Mechanisms for Intrusion Detection in Peer-to-Peer Networks

SYNOPSIS

Submitted in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

PRATIK NARANG
ID. No. 2011PHXF414H

Under the Supervision of

Prof. Chittaranjan Hota

Co-supervision of

Prof. V. N. Venkatakrishnan

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

2015

Synopsis

Introduction

Computer networking has undergone a paradigm change over the past decade, with the number of users, applications, and computing devices going through an explosive growth. The past decade saw the immense rise of the Peer-to-Peer (P2P) computing systems. In the beginning of the twenty-first century, the P2P architecture attracted a lot of attention of developers and end-users alike, with the share of P2P over the Internet in different continents being reported to be in the range of 45% to 70% [Ipoque 2008]. [Androutsellis-Theotokis & Spinellis 2004] define P2P computing systems in the following manner: *"Peer-to-peer systems are distributed systems consisting of interconnected nodes able to self-organize into network topologies with the purpose of sharing resources such as content, CPU cycles, storage and bandwidth, capable of adapting to failures and accommodating transient populations of nodes while maintaining acceptable connectivity and performance, without requiring the intermediation or support of a global centralized server or authority"*. As an increasing number of users got access to powerful processors, large storage spaces, and increasing bandwidth, P2P networks presented a great opportunity to share and mobilize resources. The runaway success of P2P applications is primarily attributed to the ease of resource sharing provided by them – in the form of music, videos, files (BitTorrent, eMule, Gnutella, etc.), sharing of computing resources

(SETI @ home project), music streaming (Spotify), IPTV (LiveStation) or Voice-over-IP based services (Skype[1]). The 2014 global Internet phenomena report by Sandvine [Sandvine 2014] points to the overall percentage of P2P traffic to be 27% in the Asia-Pacific, with BitTorrent being the dominant application of P2P.

P2P file-sharing is notorious for being a source of information leakage, piracy, spread of malware etc. It also consumes high bandwidth and has known to become a serious concern for Internet Service Providers (ISPs), Governments and other public and private organizations. P2P traffic also has many characteristics that overlap with malicious traffic. For example: multiple persistent high-throughput flows (similar to spyware), communication with centralized P2P trackers (also seen in botnets), large number of simultaneous peer connection requests, many of which are unsuccessful due to peer-churn (similar to self-propagating malware infections and port-scan attacks), communication on uncharacteristic ports, receiving requests from a peer and forwarding those requests to neighbors immediately, trying to connect using both TCP and UDP ports, etc. Consequently, in addition to consuming considerable network resources, P2P traffic also creates several issues for network security devices such as Intrusion Detection Systems (IDS), firewalls, etc.

As P2P networks are inherently modeled without any centralized server, they lack a single point of failure [Buford *et al.* 2008]. This resilience offered by P2P networks has also attracted the attention of adversaries in the form of bot-masters (a.k.a. bot-herders). A 'bot' is a computer program which enables the operator to remotely control the infected system where it is installed. A network of such compromised end-hosts under the remote command of a master (i.e., the bot-master) is called a 'Botnet'. The ability to remotely command such bots coupled with the sheer size of botnets (numbering to tens of thousands of bots) gives the

[1]Skype has now moved to a cloud-based architecture [Gillet 2013].

bot-masters immense power to perform nefarious activities. Recent botnets are known to utilize the P2P architecture for their command-and-control (C&C) communications. The bots create an 'overlay network' amongst themselves and use P2P channels to exchange commands, pass-on stolen information, etc., which offer high resilience towards network break-down and take-down attempts [Rossow *et al.* 2013, Andriesse *et al.* 2013]. Even if a few bots in the network are identified or taken down, the botnet does not break-down. Botnets have been employed for spamming, Bitcoin mining, click-fraud scams, distributed denial of service (DDoS) attacks, etc. on a massive scale, and generate millions of dollars per year in revenue for the bot-master [Kanich *et al.* 2011].

