ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 1, January 2016

A Survey of Botnet Detection Techniques and Research Challenges

Ruchi Dhole, Prof. Shobha Lolge

PG Student, Dept. of Computer Engineering, Lokmanya Tilak College of Engineering, Navi Mumbai, Mumbai University, Maharashtra, India

Assistant Professor, Dept. of Computer Engineering, Lokmanya Tilak College of Engineering, Navi Mumbai, Mumbai University, Maharashtra, India

ABSTRACT: Recently Botnet has been recognized as the most significant security threats of the Internet. These are emerging as the most serious threat against cyber-security.They provide a distributed platform for several illegal activities one of them are launching distributed denial of service attacks, phishing, and click fraud. This surveyclarifies botnet phenomenon andcoversfour classes of botnet detection techniques like DNS-based, signature-based, anomaly-based, and mining-base. The defining characteristic of botnets is the use of command and control channels through which they can be updated and directed. Latest attacks are increasingly complex, and utilize many strategies in order to perform their intended malicious/hazardous task. Attackers have developed the ability of controlling vast area of infected hosts, characterized by complex executable command set, each involved part in cooperative and coordinated attacks.

KEYWORDS: Botnet; Botnet Detection, Cyber-security, Command And Control Channel, Centralized,Decentralized, Web-Based Botnet.

I. INTRODUCTION

A Botnet can be considered as a network of bots under the remote command of a botmaster. These bots are controlled to perform illicit activities. They pose a significant threat against cyber security. They provide a distributed platform for various cybercrimes such as distribute denial of service (DDOS), malware dissemination, click fraud and phishing. All users of computers are at high risk because we all browse the same internet. Every individual should be aware of social networking attacks.

Companies and governments suffer most damage from botnet attacks. The results of these attacks can be dangerous, costing the companies significant manpower, cost and clean. DDOS attacks can disrupt the communications and infected source code can halt the critical servers. Botnets have become much more sophisticated and dangerous now a day. Few formal studies have examined the botnet issues and very little is known about the malicious behaviour of botnets. This research aims at finding out the latest and advanced techniques of botnet detection.

1.1Characterization of Botnet:

Botnet structures have three parts: Bot, Botmaster and command& control channel shown in Fig. 1.

A Bot, emanating from the term ‘Robot’ which is also called as Zombie. It is a new type of malware installedinto a compromised computer which can be controlledremotely by Botmaster for carrying out some orders throughthe received commands. After the Bot code has been installedinto the seized computers, the computer becomes a Bot orZombie. Contrary to existing malware such as virus andworm which their main activities focus on attacking theinfecting host, Bots can receive commands from Botmasterand are used in distributed attack platform.Botmaster is known as Botherder that is aperson or a group of person which control and manageremote Bots and Botnet [5]. The difference between theBotnet and the virus is whether it can be controlled or not [5].

Botnet use the command and control channel to control theBotnet. The most essential part of a Botnet is the so calledcommand and control infrastructure (C&C). Thisinfrastructure consists of the Bots and a control entity that canbe

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 1, January 2016

either centralized or distributed. The C&C infrastructuretypically serves as the only way to control Bots within theBotnet. The Bots are required to maintain a stable connection within this infrastructure in order to operate efficiently [5].

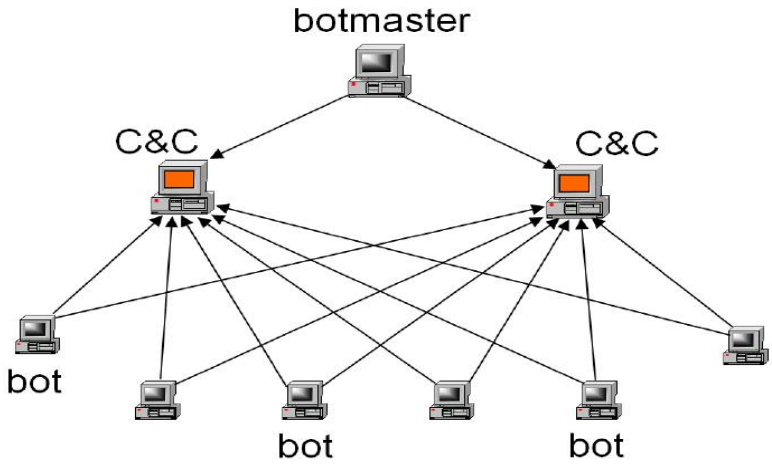

Fig.1.Botnet Structures

1.2Botnet Life-cycle:

