On the Collection of Statistical Parameters on Mobile Phones for Detection of Security Attacks and Mobile Malware

M Adeel[1], L N Tokarchuk[2], M A Azam[3], S K A Khan[4]
[1]IOARP Publishing Group, London, United Kingdom.
[2]Queen Mary University of London, London, E1 4NS, United Kingdom.
[3]University of Engineering & Technology, Taxila, Pakistan.
[4]Facebook, United Kingdom.
adeel@ioarp.org[1]
l.n.tokarchuk@qmul.ac.uk[2]
m.azam@uettaxila.edu.pk[3]
kashif_ashraf@hotmail.com[4]

Abstract

Communication between mobiles using technologies like Bluetooth, MMS, SMS, Mobile and Home Broadband can increase their susceptibility to mobile security threats. In the past few years, mobile malware has emerged as one of the major threats for modern-day mobile devices. Since the first malware outbreak in 2004, hundreds of mobile viruses, worms, Trojans and spyware and over thousands of their variants have been discovered. With such a large number of ever-evolving malware, it is challenging to develop an electronic system capable of detecting such malware. In this work, the statistical data generated as a consequence of interactions on mobile devices and the mechanism of recording this data against statistical parameters on mobile devices is discussed. We also provide a detailed taxonomy of the low and high-order statistical parameters and discuss how these parameters (besides utilisation in other scenarios) can be affectively utilised in the detection of a variety of security attacks and malware types.

Keywords: Mobile and mobile P2P networks, usage footprint, simulation parameters and statistics, malware and threat detection, simulation design.

Article Number: IOARP-JCN02-004
IOARP Journal of Communication and Networks
Volume 2, Issue 1, June 2017
Appeared on Digital Library: 15 July 2017

1. Introduction

Interactions between mobiles using technologies like Bluetooth, Multimedia Messaging Service (MMS) and Short Messaging Service (SMS) technologies [1 & 2], and World Wide Web (WWW) through Mobile or Home Broadband can seriously increase mobile phone's susceptibility to mobile security threats [3]. Bluetooth technology is known for its security vulnerabilities [4, 5 & 6], while MMS and SMS servers are a hot target of Denial of Service (DoS) and Information Hiding attacks [3 & 7]. Smartphones being capable of accessing the WWW and Wireless Local Area Networks (WLAN) opens yet another threat window for resource constrained mobile devices.

In the past few years, mobile malware [6] has emerged as one of the major threats for modern day mobile networks. Since the first mobile malware outbreak in 2004, hundreds of mobile viruses, worms, trojans and spyware and thousands of their variants have been discovered [6, 8 & 9]. Malware is capable of infecting mobile devices using three common approaches: content sharing using 3G/4G mobile Internet & WLAN, Bluetooth communication directly among different peers, and MMS & SMS messaging. Damages due to malware infections can range from a simple loss of privacy and accessing unsolicited information, to the more complex issues such as system malfunction and failure. One of the most critical motives behind modern-day malware and security attacks is the service disruption on the target devices to cause economic losses [7, 3, 10 & 11]. Fig. 1 presents the results of a survey conducted as part of this research to highlight the threat intensity levels of prominent mobile malware by the security software companies. Statistics from the survey reveal that a significant number of mobile malware is considered a medium to high level security risk on their radars.

