Spamming Botnet Detection using Neural Networks

Ickin Vural and Hein Venter

Department of Computer Science, University of Pretoria, Pretoria, South Africa

Abstract. The dramatic revolution in the way that we can share information has come about both through the Internet and through the dramatic increase in the use of mobile phones, especially in developing nations. Mobile phones are now found everywhere in the developing world. In 2002, the total number of mobile phones in use worldwide exceeded the number of landlines and these mobile devices are becoming increasingly sophisticated. For many people in developing countries their primary access point to the internet is a mobile device. Malicious software (malware) currently infects large numbers of mobile devices. Once infected, these mobile devices may be used to send spam SMSs. Mobile networks are now infected by malicious software such as Botnets. This paper studies the potential threat of Botnets based on mobile networks, and proposes the use of computational intelligence techniques to detect Botnets. We then simulate mobile Bot detection by use of a neural network.

1 Introduction

The field of computer security, which for many years focused on paradigms such as network security, information security and workstation security, is facing a paradigm shift with the ever-increasing gain in popularity of mobile devices such as smart phones and tablets. As many of the current threats to mobile devices (also known as cell phones or mobile phones) are similar to those that threaten desktop machines connected to the internet, many of the same solutions can be adapted to deal with mobile devices. Nevertheless, mobile devices present their own unique challenges such as a fragmented operating system market (such as Apple Os, Android, Windows mobile, RIM etc.), a proliferation of manufactures building devices on different standards, as well as the more limited processing and data storage capabilities of mobile devices. Security solutions have to be programmed with these limitations in mind.

This migration of computing from desktop devices to smart phones and tablets has lead to the appearance of those threats that initially only affected desktop devices. The threat that this paper addresses is the migration of spamming Botnets onto mobile devices. Botnets are now capable of infecting mobile devices and using them to send mobile spam.

The paper is structured as follows: the background section introduces the topics of spam, Botnets and Neural Networks. The following section introduces a model on combating mobile spam through Botnet detection using Neural Networks. This is followed by a description of the prototype, the actual implementation of the

Vural I. and Venter H..
Spamming Botnet Detection using Neural Networks.
DOI: 10.5220/0004089800270038
In *Proceedings of the 9th International Workshop on Security in Information Systems* (WOSIS-2012), pages 27-38
ISBN: 978-989-8565-15-0

prototype, and a section where experimental results are tabulated. The paper is then concluded.

2 Background

2.1 Spam

This section gives an overview of spam, mobile spam, spamming Botnets, Botnets on mobile devices and Neural networks. The definition of spam is discussed followed by a discussion on the different types of spam, a section on mobile spam followed by a section on the economics of spam.

2.1.1 The Definition of Spam

Unsolicited bulk mail, otherwise known as spam, is an email (electronic message) sent to a large number of email addresses, where the owners of those addresses have not asked for or consented to receive the email [1]. Spam is used to advertise a service or a product. One of the most well-known examples of spam is an unsolicited email message from an unknown or forged address advertising Viagra [2]. The lack of a universally recognized definition for spam is one of the major obstacles to creating solutions designed to minimize its harmful effects. The definition of spam could range from any unsolicited emails [3], to unsolicited commercial emails sent by any source [4], to unsolicited commercial emails from sources the recipient has never had contact with [5], to simply emails or postings transmitted in bulk quantities [6].

2.1.2 Types of Spam

Figure 1 shows the different types of spam that are commonly encountered today. Email spam is the most common form of spam and the one that most people are most familiar with. Comment spam is of the kind that inflicts the comments section of newspaper websites, where adverts are inserted in the comments section. Messaging spam, also known as spam over instant messaging (SPIM) is of the kind of spam that one would receive over an instant messenger application such as Google Talk [7]. Mobile spam, which is discussed in more detail later in this paper, is spam received on one's mobile device in the form of SMSs. Voice over internet protocol (VOIP) spam is the kind of spam that one receives through automated voice messages over a VOIP phone [8].