Over the past decade, the P2P paradigm has moved beyond its boundaries of file sharing, and has seen deployment for several applications such as the SETI @ home project, Spotify, LiveStation, etc. The P2P architecture is here to stay, and hence there is need for the current security and intrusion detection mechanisms to be more 'P2P-aware'. This need arises from the fact that P2P networks behave differently from the traditional systems in several aspects, such as the lack the traditional client-server architecture, peer-churn (the joining and leaving of peers), security issues in a distributed and decentralized environment, etc.

This thesis proposes novel mechanisms for intrusion detection in P2P networks. In the next section, we will discuss related work on intrusion detection mechanisms in P2P networks. In the subsequent sections, we will give a brief overview of the approaches proposed in this thesis and summarize our research contributions.

Related work

Initial work on detection of P2P botnets involved signature-based and port-based approaches [Schoof & Koning 2007]. Such solutions—which rely on Deep Packet

Inspection (DPI) and signatures—can easily be defeated by bots using encryption.

Most prior work has either focused on P2P traffic classification from the perspective of a more general problem of Internet traffic classification [Sen *et al.* 2004, Li *et al.* 2008, Iliofotou *et al.* 2009], or has given special attention to detection of botnets (centralized or distributed) in Internet traffic [Gu *et al.* 2008]. The challenging context of detection of stealthy P2P botnets in the presence of benign P2P traffic has not received much attention. Furthermore, building a scalable detection framework did not receive much focus during early research, and has received very little attention even in recent research (such as in [Zhang *et al.* 2014, Singh *et al.* 2014]).

For the detection of P2P botnet traffic, some of the recent work has used supervised learning approaches [Saad *et al.* 2011, Rahbarinia *et al.* 2014, Narang *et al.* 2013, Singh *et al.* 2014], unsupervised learning approaches [Zhang *et al.* 2011, Zhang *et al.* 2014] and other statistical measures [Noh *et al.* 2009, Yen & Reiter 2010].

Most of the past works have employed the classical five-tuple categorization of network flows. Indeed, one of our preliminary work [Narang *et al.* 2013] also utilized five-tuple categorization of flows to study the impact of feature selection on detection of P2P botnets. Packets were classified as 'flows' based on the five-tuple of source IP, source port, destination IP, destination port, and transport layer protocol. Flows have bidirectional behavior, and the direction of the flow is decided based on the direction in which the first packet is seen. This traditional definition of flows has been greatly employed and has seen huge success in the problems of Internet traffic classification [Karagiannis *et al.* 2005] and even in the early days of P2P traffic classification [Karagiannis *et al.* 2004]. This definition relies on port number and transport layer protocol. The latest P2P applications as well advanced P2P bots are known to randomize their communication port(s) and

operate over TCP as well as UDP. Such applications will not be well-identified by these traditional approaches. Since such a behavior is characteristic of only the latest variants of P2P applications (benign or malicious), it is obvious that past research did not refer this aspect. In response to this, some recent work has utilized super-flow and conversation based approaches which are port-oblivious and protocol-oblivious [Zhang 2013, Hang *et al.* 2013, Li *et al.* 2013]. However, these approaches will fail to detect botnet activity if P2P bots and apps are running on the same machine (which might be a rare scenario, but cannot be ruled out nonetheless). This is because conversations try to give a *bird's eye view* of the communications happening in the network, and miss certain finer details in the process.

P2P networks have been studied from the game theoretic perspective mainly with regard to incentives for sharing [Anceaume *et al.* 2005], collaboration [Ye *et al.* 2004], managing trust [Kamvar *et al.* 2003], etc. Modeling of malicious behavior in this context has received much less attention from the research community. Security in P2P networks *per se* has not received much attention from a game theoretic perspective. However, the topic of network security with game theory has attracted a lot of attention. [Kodialam & Lakshman 2003] and [Vaněk *et al.* 2012] consider optimal resource allocation by a defender in a network against potentially malicious packets by adopting a game theoretic approach of inspecting only a fraction of all packets. The work of [Kodialam & Lakshman 2003] was limited to a single source and single target, whereas [Vaněk *et al.* 2012] considered multiple targets.