A botnet has been created and maintained in five different phases including: initial infection, secondary injection, connection, malicious command and control, update and maintenance. This life-cycle is depicted in Fig. 2

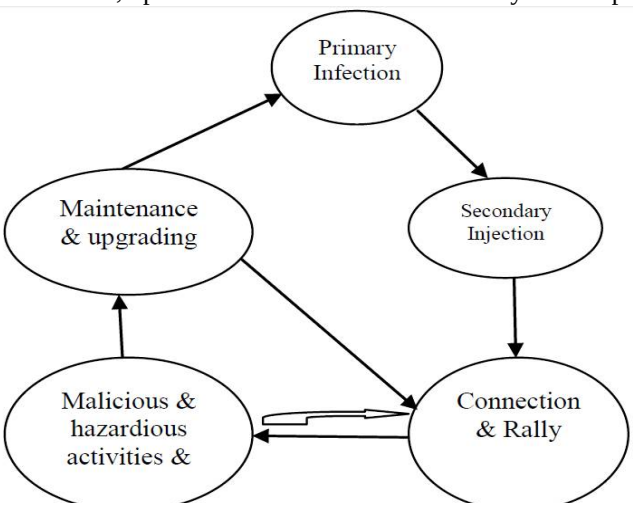

Fig.2.Botnet Life Cycle

Primary infection is the first phase of botnet life cycle, where in a host is infected and becomes a possible bot. This phase is characterized by an expected computer infection procedure, which may be passed away in different behaviour as a typical virus infection would be, for instance, through unwanted downloads of malware from websites, infected removable disks, infected files attached to email messages, etc.

Second phase is defined the secondary injection, Firstly, the first phase be successfully finished. In this phase, the infected host runs a program that searches for malware binaries in a given network database. When downloaded and executed, these binaries make the host behave as a zombie). Downloading bot binaries is usually performed by FTP, HTTP or P2P protocols.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 1, January 2016

Third phase is scheduled every time the host is restarted to ensure the botmaster that the bot is taking element in the botnet and is capable to receive commands to perform malicious activities. After establishing the command and control channel the bot waits for commands to perform malicious activities. Fourth phase is ready to perform an attack. Malicious activities may be as broad ranging as information theft, performing DDoS attacks, extortion, monitoring network traffic, spreading malware, stealing computer resources, and unprotected computers, identity theft, phishing, spamming, manipulating games and surveys, etc. Fifth phase Maintenance and up gradating is the most important phase of botnet life cycle. Maintenance is necessary if the botmaster wants to keep his army of zombies. It may be necessary to update codes for many reasons, including evading detection techniques, adding new features or migrating to another C&C. This phase is also usually measured a susceptible step. As the botmaster intends to broadcast updates as soon as possible, some behavioural patterns of the stations belonging to the network may emerge and make the botnet detectable [9].

- ✓ Bot-herder configures initial bot parameters.
- ✓ Registers a DDNS.
- ✓ Register a static IP.
- ✓ Bot-herder starts infecting victim machines either directly through network or indirectly through user interaction.
- ✓ Bots spread.
- ✓ Bot joins the Botnet through C&C server.
- ✓ Bots are used for some activity (DDoS, Identity Theft etc.)
- ✓ Bots are updated through their Bot-operator which issues update commands.

These sequential steps make botnets more unavoidable and difficult to capture. Subsequently they become more successful and devastating. Changes in behaviour are typically observed, for instance, in DNS queries and file sharing, among other areas. After bots are updated, they must establish new connections with the C&C infrastructure. Botnets, networks of malware infected machines controlled by an adversary, are the root cause of a large number of Internet security problems.

The rest of the paper is as follows. Section 2 briefs the considered different schemes based on Botnetsfor the analysis. Section 3 represents the Botnet Architectures considering different models. Section 4 represents the Botnet Detection Technique and the paper concludes with section 5.