Mobile spam is the focus of this paper and therefore the next section is devoted to discuss mobile spam in more detail.

2.1.3 Mobile Spam

Spam is not limited to e-mail as is usually thought to be the case. Spam also exists in text messaging services (SMS). In the case of an SMS, spam can cost even more that it would when received through email. For example, assume that a user has subscribed to receive a notification via SMS when he/she receives an email.

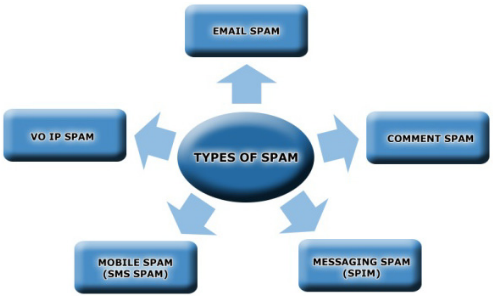

Fig. 1. Types of Spam.

Depending on the particular cellular network, the user might have to pay for every SMS received regardless if it is a spam or a valid email.

Until recently mobile networks have been relatively isolated from the Internet. Mobile networks are now well integrated with the Internet and with the proliferation of smart phones running on operating systems such as Android and Windows Mobile, threats to the Internet have started to migrate to mobile networks.

Conveniences that people access today on a desktop computer are available on mobile computing devices. Greater adaptation of these devices will encourage more users to access personal and financial data on their phones. This means that the threats to the internet such as spyware, phishing and spam are migrating to mobile devices.

Desktop operating systems are getting harder to exploit, but mobile devices have code bases that are largely unexplored, and updates to new versions with security flaws occurring frequently [9].

The increasing processing power of mobile phones, and growing features and applications included with them, make mobile phones an ideal candidate for exploitation by malware.

Malware writers are attacking commercial programs for mobile devices. A malicious program such as J2ME/RedBrowser [10], which is a Trojan horse program, pretends to access WAP web pages via SMS messages. In reality, instead of retrieving WAP pages, it sends SMS messages to premium rate numbers, thus costing the user more than intended. The next section discusses the motivation behind the sending of spam.

2.1.4 Economics of Spam

Spam makes money for those who send it, as its cost versus benefit ratio for email spam is so low. If even a small percentage of spam recipients respond to the advertised product, spammers will still make money.

Spammers generally pay nothing or very little for the sending of email spam. They accomplish this by exploiting open mail servers to send spam. The spammer need only send one email message using an open or exploited email server in a bid to reach thousands of email addresses, with the bulk of the transfer being handled by the open or exploited email server. Recipients, in turn, need to pay access costs or telephone costs in order to receive content they didn’t ask for.

ISPs have to bear the bulk of the cost for bandwidth overuse by spammers; this cost is often passed onto the consumer through increased Internet access fees or a degraded service level.

With the introduction of the "Electronic Communications and Transactions Act, 2002" unsolicited emails now have a legal definition and the sending of spam is illegal in South Africa [11]. Similar legislation exists for many other countries [12], [13]. Spammers, if identified, are liable for a fine and prosecution. Thus, spammers attempt to cover their trail to prevent identification.

The economics of SMS spam differ from email spam as network service providers have to be paid to deliver messages. Also, unlike email spam, the filters in place to filter SMS spam are not as prevalent. Thus, there is a huge incentive for SMS spammers to send SMSs from other people's devices and, thus, not pay for the SMSs being sent. Sending one million SMSs, for example, is exponentially more costly than sending one million emails. Thus, there is an incentive for SMS spammers to hijack mobile devices in a bid to send SMS spam. The following section expands on this topic by discussing Botnets.

