. of Computer Engineering, Sinhgad Institutes of Technology, Lonavala

abhinavdumbre@hotmail.com, durgesh.khalane@gmail.com
akshaychannawar007@gmail.com

***ABSTRACT**: Delay tolerant network (DTN) utilize the mobility of node and opportunistic contact among nodes for data communication. Due to limitation in resources such as buffer space and contact opportunity, DTNs are vulnerable to malware based attack. So the proposal introduces a novel malware detection technique in DTN. The proposed system deals with the several evidence matching and collection problems. The system also identifies the misbehaving nodes by collecting and validating their evidence using behavioral detection. The signature and Behavioral analysis of every node along with the evidence collection helps to track the accurate malware in DTN. The proposed system uses a hybrid collaborative malware detection technique to improve the detection accuracy.*

Keywords: Delay Tolerant Network, Proximity Malware, Adaptive look ahead, Dogmatic Filtering, Bayesian Model.

I. INTRODUCTION

Delay Tolerant Networking (DTN) 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. DTN works using different kind of approach than TCP/IP for packet delivery that is more resilient to disruption than TCP/IP. The basic idea behind DTN network is that endpoints aren't always continuously connected. In order to facilitate data transfer, DTN uses a store-and-forward approach across routers that are more disruption-tolerant than TCP/IP. However, the DTN approach doesn't necessarily mean that all DTN routers on a network would require large storage capacity in order to maintain end-to-end data integrity.

A. Store and forward message switching

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. Internet routers use memory chips and buffers to store (queue) incoming packets for a few milliseconds while they are waiting for their next-hop routing-table lookup and an available outgoing router port.

Intermittent connectivity

An intermittently connected network contains links that become available and unavailable during normal operation. This behavior is caused by mobility of nodes, lack of line- of-sight, physical disconnection, node failure, and transmission power among other factors. Link availability may be scheduled, probabilistic, or random based on the cause of disconnection.

B. Bundle Protocol

The DTN architecture implements store-and-forward message switching by overlaying a new transmission protocol called the bundle protocol on top of lower protocols, such as the Internet protocols.

Malware are a class of malicious software. They can be in the form of viruses, Trojans, worms, rootkits etc. Proximity malware is a piece of malicious code that attacks and alters the functionality of the node it attaches itself to and duplicates itself on interaction with other nodes, this duplication to other codes is called malware infection. Proximity malware can infect nodes opportunistically. These proximity malware exploit DTN and use its opportunistic contacts transmission for the malware propagation. Proximity malware in DTN pose security challenges that are not found in infrastructure mode. Cellular carrier would centrally monitor and look for abnormalities in infrastructure model.

In DTN there is no central monitoring. So the detection of proximity malware in DTN becomes necessary. 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.

Malware detection based on behavior is better than pattern matching especially in cases of polymorphic malware. Nave Bayes classifier has been used for behavioral detection of proximity malware in DTN settings. The chances of malware not getting detected if they turn malicious in a short time are possible with Nave Bayes.

II. CHALLENGES IN DELAY TOLERANT NETWORK

In the context of DTNs, we face a similar dilemma when trying to detect proximity malware. Hypersensitivity leads to false positives, while hyposensitivity 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 naïve Bayesian model, which has been applied in filtering email spams, detecting botnets, and designing IDSs, and address two DTN specific, malware-related, problems:

A. Insufficient evidence versus 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.

B. 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.

C. Liars

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.

D. Defectors

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.

III. PROPOSED SYSTEM

Behavioral characterization of malware is an effective alternative to pattern matching in detecting malware, especially when dealing with polymorphic or obfuscated malware. The proposed system applies behavioral characterization for malware detection. The proposed system overcomes the insufficient evidence risk and evidence collection risk. The proposed system also identifies the fake evidences by applying effective signature schemes.

In existing the Naive Bayesian model has been successfully applied in non-DTN settings, such as filtering email spams and detecting botnets.

A. Behavior - Based Detection

Behavior based detection differs from the surface scanning method in that it identifies the action performed malware rather than the binary pattern. The programs with dissimilar syntax but having same behavior are collected, thus this single behavior signature can identify various samples of malware.