Past research has also dealt with different aspects of 'P2P intrusion detection'. [Janakiraman *et al.* 2003] presented a collaborative, P2P approach for building a distributed, scalable IDS amongst trusted peers. [Locasto *et al.* 2005] deployed a decentralized system for efficiently distributing alerts to collaborating peers by

creating a distributed 'watch-list' from alert streams. [Duma *et al.* 2006] explore the challenge of collaborative 'P2P intrusion detection' from the perspective of insider threat.

Proposed approaches

Flow-clustering and conversation-generation for P2P botnet detection

The first two of four approaches proposed in this thesis deal with the detection of P2P botnet traffic in the presence of benign P2P traffic at a network perimeter, by exploiting behavioral differences between P2P bots and benign P2P applications. Our approaches do not rely on DPI or signature-based mechanisms which are easily defeated by botnets/applications using encryption. They do not assume the availability of any 'seed' information of bots through blacklist of IPs. They aim to detect the *stealthy* behavior of P2P botnets on the basis of their 'P2P' behavior and C&C communications with other bots, while they lie dormant in their rally or waiting stages (to evade detection by IDS which look for anomalous communication patterns) or while they perform malicious activities (spamming, password stealing, etc.) in a manner which is not observable to the network administrator.

The first approach presents a 'best of the both worlds' approach utilizing flow-based approaches as well as conversation-based approaches in a two-tier architecture. It begins with the *de facto* standard of five-tuple flow-based approach and clusters flows into different categories based on their behavior. Within each cluster, we create 2-tuple 'conversations' from flows. Conversations are oblivious to the underlying flow definition (i.e., they are port- and protocol-oblivious) and essentially capture the idea of *who is talking to whom*. For all conversations, statistical

features are extracted which quantify the inherent 'P2P' behavior of different applications. Further, these features are used to build supervised machine learning models which can accurately differentiate between benign P2P applications and P2P botnets.

Our system was extensively evaluated with real-world traces of P2P applications and botnets. Our approach can effectively detect activity of *stealthy* P2P botnets even in the presence of benign P2P applications in the network traffic. It could also detect *unknown* P2P botnets (i.e., those not used during the training phase) with high accuracy.

Noise-resistant mechanisms for P2P botnet detection

The context of P2P botnet detection is adversarial in nature. Statistical or behavioral models for detection are created using 'botnet' data which has been generated by an adversary. Hence, the adversary is in a position to evade the detection mechanisms if he can change the behavior of his bots. Thus, this necessitates the evaluation of the performance of these detection models in the presence of deliberate injection of noise by an adversary. That is, if the bot-master slightly alters the behavior and communication patterns of the bots, are these detection models *robust* and *resistant* towards such a change? To the best of our knowledge, this question has not received sufficient attention. We attempt to address this context in this work.

Our approach utilizes conversation-based mechanisms and attempts to enhance them with techniques of Discrete Fourier Transforms (DFTs) and information entropy by leveraging the *timing* and *data patterns* in P2P botnet traffic. We extract two-tuple conversations from network traffic and treat each conversation as a time-series sequence (or a 'signal'). We leverage on the fact that commu-

nication of bots amongst each other follows a certain regularity or periodicity with respect to timing and exchange of data. Bots tend to repeatedly exchange same kind of commands—which are often in the same format and of the same size. The repeated C&C communication also follows certain timing patterns—with bots generally contacting their fellow peers at predefined intervals [Tegeler *et al.* 2012]. In order to uncover the hidden patterns between the communications of bots, we convert the *time-domain* network communication to the *frequency-domain*. From each conversation, we extract features based on Fourier transform and information entropy. We use real-world network traces of benign P2P applications and P2P botnets to compare the performance of our features with traditional flow-based features employed by past research (such as [Livadas *et al.* 2006, Saad *et al.* 2011, Kheir & Wolley 2013, Zhang *et al.* 2014]).