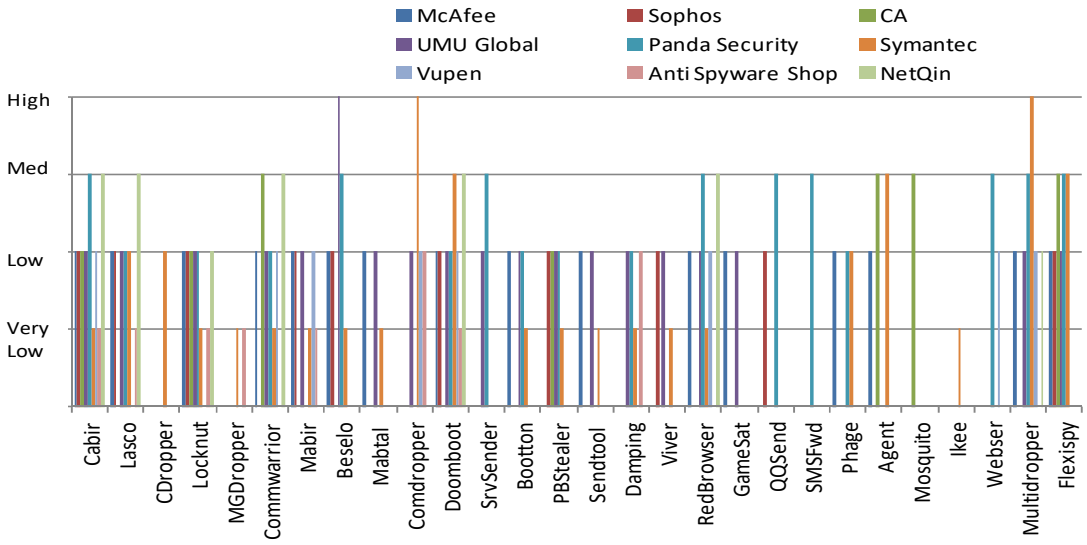

Figure 1: Survey of Malware by Antivirus Vendors

One of the critical motives behind modern-day malware and security attacks is the service disruption on the target devices to cause economic losses [7, 3, 10 & 11]. Attackers can achieve this by launching bandwidth based Denial of Service (DoS) attacks [7], or by launching attacks to cause the critical servers in the mobile networks to crash [3, 10 & 11]. Similarly, the attackers adopt strategies to overwhelm memory, processing and power resources to cause denial of critical services. Serious malware and security attacks targeting bandwidth and battery of these networks for instance, have catastrophic consequences [7, 3, 11, 10, 12 & 13] on mobile devices on intermittent or permanent basis.

With such a large and diverse number of malware and attacks, it has always been challenging to develop an efficient platform-independent electronic system capable of detecting such malware and attacks. Identification of malware and threats acts as the basis of the consequent detection, whereas the identification is strongly reliant on collection of malware or attack-specific statistics from mobile devices and (in some cases) the network.

The antivirus companies are providing with software to equip mobile devices with signature based

security software for detection of various kinds of security threats [14, 15, 16, 17, 18, 19 & 20]. Such signature based detection techniques are resource intensive as they require maintaining a signature database on resource constrained mobile devices. Malware signature databases require periodic updates which involves communication with the centralized server (for downloading updates), thus proving cost-ineffective.

Existing anomaly detection techniques for mobile malware either rely on huge repository of proximity data [5], or make use of memory intensive power signatures for abnormal behaviour detection [21], thus proving infeasible to be processed on resource constrained mobile devices.

Protection achieved through traffic filtering by mobile service providers (in the provider's network) besides being cost-ineffective and inefficient [22]. To achieve this central monitoring points are usually installed within the infrastructure of the mobile service provider to which all the contributing mobile Peer-to-Peer (P2P) nodes transmit security related statistics for a subsequent collative analysis [21, 23, 24, 25 & 26]. Cost of transmissions over the air interface for sharing such statistics can however be very high.

A review of security attack and malware detection mechanisms in general, and the statistics required by these mechanisms to perform detection in particular, reveals that the mechanism of statistics collection are prone to scalability issues, mobility constraints, and infrastructure dependent. Some of the techniques rely on cost-ineffective statistics collection, and consequently, are resource intensive.

As explained previously, one of the important motives behind mobile P2P malware is over utilization of battery power, causing throughput based flooding, incurring financial implications on victim mobiles and networks, and causing temporary or permanent loss of services. Collection of parameters pertaining to the utilization of mobile devices in terms of throughput, power utilisation and node status (i.e., if a note is active or dead), can help in detection of security threats including mobile malware.

This research, therefore, is focused on introducing statistical parameters generated as a result of different types of communication between mobile devices in mobile P2P networks. Collected on each mobile device and consolidated on the designated devices, these parameters (besides other potential uses) can be utilised in detecting security attacks including mobile malware. It is important to highlight that this paper is focused on presenting taxonomy statistical parameters and does not discuss the detection of security threats.