The behavior detector basically consists of following components which are as follows:

- *Data Collection*: This component collects the dynamic static information are captured.
- *Interpretation*: This component converts the raw information collected by data collection module into intermediate representations.
- *Matching Algorithm*: It is used to compare the representation with the behavior signature.

B. *Signature Based detection*

A malware signature consists of the summarized malicious patterns in the malware, which can be included in an alert or a patch. If a node receives the signature before it is infected by a proximity malware, it will become immune towards the specific malware.

C. *Evidence*

The evidence collection first specifies the real form of confirmation that every node have been conceptually referring. At each time interval set of nodes exchange their evidence based results which they own assessments on their neighbors with each other.

D. *Evidence filtering*

In the evidence filtering scheme, there is an initial setup phase, during the period the system should collect a unique private key and public key which helps to the nodes to observe their neighbors.

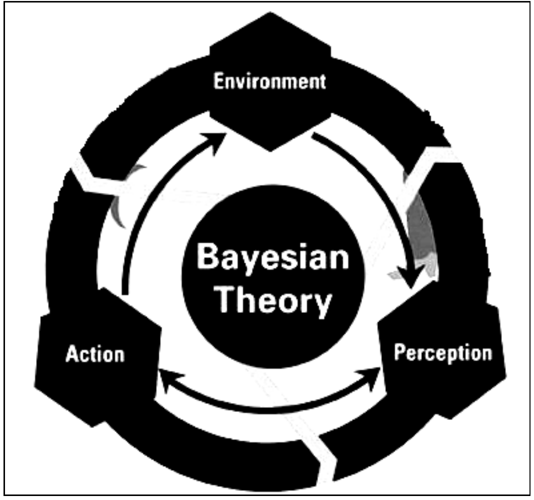

Fig 1. Bayesian Model

IV. PROPOSED TECHNIQUES

All the above mentioned challenges were solved in our proposed system using Bayesian techniques such as dogmatic filtering and adaptive look ahead, distribution.

A. *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.

With dogmatic filtering, node *i* is very conservative when its certainty about node *j*’s nature is still low. 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.

B. *Adaptive look ahead*

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

Our contributions are summarized as follows:

- We present a general behavioral characterization of proximity malware.
- Under the behavioral malware characterization, and with a simple cut-off malware elimination 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, strategy, look ahead, which naturally reflects individual nodes against malware infection.

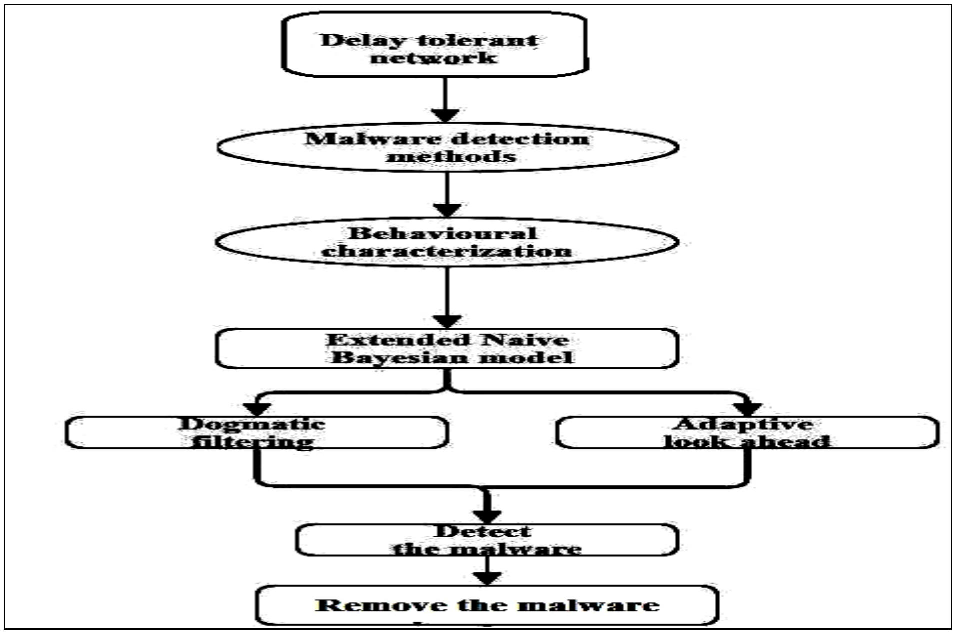

Fig 2. Malware Detection and Removal in Delay Tolerant Networks Based On Its Behavior