We build detection models with multiple supervised machine learning algorithms. We inject noise in our test data to demonstrate that our detection approach is more resilient towards variation in data or introduction of noise in the data by an adversary. With our approach, we could detect P2P botnet traffic in the presence of injected noise with True Positive rate as high as 90%.

Game theoretic strategies for IDS deployment in P2P networks

Although the decentralized and distributed nature of P2P network offers resilience towards network-breakdowns, the super-peer architecture is more sensitive in this regard since an adversary can disrupt (albeit not breakdown) the entire P2P network by attacking the super-peer nodes. For example, a DoS/DDoS attack targeted on relay nodes in Tor can lead to an increased latency and higher number of time-outs in the network. In our approach, we consider the problem of securing a super-peer based P2P network from an adversary who may become part of the P2P network by joining from any part of the network. A malicious peer can

disrupt the P2P network by attacking a super-peer through various attacks at the overlay layer, such as route table poisoning, index poisoning, or other traditional attacks (malicious payloads, etc.). Running an IDS at each peer may not be feasible since self-interested peers may not want to dedicate resources for that. Peers may try to secure the network by running IDS at certain strategic locations in the network. But, a deterministic schedule of running and positioning the IDS can be observed and thwarted by an adversary. In our work, we explore the problem of strategically positioning IDS in a P2P network with a zero-sum game theoretic approach. Our approach distributes the responsibility of running the IDS between the peers in a randomized fashion and minimizes the probability of a successful attack.

While past research has proposed techniques for sharing information between trusted peers [Janakiraman *et al.* 2003], distributing alerts among collaborating peers [Locasto *et al.* 2005] and managing trust to address insider threats [Duma *et al.* 2006], we address the issue of strategic deployment of IDS within a P2P network. Past research on different aspects of 'P2P intrusion detection' stands to gain from such strategic deployment of IDS since a deterministic schedule of running or positioning the IDS might be observed and thwarted by an adversary.

A Hadoop-based framework for detection of P2P botnets

Although many approaches have been proposed which evaluated the detection of P2P botnets in Internet traffic [François *et al.* 2011] or proposed mechanisms for the detection of P2P botnets in the presence of benign P2P traffic [Rahbarinia *et al.* 2014], building a scalable detection framework has received very little attention in past research (such as in [Zhang *et al.* 2014]).

We present a scalable, Hadoop-based framework for the detection of P2P botnets

which extracts statistical features *per host* for all P2P hosts involved in network communication. Although some past research have utilized host-level features [Zeng *et al.* 2010, Yen & Reiter 2010], these approaches did not consider the problem from the perspective of a scalable framework. Our approach relies on the header information in the network and transport layer, and extracts statistical features which quantify the 'P2P' behavior of the P2P applications running on a host. Statistical features are extracted *per host* for all P2P hosts involved in network communication, and are then used to train supervised machine learning models which can differentiate P2P botnets from P2P applications. We propose a distributed data collection architecture wherein data collectors are distributed at multiple locations inside an enterprise network and sit close to the peers, say at an Access switch or a Wi-Fi access point. This allows *inside-to-inside* communication view, which can be vital for detecting smart P2P bots inside a network which communicate to each other over LAN.

Future scope of work

The following areas can benefit from further research:

1. A thorough evaluation of the effect of injection of noise in the detection of P2P botnet traffic is required. Statistical and behavioral models need to explore and utilize heuristics or features which are resistant towards changes in communication patterns of bots.

2. Game theoretic approaches – which consider rational, utility-maximizing peers – can benefit from more detailed modeling of the players (namely the attacker and the defender) and the payoffs. Considering other forms of games such as non-zero-sum games will bring up new challenges in terms of identifying the equilibrium solution. Future work also needs to consider

solution concepts beyond the Nash equilibrium.

3. Distributed and scalable frameworks for malicious (botnet) P2P traffic need to be further improved to incorporate information from network communication (in the form of flows/conversations) as well as host-level information. Integrating these approaches with a distributed data collection approach can strengthen the detection of malicious activities.