2.2 Botnets

This section gives an overview of Botnets, the first part gives a definition of Botnets, and this is then followed by a description of mobile Botnets

2.2.1 Definition of a Botnet

A Botnet is a network that consists of a set of machines that have been taken over by a spammer using Bot software Bot software (or Bots for short) is a kind of malware that is often distributed in the form of a Trojan horse [14]. A Bot hides itself on its host machine and periodically checks for instructions from its human Botnet administrator. Botnets today are often controlled using Internet Relay Chat [15]. The owner of the computer usually has no idea that his machine has been compromised until the user's Internet connection is shut down by an ISP. Most ISPs block bulk email if they suspect it is spam. The spammers who control these Botnets typically send low volumes of mail at any one time so as not to arouse suspicions. Thus, the spam email can often be traced to an innocent individuals network address and not the spammer's actual network address. Botnets are a prized commodity on the internet and hackers are often willing to rent their hard-earned bots for money.

While the number of Botnets appears to be increasing, the number of bots in each Botnet is actually dropping. In the past Botnets with over 80 000 infected machines were common [15]. Currently Botnets with a few hundred to a few thousands infected machines are common. One reason for this decline in Bot numbers per Botnet is that smaller Botnets are more difficult to detect. Someone is more likely to notice a big Botnet and take steps to dismantle it [16]. It has also been suggested that the wider availability of broadband access makes smaller Botnets as capable as the larger Botnets of old [15].

When Procter & Gamble ran a security check of its 80,000 PCs, it found 3,000 were infected with Bots [17]. The following section elaborates on the spread of Botnets to mobile devices.

2.2.2 Botnets on Mobile Networks

Mobile devices are capable of accessing the internet through technologies such as High Speed Downlink Packet Access (HSDPA) and General Packet Radio Service (GPRS) [18]. The connection between the internet and mobile devices acts as a gateway for malware to move from the internet to mobile networks. More and more financial transactions will take place over mobile devices; this puts valuable information at risk.

The challenge for businesses and banks in the near future will be to produce secure mobile applications while ensuring ease of use at the same time [19]. The motivation for installing a Botnet that sends spam on a mobile device and installing one on a desktop differ. In the case of Botnets on desktops the motivation is to prevent the spammer's identity being revealed, in the case of mobile Botnets that is not the only motive. As was discussed in section 2.1.4 the cost of sending SMS messages is exponentially higher than the cost of sending email messages. Thus another motivation for mobile Botnets is to pass on the cost of sending the message to someone else. An implementation that would enable users to identify Botnets on their mobile devices would slow the emergence of SMS spam. The following section discusses anomaly detection; a technique that has been used in many security applications such as intrusion detection and that the authors believe can also be used to combat SMS spam.

2.3 Artificial Neural Networks

Artificial neural networks are computational methodologies that perform multifactorial analyses. Inspired by networks of biological neurons, artificial neural network models contain layers of simple computing nodes that operate as nonlinear summing devices [20]. These nodes are richly interconnected by weighted connection lines, and the weights are adjusted when data are presented to the network during a "training" process. Successful training can result in artificial neural networks that perform tasks such as predicting an output value, classifying an object, approximating a function, recognizing a pattern in multifactorial data, and completing a known pattern.

2.3.1 Learning

Learning is a process in which different events are associated with different outcomes, i.e. substantiating the cause and effect principle. This section will examine the different types of learning mechanisms.

2.3.2 Learning through Association

Learning through association is simply learning through the cause and effect principle. One example of learning by association is Boolean algebra [21]. Boolean algebra is the logical association between truth and falsehood. The logic of truth and false can be represented by AND, OR and NOT operators. Almost all arithmetic expressions can be represented by Boolean algebra and modern computers use Boolean logic for their operation. For example consider the statement a person must be at least 18 years old to drive a car, from this we can rationally deduce that someone driving a car is at least 18 years old.

Another example of association is decision tree algorithms [22]. Every node on the tree is evaluated and a decision is made as to which child node to proceed to until a solution to the problem is found. This method will fail when the parameters on the node being evaluated do not produce a sharp distinction so as to allow the algorithm to know which child node to evaluate. In cases where it is some correlation between the parameters themselves that means it is better to not just evaluate them but their combinations as well. A Neural network is designed to handle such situations. In general decision making is an attempt to find the best association between known features and known outcomes, by assigning a certain weight to each association we can select the association that is most likely.

