ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 8(34), DOI: 10.17485/ijst/2015/v8i34/85244, December 2015* ISSN (Online) : 0974-5645

Malware Detection and Elimination using Bayesian Technique and Nymble Algorithm

W. R. Salem Jeyaseelan[1*] and S. Hariharan[2]

[1]Department of Information Technology, J. J. C. E. T. Trichy - 620009, Tamil Nadu, India; salemjeyam81@gmail.com
[2]Department of CSE, T. R. P. Engineering College (SRM Group), Trichy - 621105, Tamil Nadu, India; shari1981@rediffmail.co

Abstract

Background/Objectives: DTN becomes popular because of its ability to cope with the problems in traditional infrastructural model. Like other kinds of network it is also subjected to malware attacks. **Methods/Statistical Analysis**: Pattern matching technique is so far used. But that is not secure in DTN as there is changing network topology. In this paper, a novel malware processing technique is proposed which uses Bayesian technique and Nymble algorithm. Bayesian technique is used to fabricate a secure DTN and Nymble algorithm helps in removing malware. **Findings**: Bayesian technique is used in non DTN techniques for malware processing. While using Bayesian techniques to generate a secure DTN without false positive, active attacks, passive attacks, false negative, effect of liars, effect of malware affected nodes, inadequate evidence and malware spreading. All the challenges are addressed using dogmatic filtering, adaptive look ahead, cut off strategy techniques. Nymble algorithm enhances the security of DTN. It provides ambiguity, rate limiting, subjective blacklisting and non-frame ability. **Application/Improvements**: The proposed techniques are used to identify any abnormal behavior of the nodes and complaints will be posted. Only authenticated users can post complaints. In addition to this, it provides cryptographic security to the users. The non-legitimate nodes in the network will be blocked and displayed. This improves the QoS of the DTN.

Keywords: Adaptive Look Ahead, Bayesian Technique, Dogmatic Filtering, DTN, Nymble Algorithm

1. Introduction

Delay Tolerant Network (DTN) does not needed end to end connectivity in any instance so that the possibility of connectivity from source to destination is low. In recent times it breaks the fundamental problems in networks. There are many examples of such networks in real life including military network, Vehicular Ad Hoc Networks (VANET), space networks and wildlife tracking sensor networks. There are some important properties of DTN related to the uniqueness of path and link which have a great difference from the traditional network[1,2,2,3,11,17.]

Due to consecutive modification in topology of network and movement of nodes causes, less time connectivity among the nodes. Transmission rates may be comparatively small and latency is comparatively large in some challenged networks[4,5]. The data rate may be largely symmetric with long latency of data delivery. The connectivity between the nodes fails due to the breakdown in the connectivity between nodes[1,5]. In most of the networks there is no end to end path is available. Comparatively the failure is having much more probability than the connectivity. This failure in network classified because of this network failure. These disconnections

* *Author for correspondence*

were analyzed for the traditional communication links. This couldn't make any further assumptions. The normal disconnections were happened in wireless networks due to the mobility of the source node and very less reliability of the connection[6,9].

Queuing delay dominates an end-to-end latency of data delivery. In DTN the queuing delay is come because there is more disconnection in compared to the conventional networks. Queuing delay is in seconds or typically very much less. On the contrary, queuing delay enduring for hours in traditional network. In DTN source node initiate the transmission that may be expensive because of the limited number of transmission prospect[7,7]. It states that the information has to be maintained in the message buffer for a long period. End nodes can be deployed in wide area such as disaster area and battle field. It means, the point to point data connection delay from the sender to the receiver is larger than the surviving delay between nodes which stores the data temporarily because of environmental dangers or hostile action[3,8]. In DTN, nodes are mobile and battery operated with wireless connection and thus they have limited resources. For example, if we transmit the data from source to destination through intermediates node then the data will be stored in the intermediate node until the connection to the next node is available. Therefore data buffering is confine by the limited memory space.