References

[Anceaume *et al.* 2005] Emmanuelle Anceaume, Maria Gradinariu and Aina Ravoaja. *Incentives for p2p fair resource sharing*. In Peer-to-Peer Computing, 2005. P2P 2005. Fifth IEEE International Conference on, pages 253–260. IEEE, 2005. 5

[Andriesse *et al.* 2013] Dennis Andriesse, Christian Rossow, Brett Stone-Gross, Daniel Plohmann and Herbert Bos. *Highly resilient peer-to-peer botnets are here: An analysis of Gameover Zeus*. In Malicious and Unwanted Software:" The Americas"(MALWARE), 2013 8th International Conference on, pages 116–123. IEEE, 2013. 3

[Androutsellis-Theotokis & Spinellis 2004] Stephanos Androutsellis-Theotokis and Diomidis Spinellis. *A survey of peer-to-peer content distribution technologies*. ACM Computing Surveys (CSUR), vol. 36, no. 4, pages 335–371, 2004. 1

[Buford *et al.* 2008] John Buford, Heather Yu and Eng Keong Lua. P2p networking and applications. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2008. 2

[Duma *et al.* 2006] Claudiu Duma, Martin Karresand, Nahid Shahmehri and Germano Caronni. *A trust-aware, p2p-based overlay for intrusion detection*. In Database and Expert Systems Applications, 2006. DEXA'06. 17th International Workshop on, pages 692–697. IEEE, 2006. 6, 9

[François *et al.* 2011] Jérôme François, Shaonan Wang, Radu State and Thomas Engel. *BotTrack: Tracking Botnets Using NetFlow and PageRank*. In Proceedings of the 10th International IFIP TC 6 Conference on Networking - Volume Part I, NETWORKING'11, pages 1–14. Springer-Verlag, Berlin, Heidelberg, 2011. 9

[Gillet 2013] Mark Gillet. *Skype's cloud-based architecture*. http://blogs.skype.com/2013/10/04/skype-architecture-update/, 2013. Accessed on 7th November 2014. 2

[Gu *et al.* 2008] Guofei Gu, Roberto Perdisci, Junjie Zhang and Wenke Lee. *BotMiner: Clustering Analysis of Network Traffic for Protocol- and Structure-independent Botnet Detection*. In Proceedings of the 17th Conference on Security Symposium, SS'08, pages 139–154, Berkeley, CA, USA, 2008. USENIX Association. 4

[Hang *et al.* 2013] Huy Hang, Xuetao Wei, M. Faloutsos and T. Eliassi-Rad. *Entelecheia: Detecting P2P botnets in their waiting stage*. In IFIP Networking Conference, 2013, pages 1–9, USA, May 2013. IEEE. 5

[Iliofotou *et al.* 2009] Marios Iliofotou, Hyun-chul Kim, Michalis Faloutsos, Michael Mitzenmacher, Prashanth Pappu and George Varghese. *Graph-based P2P Traffic Classification at the Internet Backbone*. In Proceedings of the 28th IEEE International Conference on Computer Communications Workshops, INFOCOM'09, pages 37–42, Piscataway, NJ, USA, 2009. IEEE Press. 4

[Ipoque 2008] Ipoque. *Ipoque Internet study 2008/2009*. http://www.ipoque.com/en/resources/internet-studies, 2008. Accessed on 4 January 2014. 1

[Janakiraman *et al.* 2003] Ramaprabhu Janakiraman, Marcel Waldvogel and Qi Zhang. *Indra: A peer-to-peer approach to network intrusion detection and prevention*. In Enabling Technologies: Infrastructure for Collaborative Enterprises, 2003. WET ICE 2003. Proceedings. Twelfth IEEE International Workshops on, pages 226–231. IEEE, 2003. 5, 9