2.3.3 Feature Identification

Identification of useful features that will be used to build associations is a pertinent issue in machine learning. A good feature is one that produces proximity to some unique region in the feature space, leading to the formation of a cluster of similar objects in the region. The greater the separation between the clusters of objects in the feature space the better the parameter. This separation is usually measured as a distance measure. A commonly used distance measure is the Euclidean distance formula. . The Euclidian distance formula, in mathematical terms, is the ordinary distance between two points that one would measure with a ruler [23].

2.3.4 Artificial Neural Network Learning

An artificial neural network models a system based on the information fed into it. If we build a good model it should be possible for the network to predict the correct output from the inputs. Real systems can be very complex and thus difficult to duplicate, neural networks can if given enough data correctly model the system that produced the data. Neural networks can be one of two types those that use supervised learning and those that use unsupervised learning [24]. Supervised learning networks learn from training data that contains many examples of possible inputs and their corresponding outputs from a real system. Thus the network attempts to mimic the training data. For unsupervised learning the training data consists of a collection of patterns with no distinction between inputs and outputs from this data the network attempts to group the patterns into different clusters. Thus the network makes a self evaluation of the possible sources of the variants in the data.

3 A Model for a Botnet Detector using Neural Networks

In this section we introduce the model for our spamming Botnet detector and explain the reasons for it being modelled in this manner as well as its advantages and disadvantages. We discuss the data that will be used to train our algorithm, the population and selection of the data. We then round this section off by explaining how the prototype would be used in a real-world situation to combat Botnet spam.

3.1 Learning with Negative and Positive Data

The detection of SMS spam is a typical classification problem where patterns, also known as signatures, need to be classified as legitimate or not. Thus, it can be regarded as a binary classification problem in which the data set, consisting of SMS messages, are classified as either spam or non spam. Neural Networks are ideal for this sort of classification problem as we can train the spam filter on legitimate (positive) and non-legitimate (negative - spam) messages. From this data, the neural network should be able to deduce whether messages are spam or not. The Neural Network Botnet Detector (NNBD) was implemented using the .Net framework.

3.2 Data Classification

As discussed in section 2.3.4 supervised learning networks learn from training data that contains many examples of possible inputs and their corresponding outputs from a real system. It is thus possible to train a neural network on a data set consisting of only spam SMSs. The neural network trained on this data should be able to accurately identify SMSs as spam or ham. To increase the accuracy of the neural network it can be trained on an additional data set. This additional data set consisting of ham messages can be used to improve the accuracy of the implementation and reduce the number of false positives (valid SMSs incorrectly classified as spam). Section 4.4 further describes the data selection process.

3.3 Remote Analyses Vs. Analysis on the Device

There are two ways in which one could analyse a mobile device user's SMS data. The first possibility would be to analyse the SMSs sent on the ISP's servers and use this to build a profile of the user's SMS sending behaviour. There are several drawbacks to this approach. Firstly, there are privacy implications of analysing SMS data on an ISPs server, but even if these concerns were addressed, the classification algorithm would have to determine whether or not a message is valid or invalid without the user's input. Thus, it would not be able to learn based on user feedback. The second solution is to implement the detection algorithm on the device, thus, taking care of the privacy implications as well as allowing for user feedback in the learning process. The major disadvantage of this approach is that the prototype needs to be installed on a mobile device which has limited processing power and storage space especially when compared to an ISP's server. Thus, the prototype had to be programmed to use minimum storage and be optimised to make use of as little memory as possible.