II. LITERATURE SURVEY

From the past decade, the search for highly effective and efficient techniques of detecting botnet is a dynamic focus of research. The comprehensive works of Piyush [1], Kim [2] and Dittrich[3] provide some of the most influential surveys on the security in peer to peer networks and attacks done on this highly vulnerable network. The extensive work of Daniel [2] also outstands to describe the functionality peer to peer system and botnet attacks and detection techniques in those systems. Mohini [6], David [9] done a great work with botnet detection by getting information of different systems and analyse it. Finally, the recent study of [4], [8] and [12] gives the enhanced work on machine lerning type of detection technique and give us the most promising methods for botnet detection with greater accuracy.

III. BOTNET ARCHITECTURES

Communication is the next major problem of the botnet attacker. Most attackers would communicate to bots but do not interact to the exposed bots. The Command and Control mechanism creates an interface between the bots, C&C servers and the Botmasters to transmit data among them. According this channel, there are three different Botnet architectures: the Centralized model, the Decentralized model and Hybrid model.

3.1 Centralized Model:

A centralized topology is characterized by a central point forwarding messages among clients or publishing them [1]. The master selects a host to be the contacting point of all Bots. It can be a compromised machine or a legitimate provider for public service. When the victim is infected it will ‘connect’ to the C&C server and then will wait or check

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 1, January 2016

for pending commands from the Botmaster [5]. Fig. 3 shows the basic communication architecture for a Centralized model. They are characterized by:

- ✓ Low latency due to the small number of hops required to transmit the orders from the botmaster.
- ✓ Direct connection to order distribution nodes, which would compromise the security of the network in case of accidental detection of a node.
- ✓ Implemented using different communication protocols, but most typically the IRC and HTPP.

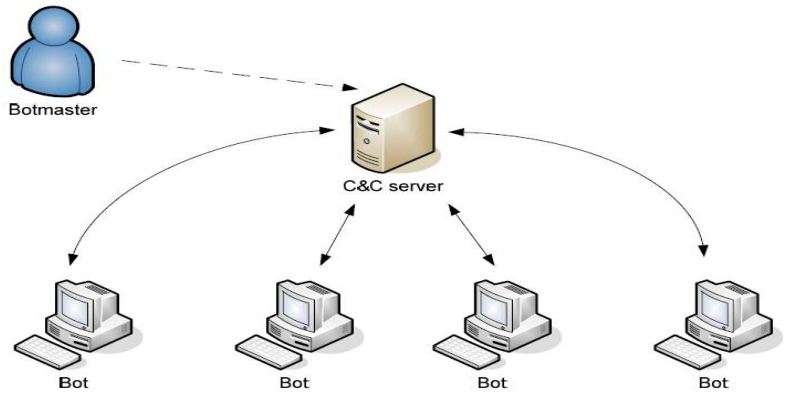

Fig.3.Centralized Model

3.2 Decentralized Model (Peer to Peer Model):

The decentralized form allows the bots to act autonomously. Bots can establish connections with other bots. They also send requests for additional commands to the botnet. Here, the bots are loosely coupled. These links are useful for communicating with other bots within botnets. These botnets are also known as peer-to-peer botnets [1].

The information regarding the other peers is distributed all over the botnet. Bots can send and receive their revision numbers while communicating. If the revision number varies, the older version gets updated to the newer version automatically. So, monitoring such an activity becomes very difficult. Peer to Peer is anetwork in which any node in a network can act as both a client anda server.P2P botnets aim at removing the failure point which is themain limitation and vulnerability of centralized networks.Fig.4 shows decentralized Model. P2P communication system is much strong, complex and typicallyno guarantees on message delivery or latency. Transferringcommand of P2P botnet is a slow to compare with centralizedbotnet. This means that the compromise of a single bot does notnecessarily mean the loss of the entire botnet [].