In delay tolerant network routing protocols are broadly categorized as deterministic routing and stochastic routing[9,19]. In deterministic routing strategy network topology and its characteristics are assumed to be known for computing the optimal route from source to destination. In this category, exact acquaintance of topology and specific protocols are developed. In stochastic routing strategy network topology and its characteristics are not known[10,22]. There are two types in stochastic routing. They are passive protocol routing and active protocol routing. In passive routing protocols moving path of node is not changed in forwarding process of message. It provides low delay and this can be used where there is less knowledge about topology and mobility of nodes. In following, there are several protocols using passive stochastic routing. Epidemic routing is kind of flooding scheme. The message is received by the node and the copy of that message is forwarded to all nodes which are nearer to source node. Thus message is spread to the intact network and each node has the same data packet. It will choose the most desirable path to reach the destination with small delay. In active routing protocols moving path of nodes are controlled in forwarding process of message active routing protocols methods are very complicated and costly compared to passive routing protocol method. Performance will be enhanced in active routing protocol method in terms of delay. Though DTN solves some of the major challenges in network[11,16]. It is also vulnerable to attack. DTN distributed nature is exploited by proximity malwares[12,23]. In order to build a secure DTN in this paper Bayesian techniques[13,21] and Nymble[6,14] algorithm have been used.

2. Related Works

The major disadvantages in DTN examined in which basic Public Key Infrastructure (PKI) is not appropriate. The Certificate Authority (CA) provides the access privilege users with their public key to another user's within the network. In connection failure stage, transmission of the encrypted message in online cannot encourage the accessing of the arbitrary receiver's public key or certificate. The CA reordered the Certificate Revocation List (CRLs) based on the frequently changed revocation key. The PKI focuses this updation. If the CRLs were missing, then the sender public key is not authenticating public key or certificate in a DTN by the receiver. Sushant and Rabin imply the use of IBC to conquer the above difficulties, the identity of the sender and associated public policies were changed in each entity of the public key. Issues with the use of IBC for security in DTNs[12,15] were investigated by Kveton and Dash. They imply the key management problem is not solved with the use of IBC in DTNs. In IBC-based DTNs solves by merging both geographical identifiers and identities.

Evan cookie and Farnam outlined the origins and structure of bots and botnets and use data from the operator community, the Internet Motion Sensor project and a honey pot experiment to illustrate the botnet problem today. Global Internet threats are undergoing a profound transformation from attacks deliberate solely to disable infrastructure to those that also target people and organizations. This terrifying new class of attacks directly impacts the day-to-day lives of millions of people and endangers businesses around the world. For example, new attacks steal personal information that can be used to damage reputations or lead to noteworthy financial losses. Current mitigation techniques focus on the symptoms of

the problem, filtering the spam, hardening web browsers or building applications that warn against phishing tricks. While tools such as these are important, it is also critical to interrupt and dismantle the infrastructure used to perpetrate the attacks.

Jeffry Bernet made some hypothesis[4,16] to reduce mathematical complexity that may occur due to evidence collection problem in DTN. The assumption made here is that each piece of the evidence either confirms or denies a single proposition rather than a disjunction[14,17]. For any domain in which the assumption is justified, the savings are available. At this point, a fallback position must be preferred and if our luck holds, the resulting system exhibits behavior interesting enough to qualify as a success. Typically, a fallback position takes the form of a uniformity assumption allowing the utilization of a non-domain-specific reasoning mechanism: for example, the numerical evaluation procedures employed in mycin and internist, the simplified statistical approach and multi valued logic. The hearsay-ii speech understanding system provides another example of a numerical evaluation and control mechanism however, it is highly domain-specific.