V. 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 nodes 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.

If node *i* has N (pairwise) encounters with its neighbors and sN of them are assessed as suspicious by the neighbors, its suspiciousness Si is defined as the evidence provided by others directly in the cut-off decision, adaptive look ahead indirectly uses the evidence by

$$Si = \lim_{N\to\infty} \frac{sN}{N}$$

By calculating suspiciousness value Si, We draw a fine line between good and evil, and judge a node. 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.*

VI. RELATED WORK

A. *Proximity malware and mitigation schemes*

Su et al. collected Bluetooth traces and demonstrated that malware could effectively propagate via Bluetooth with simulations. Yan et al. developed a Bluetooth malware model. Bose and Shin showed that Bluetooth can enhance malware propagation rate over SMS/MMS. Cheng et al. analyzed malware propagation through proximity channels in social networks. Akritidis et al. quantified the threat of proximity malware in wide-area wireless networks. Li et al. discussed optimal malware signature distribution in heterogeneous, resource-constrained mobile networks. In traditional, non-DTN, networks, Kolbitsch et al. and Bayer et al. proposed to detect malware with learned behavioral model, in terms of system call and program flow. We extend the Naïve Bayesian model, which has been applied in filtering email spams detecting botnets, and designing IDSs 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.

B. 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. While early work in mobile networks used a variety of simplistic random i.i.d. models, such as random waypoint, recent findings show that these models may not be realistic. Moreover, many recent studies, based on real mobile traces, revealed that a nodes mobility shows certain social network properties. The system may extend the malware detection work using other type of detection techniques such as game theory. The malware detection can be enhanced with the user specified rules for personalized opinion based malware filtering. 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.

VII. CONCLUSIONS

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 versus 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] Behavioral Malware Detection in Delay Tolerant Networks, Wei Peng, Student Member, IEEE, Feng Li, Member, IEEE, Xukai Zou, Member, IEEE, and Jie Wu, Fellow, IEEE

[2] Behavior Based Malware Processing in DTN Using Bayesian Model, K. Dhivya and Mr. W. R. Salem Jeyaseelan, M. Tech IT, J.J. Collegeof Engineering and Technology, Trichy, India. Assistant Professor, J.J College of Engineering and Technology, Trichy, India.

[3] A Hybrid Scheme for Malware Detection in Delay Tolerant Networks, M.Phil Scholar, Department of Computer Science, Dr.SNS Rajalakshmi College of Arts and Science, Coimbatore, India

[4] Defending Mobile Phones from Proximity Malware, Gjergji Zyba, Geoffrey M. Voelker

[5] Secure Data Retrieval for Decentralized Disruption-Tolerant MilitaryNetworks, Junbeom Hur and Kyungtae Kang, Member, IEEE, ACM

[6] Routing in Delay/Disruption Tolerant Networks: A Taxonomy, Survey and Challenges, Yue Cao and Zhili Sun,Member, IEEE

[7] SMART: A Social and Mobile Aware Routing Strategy for Disruption Tolerant Networks, Konglin Zhu, Wenzhong Li, Xiaoming Fu

[8] Optimal Forwarding in Delay Tolerant Networks with Multiple Destinations, Chandramani Singh, Anurag Kumar, Rajesh Sundaresan, Department of Electrical Communication Engineering

[9] Routing Approaches in Delay Tolerant Networks: A Survey, R. J.D'Souza National Institute of Technology Karnataka, Surathkal, India

[10] Malware Detection Based on Behavior in Delay Tolerant Network Using Support Vector Machine, K. M. Rajiha, Dr. D. C. Joy Winnie Wise, S. N. Ananthi