This paper is organised as follows. Section 2 provides a detailed discussion on acquiring statistical parameters including a discussion on the concepts of communities in context of this research, protocols for communication within communities, and collection and analysis of statistical parameters. Section 3 provides a detailed taxonomy of low-order statistical parameters extracted from the social interactions on mobile devices, and high-order statistical parameters acquired after processing the low-order statistical parameters.

2. Towards Acquiring Statistical Parameters

This section elaborates on the concept of communities in the mobile P2P networks to enable the collection of statistical parameters. This section also elaborates on the mechanism of communication of statistical parameters from the participating mobiles to the designated nodes, and different levels of processing of acquired parameters at the designated nodes.

2.1. Concept of Communities

Collection of statistics against various parameters is reliant upon communication between mobile devices in a community. A community is a collection of directly and indirectly connected mobile devices as mobile P2P networks. Mobile nodes are capable of collecting node-specific statistics such as throughput, battery power, and the number of connected neighbours. After initial processing of these statistics, each node, as demonstrated in Fig. 2, transmits the information to the designated agent node using intermediate devices. The processing of these statistics collected from throughout the community can help acquire valuable information including that on security attacks in the community.

Although beyond the scope of this paper, it is important to mention that the agent node relies on opportunistically collected statistical information acquired through periodic interactions with other mobiles in the community. Using the concept of temporal spatial centrality, various mobile P2P agents can coexist in a community. Each agent node is responsible for collating usage statistics from the participating nodes in proximity. An assessment of these statistics against abnormal throughput, battery usage, and neighbour depletion thresholds, can help identify security attacks including mobile malware.

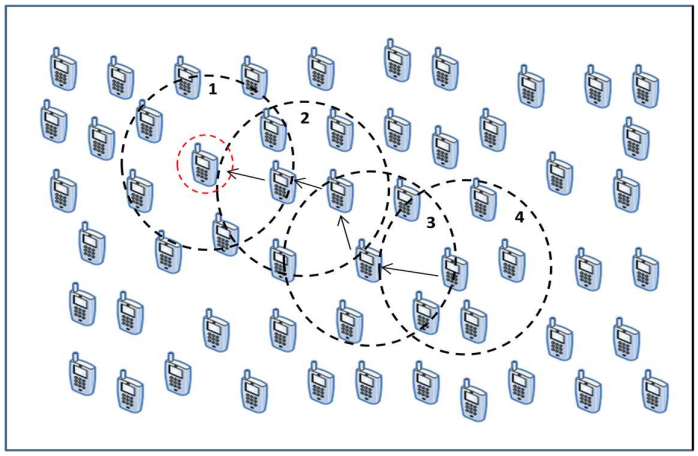

Figure 2: Interactions in Mobile P2P Communities

2.2. Communication within the Communities

Different protocols can be utilised to connect devices in the community to share statistical information with the agent nodes. The technology for connectivity however is Bluetooth communication in all cases. Ad Hoc routing protocols such as AODV [27] and DSDV [28] can be utilized for transmission, however, in Bluetooth based connectivity between resource constrained mobile P2P environments, such connectivity is not feasible. Ad Hoc based routing protocols require a topology preconfigured which is challenging to acquire in highly volatile mobile P2P network environments. Moreover, such protocols require every node to have a complete or partial knowledge of the routing which may prove resources intensives for participating mobile nodes.

Although resource intensive, opportunistic broadcast based forwarding also has highest guarantees on message delivery. Epidemic forwarding [29] which is considerably resource-friendly as compared to broadcast based forwarding has the similar level of guarantees on message delivery ratio. Social based forwarding such as Bubble Rap [30] makes use of popular nodes as efficient distributors of information. Popular nodes are primarily the nodes with a high rank in terms of previous social interactions of a node with other nodes. Lobby Index [31] is yet another social based forwarding protocol that utilizes lobby index as a node forwarding criterion where lobby index is reliant on current social interaction of a node. Lobby Influence [32] based social forwarding makes use of lobby influence as forwarding criterion and has the highest delivery ratio amongst all social based forwarding protocols. Opportunistic epidemic forwarding having the highest guarantees in terms of delivery ratio and resource friendly compared to Ad Hoc routing and opportunistic forwarding is the most recommended forwarding mechanism while Lobby Influence based social forwarding can also be utilized for transmissions between the mobile nodes.