3.4 Flow Diagram for Botnet Detector

Figure 2 visualizes how the Botnet detector is designed to work. The mobile user enters a text message and sends it to a recipient. This message is intercepted and certain message characteristics such as the number of capital letters (the full list of characteristics is defined in section 4.1) are also extracted for analysis by the Botnet

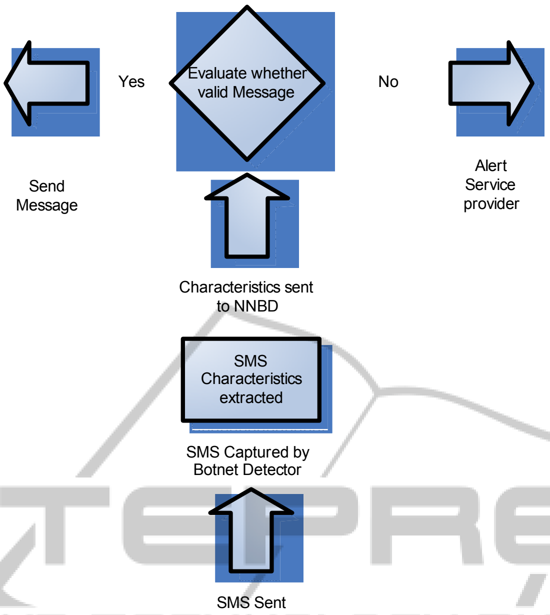
Yes
Evaluate whether valid Message
No
Send Message
Alert Service provider
Characteristics sent to NNBD
SMS Characteristics extracted
SMS Captured by Botnet Detector
SMS Sent

Fig. 2. Model of the Botnet detector.

Detector before the message is sent (the neural network should not send out messages identified as spam messages). The characteristics are sent to the neural network which then determines whether the message is valid or not by inputting the SMS message into the neural network (this is explained further in the next section). If the NNBD can determine that the message is valid, the message is sent onwards. Else the network provider will be alerted by the NNBD so that the service provider can investigate the malware that is sending these spam messages and remove it from the mobile device

The premise behind the NNBD implementation is that the authors believe that a spamming Botnet that has installed itself on a user's mobile device, and is sending out spam SMSs without the knowledge of the mobile devices user, can be detected by the NNBD. Thus, if these spam messages are blocked, the NNBD would have succeeded in preventing the sending of spam as well as saving the mobile device owner the cost of the spam SMSs being sent. Additionally, the NNBD would alert the mobile device user to the presence of Botnets on their device so that the malware, that has installed itself on their device, can be removed.

4 Implementation of Mobile Botnet Detector – A Prototype

This section describes how the prototype was implemented. The authors first discuss the message characteristics chosen to extract and train the Neural Network module with. Next the section describes the process of training the neural network. This is followed by a section documenting the testing of the neural networks accuracy as well as a section describing how the data was selected for this implementation.

4.1 Message Signatures

The prototype creates a signature (pattern) for each message sent by the mobile device. The signature consists of the following characteristics that are analysed by the neural network to determine the validity of the message:

- Does the SMS message contain links?
- Does the SMS message contain telephone numbers?
- The number of words in a message
- The ratio of punctuation characters to words
- The ratio of links to words in the message
- The ratio of capital letters to words in the message
- The ratio of misspelt words in the SMS message
- Bayesian spam probability

The specific characteristics mentioned above were chosen by the authors to define the message signature as they allow the implementation to build a profile of the user's sending behaviour. The characteristics chosen are simple to capture, yet indicative of sending behaviour. Use of punctuation, capital letters and their ratios to word in a message may reveal much about an SMS message.

The majority of spam emails contain a link to a URL, thus it makes sense to mark the presence of URLs in the SMS message (these links might lead to a website which sells a product that the spammer is attempting to advertise). The presence of telephone numbers is also be a useful bit of information to mark as the spammer might include a telephone number for the recipient of the spam message to call in order to enquire about the product or service advertised, quite possibly this call may be charged at a premium rate. The ratio of misspelt words is also a useful characteristic to monitor as spammers will often hide words. They will often send an SMS with phrases like "v1agr@a" instead of Viagra in order to thwart a spam filter. The Bayesian spam probability, which calculates the probability of a message being spam, will become the dominant feature if the other characteristics fail to identify the message as being spam. To calculate The Bayesian spam probability examines all string tokens within a message and calculates whether the token appears often in spam message. This Bayesian probability is used to calculate the spam probability of the message. Additional characteristics may be added in future to the prototype to increase the accuracy of the implementation. This would enable us to build a better profile of the users messages.