In non-congested ISP, the multiple access of medium introduces the threats like Distributed Denial of Service attacks or the propagation of a new worm. These attacks are very tough to identify by the backbone links. The recent emerging trend hasn't provides such kind of reliable identification parameters. This can be used with the latest network management Siaterlis. C and Maglaris. B identified and encountered attacks[18,18]. The observations are based on a network traffic analysis for a period longer that were conducted with the use of common DDoS tools in the production network of an academic ISP. Based on different types of passive measurements data can be analyzed to ISP's. Using this variety of huge identification parameters that could give network administrators insight to malicious activities passing through their networks are identified.

Tao Peng proposed a method for identify the DoS attack in the distributed network. It can enhance the effect of identification by providing network address among the nodes[20][19].They discussed about the failure in the bandwidth detection schemes which are based on monitoring the traffic volume, compared with this our scheme is very effective for highly distributed denial of service attacks[13,15,20,21]. It exploits an intrinsic feature of DDoS attacks, which makes it hard for the attacker to counter this detection scheme by changing their attack signature[10,22]. It uses a sequential nonparametric change point detection method to improve the detection accuracy without requiring a detailed model of normal and attack traffic. The detection efficiency can be enhanced with the use of distributed beliefs in multiple number user scenarios.

By using Bayesian model, Wei Peng identifies malwares in DTN. Malwares are identified based on their performance in the network[8,23]. This technique is implemented successfully by setting the filtering email spams detecting botnets.

3. Proposed System

Our proposed system focuses the proximity malware in behavioral model approach. This behavioral model the effective pattern matching is used for detecting malware. It overcomes the existing systems' drawbacks like system call and program flow. If a download limit was exceeded by the user then it tries to login to the system without perceptive passwords, tries to access the files which they are not allowed to do, tries to alter the secret files etc. then they are considered as misbehavioring peers. By using Nymble algorithm they are blocked from the system. This proposed model, the observation is performed on the malware-infected nodes' behaviors were damaged due to more number of accessing a single node. This causes the data loss in individual user, but the identification of infected nodes was performed by measuring the abnormal behaviors of those nodes.

Many challenges occur while building a DTN. Evidence consolidation problem will be there. The sender and receiver make their transaction with the identity. This can be collected and maintained. But contacting node may be malware infected node so inadequate evidence problem will be there. Malware contaminated nodes may also share false evidence about the nodes. Sometimes good nodes may be considered as malicious nodes. Similarly malicious nodes may be considered as good nodes. Liar nodes are nodes which shares false evidence about peers. By using Bayesian technique such as dogmatic filtering and adaptive look ahead all these above problems can be solved.

3 .1 System Architecture

Figure 1 shows the system architecture, which consists of four vital components. They are user peer, monitoring peer, pseudonym manager and server. The security for the system is provided by Pseudonym manager. It will afford pseudonym (secret key) to the users. During data transmission, user will use the secret key. PM will reside into the server. Secret key will be stored in encrypted format so that intruders cannot get the secret keys from the system. Users will be connected to the server. Username and passwords for the users is given by the server. Users use these to enter into the system. Unique username and passwords will be generated so only legitimate users know the correct username and password. If the server finds any misbehavior in the network, it will send complaints about misbehaving peers to the monitoring peer. Activities of the peers in the network are monitored by monitoring peer. It will obstruct the misbehaving peers in the network and preserve blacklist.

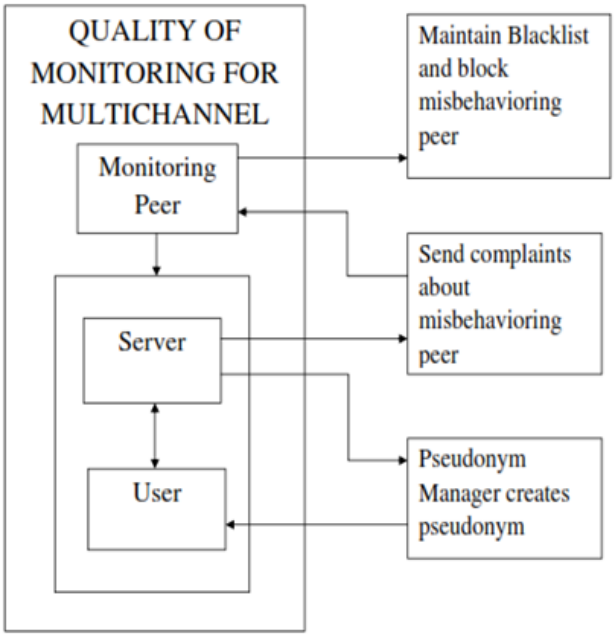

Figure 1. System architecture.