2.3. Collection and Processing of Statistical Parameters

The process of collection of statistical parameters starts at the participating nodes where stage-1 statistical parameters are extracted in form of transmission, battery consumption during each transmission, and the neighbour connections at different instances. The acquired statistical parameters are processed initially on the participating nodes to generate rather high-order composite parameters to record throughput, battery usage and available batter, and the status of node associations over a period of measurement. Composite parameters (along with some stage-1 parameters) are subsequently transmitted opportunistically to the designated agent nodes within the community. Designated agent node processes the statistics acquired from the community in form of stage-1 parameters or stage-2 composite

parameters to generate collative community based composite parameters and further high-order statistical parameters such as stage-3 threat conditions and stage-4 family identifiers. Although the statistical parameters acquired at stage 1 can be utilised for different types of analyses such as performance, usage, and billing, this work is focused on utilisation of statistical parameters to detect security attacks on mobile devices. Although beyond the scope of this discussion, the statistical parameters have proven effective in detection of security attacks such as flooding attacks, battery depletion attacks, and various mobile malware families. Fig. 3 provides an explanation of the mechanism where a participating node is transmitting throughput, battery power, and associations related information to the designated agent node to acquire community-specific high-order statistical parameters consequently the attack identification.

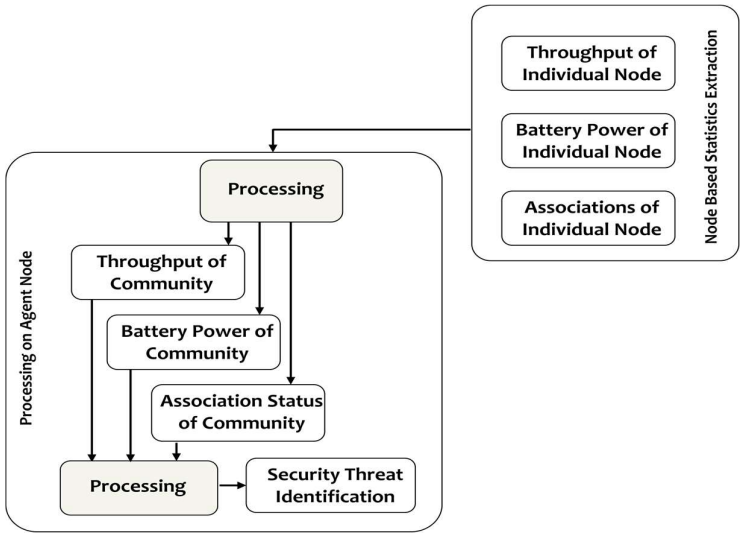

Figure 3: Processing Statistical Parameters

Before we move on to the taxonomy of statistical parameters, it is important to briefly discuss the critical high-order statistical parameters generated at the designated agent nodes by processing the stage-2 composite parameters (and some stage-1 parameters). These statistical parameters include transmit throughput, battery power, and neighbour status of the community.

2.3.1. Throughput

Overall throughput of proximity based community can be recorded as a high-order statistical parameter besides recording the throughput of individual traffic types i.e., Bluetooth, MMS, and SMS. Total throughput of the community can be calculated by adding transmitted and received throughput of all the nodes in community for Bluetooth, MMS, and SMS services.

2.3.2. Battery Power

Overall battery power of the proximity based community at any instance during simulation can also recorded as a high-order parameter. Total battery depletion in the community is acquired by adding the respective depleted batteries of nodes in the community. The total and average battery depletions can also be calculated for different Bluetooth, MMS, and SMS traffic.