4.2 Training the Neural Network

The training of our neural network is a three stage approach. First, we add the messages we are going to train the neural network on, to a spam list and a ham list (a ham list is a list of valid SMS messages). Once we have done this, we can then generate the message signatures for each message, which we save to a list. Finally, we pass the message signatures from our saved list into the training procedure of our neural network, which, in turn, causes the network to learn what message signatures represent spam.

4.3 Testing the Neural Network

To test the neural network we run the network against a collection of spam and ham messages extracted from one of the authors' mobile phones. The neural network looks at the message signatures as identified previously in section 4.1 and generate a numerical statistic for each signature between 0 and 1. The NNBD then takes the collection of statistics generated, and feed them as input vectors into our neural network. The neural network will have one output, which will be a probability between {0, 1} on whether the SMS is spam or ham..

4.4 Data Selection

The data that was used to train the NNBD implementation was selected by using SMS messages collected by the SMS Spam Collection project [25]. In total 4815 of these valid SMSs were used to create a white list of legitimate SMS messages. The second set, i.e. the black list of invalid SMSs, totalled 746 (including blacklisted URLs). In the following section we tabulate the experimental results.

5 Tabulation of Results

The results of the experiment are tabulated as follows and show the accuracy in detecting spam SMSs in Table 1. For this paper the authors compared the results from the NNBD against a previous prototype they had built that used a Bayesian filter to detect spamming Botnets.

Table 1. Results.

	Valid Message	Invalid Message	Total error
Neural Network	100%	95%	2.5%
Bayesian Filter	100%	90%	5%

As can be seen in Table 1, the NNBD has an accuracy of 100% in identifying valid messages, and an accuracy of 95% in identifying spam messages giving us a total error of 2.5%. The total error is calculated by determining how many of the spam and ham messages (40 in total, 50% spam and 50% ham) were incorrectly identified. This compares well to the prototype that uses a Bayesian probability to detect spam as shown in the second line of Table 1.

6 Discussion

Spam cannot be eliminated solely with technological solutions. To reduce spam, people must ideally stop responding to spam messages by actually purchasing items advertised in spam messages. Never the less, technical solutions play a strong role in combating spam. By allowing less spam to get through, we can reduce the incentive for spammers to spam and increase the cost of sending spam for spammers, thereby reducing spam.

The authors believe that this neural network, when implemented, could reduce spam significantly. This NNBD could be implemented either on a mobile device or on a network provider's server. The advantage of installing it on a mobile device is that the user can be prompted to confirm whether a particular message is indeed spam or not if the neural network is unsure. The disadvantage of installing the NNBD on a user's mobile device is the degradation of performance that could be experienced when the NNBD calculates its spam probability. The advantage of situating the NNBD on a network service provider's server is that processing can be done by dedicated servers. The disadvantage, of course, is that the user cannot be prompted for feedback and confirmation, which might result in certain messages being incorrectly classified.

7 Conclusions and Further Work

Spamming Botnets have the potential to become more common with the increasing processing power of mobile phones make mobile phones more attractive for exploitation by malware. The aim of this research is to provide a tool that is not only capable of identifying spam SMSs being sent from a user's mobile device, but also to allow the spam filter to learn new spam features as spammers continuously change the spam features to confuse spam filters. The NNBD is has been shown to be capable of correctly identifying 95% of all Spam messages with a zero false positive rate. This shows that the NNBD can be used to detect spamming mobile Botnets that have infected mobile devices.

The authors believe that over time, with more spam messages added to the black list, the accuracy of this implementation would improve. The authors hope to apply these ideas out on an Instant Messaging platform as well as on SMS messaging.