Figure 2 shows the system overview. Nymble Manager is the monitoring peer. It will examine the entire system. Nymble algorithm is used for blacklisting the users. Once the users are blocked they cannot enter into the system. Nymble manager maintains the blacklist. A DTN is created. Users are connected to the server and Pseudonym Manager. PM provides secret keys. Username and passwords is provided by the server. Data transfer takes place. If the user tries to exceed its allocated download limit, tries to access secret files which they are not allowed, tries password several times, tries to alter secret files etc. server will send complaints about that peer to the Nymble Manager. Nymble Manager will block the misbehaved peers using Nymble technique.

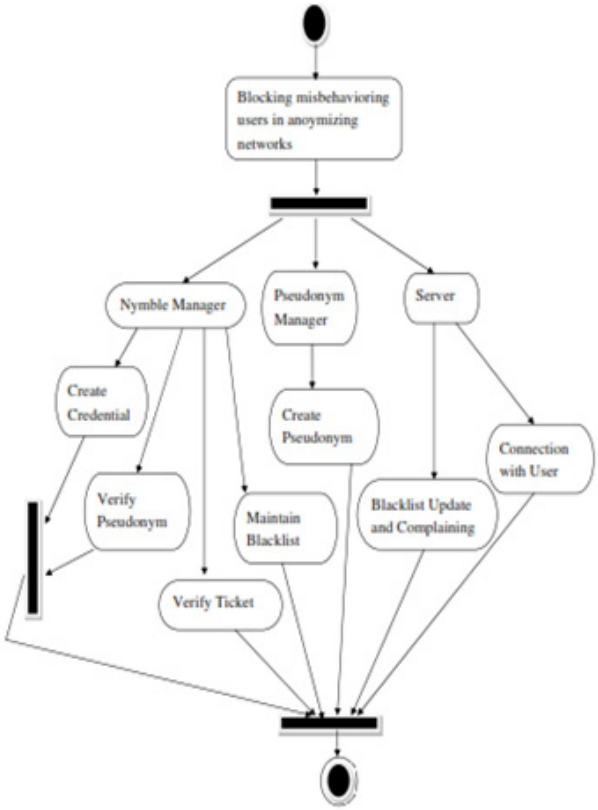

Figure 2. System overview.

3.2 Bayesian Techniques

Dogmatic filtering is used in household watch model. The observations are considered in this model which has a node assess the other nodes' verification by its own. Only based on this observation, resolution will be taken. It will accept the evidence from other nodes only if it doesn't deviate its own decision. In our proposed system, the server node will be using dogmatic filtering. It will monitor the peers. If it finds any misbehavior it will send complaints about that node to the monitoring peer. Adaptive look ahead technique is used in neighborhood model. In this model evidence from the other nodes will also be considered. Adaptive look ahead indirectly uses the evidence from other nodes. Monitoring peers uses this technique. It will acknowledge the complaints from

server node. It will make the cut off decision based on this complaint and its own assessment also. Thus both the models are used in the proposed system. By monitoring the system Denial of Service can be prevented. Malicious nodes in DTN can be identified by Dogmatic filtering and adaptive look ahead techniques. After identification they are detached by using Nymble algorithm.