2.3.3. Neighbour Status

Neighbour status in the proximity based community can also be recorded by the agent node as a high-order statistical parameter to evaluate the impact of a legitimate or malicious activity on the active, idle, and possibly dead nodes. Average active nodes in the community can be acquired through adding active neighbours of all the nodes in community by the total number of nodes.

3. Taxonomy of Statistical Parameters

After discussing the mechanism for the collection of statistical parameters and their processing, this section provides a detailed taxonomy of different levels of low-order statistical parameters extracted from the social interactions on mobile devices, and the high-order statistical parameters acquired after processing the low-order statistical parameters. Section 3.1 lists and briefly describes stage-1 parameters, section 3.2 lists stage-2 composite parameters, section 3.3 discusses threat conditions while section 3.4 provides a list of and briefly describes stage-4 family identifiers. It is important to note that in this section, a cycle (C) represents a fixed duration of time in which the statistical parameters are recorded. Ci represents the current cycle while Ci-1 represents the previous cycle. A comparison of statistical parameters extracted in different cycles yields critical information for various types of analyses.

3.1. Parameters (Parameters)

Stage-1 parameters contain the information generated as a consequence of Bluetooth, MMS, and SMS based social interactions on participating and the agent nodes. Each node records statistical information against stage-1 statistical parameters. Some of the significant stage-1 parameters have been presented in Table 1.

Table 1: Parameters (Stage-1)

Parameters	**Description**
On Participating Node	
TIME_STAMP (Numeric)	Time stamp (for an activity)
BTRY_STRT(mAh-numeric)	Battery at start of transmission
BTRY_END (mAh-numeric)	Battery at the end of transmission
NUM_NBRS (Numeric)	Number of active neighbours
NUM_BT_FILE_SNT (Numeric)	Number of Bluetooth files transmitted
NUM_MMS_SNT (Numeric)	Number of MMS messages transmitted
NUM_SMS_SNT (Numeric)	Number of SMS messages transmitted
On Agent Node	
BTRY_DEP_THSD (Numeric)	Battery depletion threshold
NBR_DEP_THSD (Numeric)	Neighbour depletion threshold
BT_THD (Numeric)	Bluetooth transmission threshold
MMS_THD (Numeric)	MMS transmission threshold
SMS_THD (Numeric)	SMS transmission threshold

3.2. Composite Parameters (Stage 2)

Either on the participating node or the designated agent node, stage-1 parameters are processed periodically to acquire stage-2 composite parameters them into composite parameters through parameter regression. These composite parameters along with some stage-1 parameters are periodically transmitted to the relevant agent nodes using opportunistic transmissions. The agent nodes collate the information acquired from all the participating nodes in form of basic stage-1 parameters and composite parameters, and generate agent-based composite parameters. Table 2 presents critical composite parameters on participating and agent nodes.

Table 2: Composite Parameters (Stage-2)

Composite Parameters	Description
On Participating Node	
CUM_BTRY_CNSD (mAh-numeric)	Node Battery consumed in duration of reporting time
CUM_BT_SNT (Numeric)	$NUM\_BT\_FILE\_SNT_{ci}$ + $NUM\_BT\_FILE\_SNT_{ci-1}$ + … + NUM_BT_FILE_SNT (cycle 1 in current reporting time)
CUM_SMS_SNT (Numeric)	$NUM\_SMS\_SNT_{ci}$ + $NUM\_SMS\_SNT_{ci-1}$ + … + NUM_SMS_SNT (cycle 1 in current reporting time)
CUM_MMS_SNT (Numeric)	$NUM\_MMS\_SNT_{ci}$ + $NUM\_MMS\_SNT_{ci-1}$ + … + NUM_MMS_SNT (cycle 1 in current reporting time)
On Agent Node	
CL_NUM_NBRS (Numeric)	Active neighbours in community
CL_BTRY_CNSD (mAh-numeric)	Battery consumed in community
CL_BT_SNT (Numeric)	Bluetooth transmissions in community
CL_MMS_SNT (Numeric)	MMS transmissions in community
CL_SMS_SNT (Numeric)	SMS transmissions in community
CL_NUM_NBRS_DEP (Numeric)	Neighbours depleted in community
CL_AVG_BTRY_CNSD (mAh-numeric)	Average battery consumed in community
CL_AVG_BT_SNT (Numeric)	Average Bluetooth files transmitted
CL_AVG_MMS_SNT (Numeric)	Average MMS files transmitted
CL_AVG_SMS_SNT (Numeric)	Average SMS files transmitted