References

1. Internet Service Providers' Association, 2008. 'What is Spam?' Available: http://www.ispa.org.za/spam/whatisspam.shtml. [April 2009]
2. Spam-site "Samples of Spam" http://www.spam-site.com/spam-sample.shtml [September 2011]
3. B. G. Kutais, "Spam and Internet Privacy", 'Journal of High Technology Law Suffolk University Law School'
4. Consumer fraud reporting, "Spam Emails and Spamming", http://www.consumerfraudreporting.org/spam.php [September 2011]
5. Federal Communication Commissio, "Spam: Unwanted Text Messages and Email", http://www.fcc.gov/guides/spam-unwanted-text-messages-and-email [September 2011]
6. Spamhaus "The Definition of Spam", http://www.spamhaus.org/definition.html [September 2011]
7. Earth Web. "Think Spam is tough? Try Fighting Spim" http://itmanagement.earthweb.com/secu/article.php/3365931 [September 2011]

8. R. Dantu and P. Kolan. Detecting Spam in VoIP Networks. In Proceedings of USENIX Steps to Reducing Unwanted Traffic on the Internet Workshop (SRUTI), July 2005.
9. Security Vision from McFee Avert Labs, 2007 The Future of Security McFee, 2010 Available: http://vil.mcafeesecurity.com/vil/content/v_138726.htm
10. McFee, 2010 Available: http://vil.mcafeesecurity.com/vil/content/v_138726.htm
11. Acts Online, 2002. Electronic Communications and Transactions Act ,2002. Available: http://www.acts.co.za/ect_act/. [April 2009]
12. Australian Government Department of Broad Band Communications and the Digital Economy. "Spam". http://www.dbcde.gov.au/online_safety_and_security/spam [September 2011]
13. Industry Canada. "Government of Canada Introduces Anti-Spam Legislation" http://www.ic.gc.ca/eic/site/ecic-ceac.nsf/eng/gv00521.html#Q1 [September 2011]
14. Seach Security.com "botnet (zombie army)". http://searchsecurity.techtarget.com/ definition/botnet [September 2011]
15. E. Cooke, F. Jahanian, and D. McPherson. The zombie roundup: Understanding, detecting, and disrupting botnets. In USENIX SRUTI Workshop, pages 39–44,2005.
16. Ryan Vogt, John Aycock, and Michael J. Jacobson, Jr. "Army of Botnets", Proceedings of the 2007 Network and Distributed System Security Symposium, pp. 111–123, 2007
17. The Economist "Big brother bosses" September 11 2009 Available: http:// www.economist.com/businessfinance/displaystory.cfm?story_id=14413380 [September 2009].
18. Sumit Kasera and Nishit Narang. , 2005, 3G Mobile Networks. Architecture, Protocols and Procedure. Tata MCGraw-Hill Publishing Company, limited edition.
19. Emerging Cyber Threats Report for 2009, Georgia Tech Information Security Center, October 15, 2008
20. Judith E. Dayhoff Ph.D,, James M. DeLeo Supplement: Conference on Prognostic Factors and Staging in Cancer Management: Contributions of Artificial Neural Networks and Other Statistical Methods, Volume 91, Issue Supplement 8, pages 1615–1635, 15 April 2001
21. NIGEL P. COOK, Introductory Digital Electronics, Prentice Hall (1997)
22. Safavian, SR., and D.Langrebe, A survey of Decision Tree Classifier Methodology, IEEE Transactions on Systems, Man and Cybernetics (1991)
23. Elena Deza & Michel Marie Deza (2009) Encyclopedia of Distances, page 94, Springer.
24. Anthony Zaknich, Artificial Neural Networks :An Introductionary Course available 'http:// www.maths.uwaedu.au/~rkealley/ann_all/index.html
25. José María Gómez Hidalgo , Guillermo Cajigas Bringas , Enrique Puertas Sánz , Francisco Carrero García, Content based SMS spam filtering, Proceedings of the 2006 ACM symposium on Document engineering, October 10-13, 2006, Amsterdam, The Netherlands [doi>10.1145/1166160.1166191]