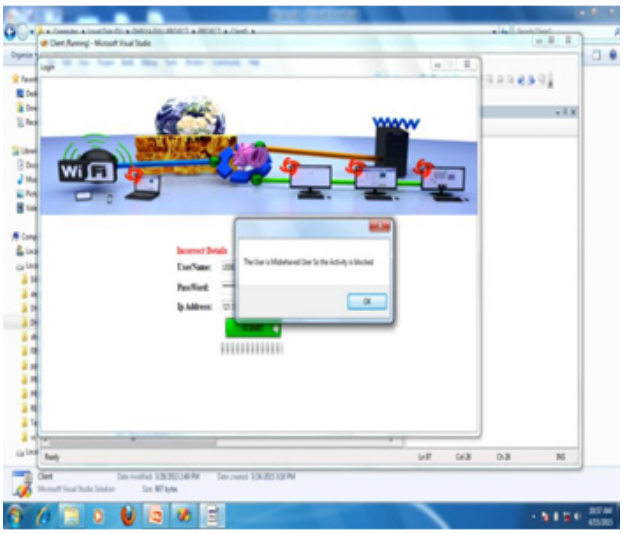

Figure 3. User misbehavior.

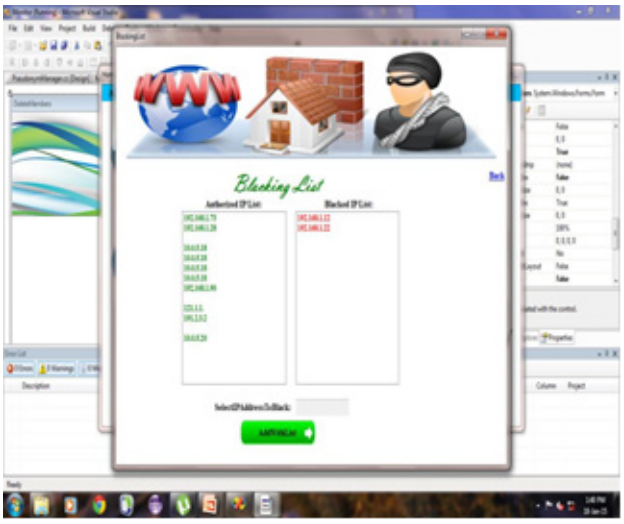

Figure 4. Authorized list and blacklist.

3.3 Nymble Algorithm

Nymble system possesses the following properties: - subjective blacklisting, fast authentication, anonymous authentication, backward un-linkable, rate limited anonymous connections, revocation audibility and also addresses the Sybil attack. The security goals of this system were anonymity, blacklisting, rate limiting and non-frameability. Blacklist ability ensures that the misbehavioring user in the home network have been blocked and the honest server intimates about the misbehaved node in the current linkability window. These windows information's were stored successfully, so the unauthorized user couldn't able to make Nymble-connect. The rate limiting assures that no user can effectively Nymble connect to it more than once within single period of time. Non-frame ability guarantees that any honest user who is legitimate according to an honest server can Nymble-connect to the server. This prevents the attacker from framing a legitimate honest user. Anonymity protects the anonymity of honest users, regardless of their legitimacy according to the server. If the connection is authorized means, server learnt more information beyond whether the user behind a Nymble-connection. Else it couldn't gather any information. If a user misbehaves in the network its IP is blocked by this algorithm.

4. Implementation

Based on operation proposed system is divided into five blocks.

4.1 Developing Pseudonym Manager

The Pseudonym Manager (PM) provides the security on authorized users. The communication between the users were controlled, resources and IP address were blocked for the unauthorized users by the PM.

4.2 Server Registration with Auditing

At the time of participation in monitoring system, the authorized user identity initiates to the PM by sending type-Auth control. By using this, the server registers the user and their U_{id} with the PM.

The server registered the user in the linkability window at most one time.

The PM listed the registered user, in which they were new to the server.The server access time duration and linkability of the node has been examine by the PM.

The unauthorized users were blocked in the network. This can be maintained by server. The authorized users being kept in private in the network.