3.3. Threat Conditions (Stage 3)

Acquired through processing various stage-1 and stage-2 parameters, stage-3 parameters represent threat conditions primarily in the community of nodes. Threat conditions are mainly Boolean and return a true value when the condition pertaining to the threat condition is true. Bluetooth flooding i.e. bf for instance is true when average of the Bluetooth transmissions by all the participating nodes in the proximity during specified reporting time, is greater or equal to the BT_THSD (i.e. Bluetooth transmission threshold). BT_THSD is based on the average Bluetooth transmissions by the same community under normal circumstances. Table 3 presents threat conditions compiled on the agent node.

Table 3: Threat Conditions (Stage-3)

Threat Conditions	Description
bf (Bluetooth Flooding)	CL_AVG_BT_SNT>=BT_THSD
mf (MMS Flooding)	CL_AVG_MMS_SNT>=MMS_ THSD
sf (SMS Flooding)	CL_AVG_SMS_SNT>=SMS_THSD
bd (battery Depletion)	CL_AVG_BTRY_CNSD>=BTRY_ DEP_THSD
nd (Neighbours Depletion)	CL_NUM_NBRS_DEPT>=NBR_DEP_ THSD

3.4. Family Identifiers (Stage 4)

Family identifiers are stage-4 parameters. Processing low-order parameters, composite parameters, and threat conditions, high-order stage-4 family identifiers can be acquired. The processing of parameters can result in identification of malware families such as Bluetooth Propagator (BP), MMS Propagator (MP) and SMS Incrementer (SI), and security attacks such as battery depletion (BD) and flooding attacks (FA). BP

malware family can be identified if condition bf && bd is true. MP Malware family can be detected if (bf && mf && bd) and/or nd is true. Similarly, if bd due to any reason is true for the specified reporting interval, it causes BD family identifier to be true while if condition bf||mf||sf is true, the FA family identifier is true.

It is important to mention that the simulation environments such as MPeersim, besides using behavioural parameters, have affectively utilized statistical parameters discussed above to conduct various types of analyses for mobile and mobile P2P network scenarios.

4. Conclusion

This paper has provided a detailed taxonomy of low-order statistical parameters extracted from the social interactions (Bluetooth, MMS, and SMS) on mobile devices, and the high-order statistical parameters acquired after processing the low-order statistical parameters. Concept of communities and the communication between nodes in the community has been discussed. A mechanism for sharing low-order parameters with the designated agent nodes along with a discussion on the processing of low-order parameters to acquire high-order statistical parameters has been provided. A brief overview of various utilisations of these statistical parameters including a thorough discussion on the use of high-order composite parameters and threat conditions in collaborative detection of security attacks and mobile malware has been provided.

References

[1] Andrew G. Miklas et al, “Exploiting Social Interactions in Mobile Systems” in proceedings of the 9th International Conference on Ubiquitous Computing, UbiComp 2007, Austria, September 2007

[2] Zhichao Zhu et al, “A Social Network Based Patching Scheme for Worm Containment in Cellular Networks” in proceedings of the IEEE INFOCOM 2009, Pennsylvania, USA, April 2009

[3] Radmilo Racic & Denys Ma, “Exploiting MMS Vulnerabilities to Stealthily Exhaust Mobile Phone’s Battery”, In proceeding of Securecomm 2006, California, USA, August 2006.

[4] Mikko Hypponen, “Mobile Malware”, Proceedings of 16th Usenex Security Symposium, Boston, USA, August 2000. [Online], Available: http://www.usenix.org/ events/sec07 /tech/hypponen.pdf.