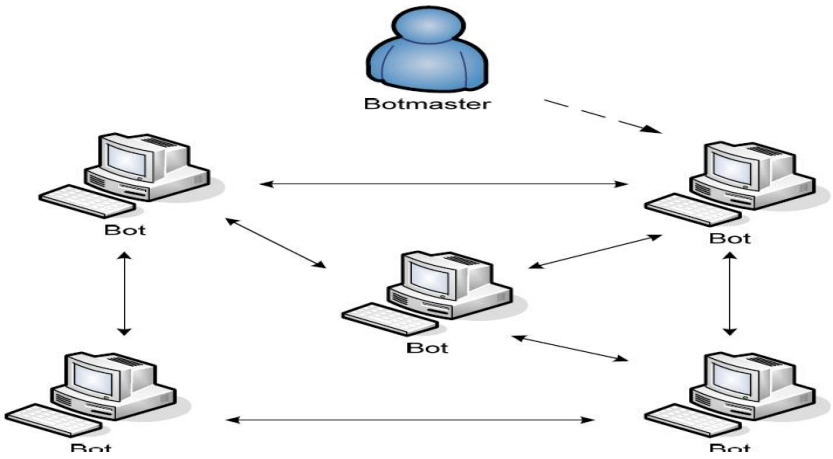

Fig.4.decentralized Model

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 1, January 2016

3.3 Hybrid Model

A hybrid peer-to-peer botnet based on the unstructured P2Pprotocols []. A hybrid botnet will be divided into servent andclient bot. The servent bot receives the commands from the botmaster, and it forward to the client. Fig.5 shows the Hybrid model very well. The hybrid P2P botnet is equivalent to a C&C botnet whereservent bots take the role of C&C servers: the number of C&Cservers (servent bots) is greatly enlarged, and they interconnect witheach other. In hybrid P2P botnet in comparision in current botnet isharder to shut down monitored and hijacked.