4.3 User Registration with Auditing

In the link ability window, the identity of the user U_{id} should be registered with the PM. The user registration protocol mentions the type-basic connection to the PM of the user. The registered user has been verified by the PM. IP address of the users from the network was gathered by PM and confirms about the available. This provides the guaranty on IP address duplication. PM sense the current linkability window for the process else it has been terminated with failure. This leads for easy auditing for the user and provides much more protection against the blacklisted users.

4.4 Auditing and Filling for Complaints

The server collects the user ticket and files the complaints for the server references. During the connection establishment, the unauthorized users were blocked from accessing, in which they were in the blacklist and they were monitored.

4.5 Blacklisting a User

The current linkability window which has been connected to the blacklist user in later, the server won't allow them and mentioned them as misbehave users. The misbehavioring previous links can be made to change the unlinkable list. It provides the reverse unlinkable list and skewed blacklist, then the link connection remains unlinkability.

5. Results and Discussion

Login screen of the project is created. Only authenticated users can participate in the network. Authenticated users are given username and password during registration. Authenticated users must enter correct username and password to enter into the system. Peers are connected to the monitoring system using their IP address. Pseudonym Manager will provide secret key during registration. It provides cryptographic security. Peer details are registered. Server provides username and password, which will be used during communication. If any abnormal behavior of peer is identified complaints can be posted. Only authenticated users can post complaints. Figure 3 has shown the misbehavioring users. Client can login to the system by giving correct username and password that is provided to them during registration. Files can be sent to the server and also other clients in the Delay Tolerant Network. File transfer takes place. Files having particular size can only transferred to other systems. Thus the delay in the file transfer can be adjusted by knowing the status of transfer. Malicious user is trying to invade the system. They do not know the correct username and password. If they try the passwords several times, they will be blocked from the network.

Monitoring peer monitors the system. It provides cryptographic security to the users. If a user misbehaves in the network, they will be blocked. Both the legitimate users and blocked users are displayed. Monitoring peer can view such list. If a user misbehaves in the network, their IP is selected and they are added to the blocked list. Figure 4 shows authorized list and blacklist IP addresses of the users in the network are displayed. Misbehaved users are selected and deleted from the list.

6. Conclusion

All the peers are connected with each other in DTN. A secure DTN is built by using Nymble algorithm. If a user misbehaves in DTN they will be blocked. They cannot misuse the resources. Both active and passive attacks are prevented. Dogmatic filtering and adaptive look ahead techniques were used to build a secure DTN. While building a DTN several challenges occur. Some of them are insufficient evidence, false evidence sharing, evidence consolidation, effect of liars and defectors. All these challenges can be solved by using dogmatic filtering and adaptive look ahead techniques. This technique follows household watch model. In server peer dogmatic filtering technique is used. Adaptive look ahead follows neighborhood watch model. In this technique evidence from other nodes are accepted. Monitoring peer uses both techniques. Once the misbehavioring users in the network are identified they are blocked by using Nymble algorithm. Nymble algorithm further enhances the security of the network by providing anonymity, subjective blacklisting, revocation audibility and Non-frameability.

7. Future Enhancement

In future, the high rate DDoS attack can be detected by computing the entropy and frequency values of the incoming packets. The incoming bandwidth level exceeds the ISP allocated bandwidth. The ring level protection of FireCol is assigned only to the subscribed users of that

particular ISP. Intruders now resort to Low Rate DDoS attacks, as there are not many algorithms that successfully prevent it. Successful DDoS prevention algorithm must be equipped to prevent both High Rate and Low Rate DDoS attacks.