[5] Pu Wang et al, “Understanding the Spreading Patterns of Mobile Phone Viruses”, Science, Vol. 324, No. 5930, pp. 1071-1076, May 2009.

[6] Mikko Hypponen, “ Malware Goes Mobile”, Scientific America Inc., 2006.

[7] Fabio Ricciato et al, “A review of DoS attack models for 3G cellular networks from a system-design perspective” Computer Communications, pp 551–558, November 2010

[8] McAfee Threats Report, Fourth Quarter 2010. [Online], Available: http://www. mcafee.com/resources/reports/rpquarterly-threat-q4-2010.pdf.

[9] McAfee Threat Report, Second Quarter, 2011. [Online], Available: http://www. mcafee.com/us/resources/reports/rp-quarterly-threat-q1-2011.pdf.

[10] Doomboot [Online], Available: http://www.f-secure.com/v-descs/doomboot_a.shtml.

[11] Patrick Traynor, W. Enck, “Exploiting open functionality in SMS-capable cellular networks” Journal of Computer Security, Vol. 16, pp. 713–742, 2008.

[12] CommWarrior [Online], Available: http://www.f-secure.com/v-descs/commwarrior. shtml.

[13] Cabir [Online], Available: http: //www.f-secure.com/v-descs/cabir.shtml.

[14] McAfee [Online], Available: http://www.mcafee.com.

[15] Sophos [Online], Available: www.sophos.com.

[16] CA Antivirus [Online], Available: www.ca.com.

[17] UMU Global [Online], Available: www.umuglobal.com.

[18] Panda Security [Online], Available: www.pandasecuriy.com.

[19] Symantic [Online], Available: www.symantec.com.

[20] NetQin [Online], Available: www.netqin.com.

[21] Hahnsang Kim et al, "Detecting Energy-Greedy Anomalies and Mobile Malware Variants", in proceedings of MobiSys 2008, Colorado, USA, June 2008

[22] Steven Furnell, "Handheld Hazards: The Rise of Malware on Mobile Devices", Elsevier Computer Fraud & Security , vol. 2005, no. 5, pp. 4-8, 2005.

[23] Heng, J., Wong, S., Yang, H., Lu, S., "Smartsiren: Virus detection and alert for smartphones", in proceedings of the ACM MobiSys 2007, San Juan, Puerto Rico, June 2007.

[24] Trimothy K. Buennemeyer et al, "Battery Sensing Intriusion Protection System", in proceedings of IEEE workshop on information assurance, New York , USA, June 2006.

[25] Guanhua Yan, Leticia Cuellar, "Blue-Watchdog: Detecting Bluetooth Worm Propagation in Public Areas", in proceedings of 39th Annual IEEE Conference on Dependable Systems and Networks, Lisbon, Portugal, June 2009.

[26] Aubrey-Derrick Schmidt et al, "Monitoring Smartphones for Anomaly Detection", in proceedings of Mobileware 2008, Innsbruck, Austria, February 2008.

[27] AODV [Online], Available: http://www.ietf.org/rfc/rfc3561.txt.

[28] Abdul Hadi Abd Rahman, "Performance Comparison of AODV, DSDV and I-DSDV Routing Protocols in Mobile Ad Hoc Networks", European Journal of Scientific Research, pp.566-576, 2009.

[29] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," Technical report, Duke University, April 2000.

[30] P. Hui, J. Crowcroft and E. Yoneki, "Bubble rap: social-based forwarding in delay tolerant networks", in proceedings of the ACM Symposium on Mobile ad hoc networking and computing, MobiHoc 2008, Hong Kong, May 2008.

[31] A. Korn, A. Schubert and A. Telcs, "Lobby index in networks", Physica A: Statistical Mechanics and its Applications, Vol. 388, Issue 11, pp. 2221-2226, 2009.

[32] S. K. Khan *et al.,* "Lobby Influence: Opportunistic forwarding algorithm based on human social relationships" in proceedings of IEEE International Conference on Pervasive Computing and Communications, Lugano, Switzerland, March 2012.