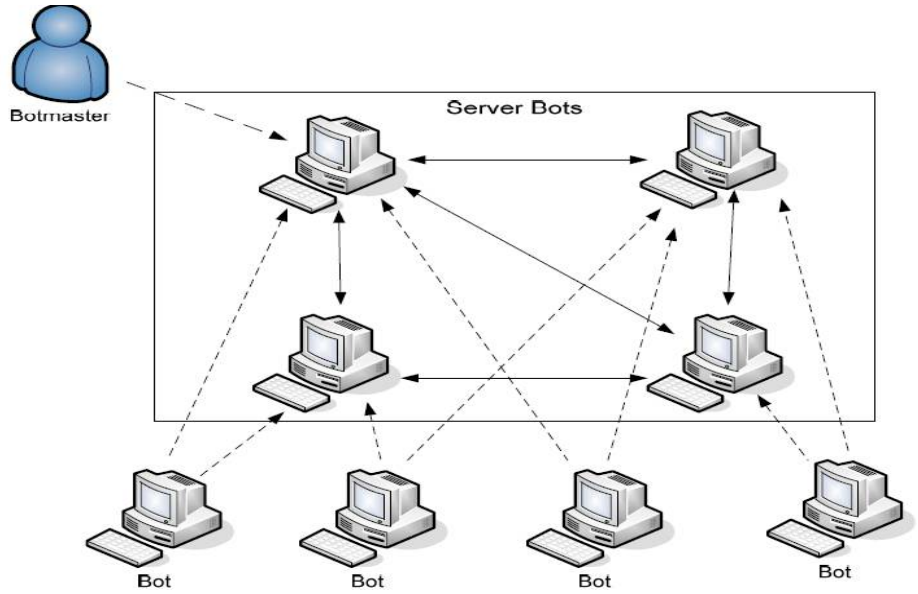

Fig.5.Hybrid Model

IV. BOTNET DETECTION TECHNIQUE

One of the most important aspects of this distinction betweendifferent types of organizations is the different data types that are available. Detecting a botnet often needs advanced analysing capabilities which are related to the selected data for analysis track and the characteristics of issues performed.An enterprise network might haveaccess to DHCP logs, DNS resolver data, address allocationdata, complete packet traces for each host, email serverlogs, policy data, as well as antivirus scanning logs. A networkservice provider on the other hand might only haveaccess to sampled or unsamplednetflow data and perhapssome limited packet tap data. While it is possible to inferactivity such as DNS requests or SMTP activity, the accuracyand confidence in this data would depend on the netflowsampling being used. In detection based on used data, Krugel et al. define intrusion detection as “the process of identifying and responding to malicious activities targeted at computing and network resources”. Intrusion Detection System (IDS) [13] discriminates intrusion attempts from normal system usage. IDS based detection required extra payload on bandwidth. Famous tools like Snort and Bro consume high resource when they deal with huge amount of payload data of in today high speed networks. Anomaly Based IDS also known as behaviour-based IDS, compare input data with the expected behaviour of the system. The system can detect unknown attacks because of their anomaly (irregular) based nature; they may give false positive alarms. Flow based solutions are more comfortable than intrusion detection systems. Flows are monitored by specialized accounting modules usually placed in network routers.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 1, January 2016

4.1 DNS Based Techniques:

Data regarding name resolution can be obtained by mirroring data to and from the local DNS servers [5] or resolvers and can be used to detect both botnetattack behaviour such as email spam (MX querylookups), as well as botnet communication behavioursuch as DNS lookups for suspicious domains.In order to access the C&C server bots carry out DNS queries to locate the particular C&C server that is typically hosted by a DDNS (Dynamic DNS) provider. So DNS monitoring will be easy approach to detect Botnet DNS traffic and detect DNS traffic anomalies. This is most famous and easy technique to botnet detection but it will be tough to detect recent advanced botnet through this technique.

4.2NetflowBased Techniques:

Netflow [5] data represents informationgathered from the network by sampling traffic flowsand obtaining information regarding source and destinationIP addresses and port numbers. At a courselevel, this data is useful for identifying malicious communicationpatterns and course grained attacks, but oftenvisibility is limited to the peering edge of a network,missing large amounts of backbone (ISP) orswitched (enterprise) traffic.

4.3 HoneypotBased Techniques:

The use of sacrificial hosts, placed inthe network with the express intention of them beingturned into bot members, can be a powerful tool forgaining insight into botnet means and motives withoutactually involving production hosts. A honeypot can be defined as an “environment where vulnerabilities have been eliberately introduced to observe attacks and intrusions”. Unfortunately, aspropagation techniques tend towards social engineering,these honeypots must increasingly emulate notonly user systems but the users themselves to be useful [7].

4.4Intrusion Detection SystemBased Techniques:

It can be characterized mainly in two ways.

a) Signature Based Botnet Detection: Rule based intrusion detection systems like Snort are running by using known malware signatures. They monitor the network traffic and detect sign of intrusions. It is obvious that payload information of network traffic is transformed and embedded into the signature or rule. The IDS detects malicious traffic fitting the communication parameters proposes a framework, “BotHunter” [11], to correlate IDS based detection alerts.

b) Anomaly Based Botnet Detection: This approach tries to detect Botnet based on number of network traffic anomalies such as high network latency, high volumes of traffic, traffic on unusual ports, and unusual system behaviour that could show existence of bots in the network. This approach can detect unknown Botnets. It can be categorized as Host based and Network Based Detection. In host based detection technique, a detection strategy which monitors and analyses the internals of a computer system instead of network traffics on its external interfaces. Limitation with this system is high false positive. A network-based technique is a detection strategy which tries to detect Botnets by monitoring network traffics. We can classify Network-based techniques into two categories: Active monitoring and passive monitoring. In Active monitoring, it injects test packets in network to measure the reaction of network such that gaining extra traffic on network. ABotsniffer[10] that uses network-based anomaly detection to identify Botnet C&C channels in a local area network.

4.5 Data Mining Based Technique:

Data mining aims to recognize useful patterns to discover regularities and irregularities in large data sets. Packet flow provides full information of flow data but in large file type. Anomaly based techniques are mostly based on network behaviour anomalies such as high network latency, activities on unused ports [4]. Data mining technique can be applied for optimization purpose. It enables to extract sufficient data for analysis from network log file. Most useful data mining techniques includes correlation, classification, clustering statistical analysis, and aggregation for efficiently knowledge discovery about network flows [8]. Flow correlation algorithms are useful to compare flow objects based on some characteristic other than packet content. This technique is very effective when content of packet is not available of encrypted, e.g. might compare arrival time. These kinds of algorithms utilize the characteristic values as inputs into

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 1, January 2016

one or more functions to create a metric used to decide if the flows are correlated [12]. Classification algorithms assume that incoming packet will match one of the previous patterns.