8. References

1. Cheng S, Ao W, Chen P, Chen K. On modeling malware propagation in generalized social networks. IEEE Trans. 2011 Jan; 15(1):25–7.
2. Daly EM, Haahr M. Social network analysis for information flow in disconnected delay-tolerant MANETs. IEEE Trans on Mobile Computing. 2009 May; 8(5):606–21.
3. Cookie E, Jahanian F, Mc Pherson D. The Zombie Roundup: Understanding, Detecting and Disrupting Botnets. IEEE Trans on Mobile Computing. 2010; 3(8):
4. Bernett JA. Computational methods for a mathematical theory of evidence. Classic works of the Dempster-Shafer Theory of Belief Function studies in Fuzziness and Soft Computing. 2008; 219:197–216.
5. Wright J, Brown I. Privacy challenges in delay tolerant and restricted route networks. IEEE Trans on Mobile Computing. 2010 May; 8(8):1132–46.
6. Tsang PP, Kapadia A. Nymble: Blocking misbehavioring users in anonymizing networks. IEEE Trans on Dependable and Secure Computing, 2011 Mar-Apr; 8(2):256–69.
7. Kapadia S, Krishnamachari B, Zhang L. Data delivery in DTN: A Survey. ACM SIGCOMM workshop on DTN: Newyork, USA; 2005. p. 252–9.
8. Peng W, Li F, Zuo X, Wu J. Behavioral malware detection in delay tolerant networks., IEEE Trans on Parallel and Distributed System. 2014 Jan; 25(1):53–63.
9. Akritidis P, Chin WY, Lam VT, Sidiroglou S, Anagnostakis K. Proximity Breeds Danger: Emerging Threats in Metro-Area Wireless Networks. Proc 16th USENIX Security Symp; 2007.
10. Agosta JM, Diuk-Wasser C, Chandrashekar J, Livandas C. An adaptive anomaly detector for worm detection. Proc Second USENIX Workshop Tackling Computer System Problems with Machine Learning Techniques (SYSML); 2007.
11. Bucheggar S, Le Boudee J. Self policing mobile adhoc networks by reputation systems. IEEE Comm Magazine. 2005 Jul; 43(7):101–7,
12. Dash D, Kveton B, Agosta J, Schooler E, Chandrashekar J, Bachrach A, Newman A. When gossip is good: Distributed probabilistic inference for detection of slow network intrusions. Proc 21st Nat'l Conf Artificial Intelligence (AAAI); 2006. p. 1115–22.
13. Berell JL, Poggi N, Gavaldia R. Adaptive distributed mechanism against flooding network attacks based on machine learning. Proc IEEE Infocom: 2002. p. 43–50.
14. Hwang K, Tanacaiwiwat S, Dave P. Proactive intrusion defense against DDoS flooding attacks, IEEE Security and Privacy Magazine. 2008.
15. Kolbitsch C, Comparetti P, Kruegel C, Kirda E, Zhou X, Wang X. Effective and efficient malware detection at the end host. Proc 18th Conf USENIX Security Symp; 2009. p. 351–66.
16. Li F, Yang Y, Wu J. CPMC: An efficient proximity malware coping scheme in smartphone-based mobile networks. Proc IEEE INFOCOM; 2010 Mar 14-19. p. 1–9.
17. Ho M, Fall K. POSTER: Delay tolerant networking for sensor networks, Proc of IEEE Workshop on Sensor Network Protocols and Applications; 2003.
18. Siaterlis C, Maglaris B . Detecting DDoS attacks with passive measurement based heuristics, Proc 9th International Symposium on Computers and Communication. 2004 Jun 28-Jul 1. p. 339–44.
19. Jain S, Patra R. Routing in DTN. Proc SIGCOMM '04; USA. 2004.
20. Peng T, Leckie C, Ramamohanrao K. Detecting distributed denial of service attacks by sharing distributed beliefs. Proc of 10th UNISEX Security Symposium; 2009.
21. Bayer U, Comparetti P, Kirda E. Scalable, behavior-based malware clustering. Proc of UNISEX Annual Technical Conference; 2004.
22. Villamarın-Salomon R, Brustoloni J. Bayesian bot detection based on DNS traffic similarity. Proc ACMymp. Applied Computing (SAC); 2013. p. 2035–41
23. Zyba G, Voelker G, Liljensta M. Defending mobile phones from proximity malware. Proc IEEE. INFOCOM; 2009.