V. CONCLUSION AND FUTURE WORK

Botnets pose a significant and growing threat against cyber-security as they provide a key platform for many cyber-crimes such as Distributed Denial of Service (DDoS) attacks against critical targets, malware dissemination, phishing, and click fraud. Despite the long presence of malicious botnets, only few formal studies have examined the botnet problem and botnet research is still in its infancy. This paper surveys botnet and botnet detection. In this survey botnet detection techniques based on passive network traffic monitoring are classified into five classes including netflow based, signature-based, anomaly-based, DNS- based, and mining-base.

REFERENCES

1. PiyushChandekar and Dr. S. Chopra, "Efficient Sybil Attack Defence Mechanisms in Large Social Networks," International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE) Vol. 3, Issue 10, October 2015.
2. J.-T. Kim, H.-K. Park, and E.-H. Paik, "Security issues in peer-to-peersystems," in Advanced Communication Technology, 2005, ICACT 2005.The 7th International Conference on, vol. 2. IEEE, 2005, pp. 1059–1063.
3. D. Dittrich and S. Dietrich, "P2p as botnet command and control: adeeper insight," in Malicious and Unwanted Software, 2008. MALWARE2008. 3rd International Conference on. IEEE, 2008, pp. 41–48..
4. Felix Brezo, Jose Gaviria de la Puerta, XabierUgarte-Pedrero,Igor Santos, Pablo G. Bringas"A Supervised Classification Approach for DetectingPackets Originated in a HTTP-based Botnet"Clei Electronic Journal, Volume 16, Number 03, Paper 02, December 2013.
5. Daniel Plohmann, ElmarGerhards,Felix Leder, "Botnets: Detection, Measurement, Disinfection &Defence" ENISA 2011.
6. Mohini N. Umale, Prof. A. B. Deshmukh, Prof. M. D. Tambakhe, "Review on Botnet Threat Detection in P2P" International Journal on Recent and Innovation Trends in Computing and Communication Volume: 3 Issue: 2, 2015.
7. Pratik Narang, Subhajit Ray, ChittaranjanHota"PeerShark: Detecting Peer-to-Peer Botnetsby Tracking Conversations" IEEE Security and Privacy Workshops 2014.
8. Ritu and RishabhKaushal"Machine Learning Approach for Botnet Detection", IEEE 2010.
9. David Barroso"Botnets – The Silent Threat", ENISA 2007.
10. GuofeiGu, Junjie Zhang, and Wenke Lee "BotSniffer: Detecting Botnet Command and Control Channelsin Network Traffic", IEEE 2008.
11. Guofei Gu1, Phillip Porras, Vinod Yegneswaran, Martin Fong and Wenke Lee "BotHunter: Detecting Malware Infection through IDS-Driven Dialog Correlation" IEEE 2010.
12. PijushBarthakur, MrinalKantiGhoseand ManojDahal, "A Framework for P2P Botnet Detection Using SVM", International Conference on Cyber-Enabled Distributed Computing and Knowledge Discover 2012.
13. J. Zhang, R. Perdisci, W. Lee, U. Sarfraz, and X. Luo, "Detectingstealthy p2p botnets using statistical traffic fingerprints," in DependableSystems & Networks (DSN), 2011 IEEE/IFIP 41st InternationalConference on. IEEE, 2011, pp. 121–132.

BIOGRAPHY

Ruchi Dhole is pursuing M.E (Computer) from Lokmanya Tilak College of Engineering, Navi Mumbai, Mumbai University. She did his graduation B.E (Computer) from Mumbai University, Maharashtra. She is currently working on Conversation based detection of malicious peer-to-peer botnets.

Shobha Lolge is working as assistant professor in Dept. Of Computer Engineering at Lokmanya Tilak College of Engineering, Navi Mumbai, Mumbai University. She has academic experience of 12 years at UG and PG level courses of University of Mumbai. She has guided many projects at UG and PG level. Her area of interest are Database Management, Software Engineering, Mobile Computing, Artificial intelligence.