[Kamvar *et al.* 2003] Sepandar D Kamvar, Mario T Schlosser and Hector Garcia-Molina. *The eigentrust algorithm for reputation management in p2p networks*. In Proceedings of the 12th international conference on World Wide Web, pages 640–651. ACM, 2003. 5

[Kanich *et al.* 2011] Chris Kanich, Nicholas Weavery, Damon McCoy, Tristan Halvorson, Christian Kreibichy, Kirill Levchenko, Vern Paxson, Geoffrey M. Voelker and Stefan Savage. *Show Me the Money: Characterizing Spam-advertised Revenue*. In Proceedings of the 20th USENIX Conference on Security, SEC'11, pages 15–15, Berkeley, CA, USA, 2011. USENIX Association. 3

[Karagiannis *et al.* 2004] Thomas Karagiannis, Andre Broido, Michalis Faloutsos and Kc claffy. *Transport Layer Identification of P2P Traffic*. In Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement, IMC '04, pages 121–134, New York, NY, USA, 2004. ACM. 4

[Karagiannis *et al.* 2005] Thomas Karagiannis, Konstantina Papagiannaki and Michalis Faloutsos. *BLINC: multilevel traffic classification in the dark*. In ACM SIGCOMM Computer Communication Review, volume 35, pages 229–240. ACM, 2005. 4

[Kheir & Wolley 2013] Nizar Kheir and Chirine Wolley. *Botsuer: Suing stealthy p2p bots in network traffic through netflow analysis*. In Cryptology and Network Security, pages 162–178. Springer, 2013. 8

[Kodialam & Lakshman 2003] Murali Kodialam and TV Lakshman. *Detecting network intrusions via sampling: a game theoretic approach*. In INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies, volume 3, pages 1880–1889. IEEE, 2003. 5

[Li *et al.* 2008] Jun Li, Shunyi Zhang, Yanqing Lu and Junrong Yan. *Real-time P2P traffic identification*. In Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008, pages 1–5, USA, 2008. IEEE. 4

[Li *et al.* 2013] Liyun Li, Suhas Mathur and Baris Coskun. *Gangs of the internet: Towards automatic discovery of peer-to-peer communities*. In Communications and Network Security (CNS), 2013 IEEE Conference on, pages 64–72, USA, 2013. IEEE. 5

[Livadas *et al.* 2006] Carl Livadas, Robert Walsh, David Lapsley and W Timothy Strayer. *Using machine learning techniques to identify botnet traffic*. In Local Computer Networks, Proceedings 2006 31st IEEE Conference on, pages 967–974. IEEE, 2006. 8

[Locasto *et al.* 2005] Michael E Locasto, Janak J Parekh, Angelos D Keromytis and Salvatore J Stolfo. *Towards collaborative security and p2p intrusion detection*. In Information Assurance Workshop, 2005. IAW'05. Proceedings from the Sixth Annual IEEE SMC, pages 333–339. IEEE, 2005. 5, 9

[Narang *et al.* 2013] Pratik Narang, Jagan Mohan Reddy and Chittaranjan Hota. *Feature Selection for Detection of Peer-to-peer Botnet Traffic*. In Proceedings of the 6th ACM India Computing Convention, Compute '13, pages 16:1–16:9. ACM, 2013. 4

[Noh *et al.* 2009] Sang-Kyun Noh, Joo-Hyung Oh, Jae-Seo Lee, Bong-Nam Noh and Hyun-Cheol Jeong. *Detecting P2P botnets using a multi-phased flow model*. In Digital Society, 2009. ICDS'09. Third International Conference on, pages 247–253. IEEE, 2009. 4

[Rahbarinia *et al.* 2014] Babak Rahbarinia, Roberto Perdisci, Andrea Lanzi and Kang Li. *PeerRush: Mining for unwanted P2P traffic*. Journal of Information Security and Applications, vol. 19, no. 3, pages 194 – 208, 2014. 4, 9

[Rossow *et al.* 2013] Christian Rossow, Dennis Andriesse, Tillmann Werner, Brett Stone-Gross, Daniel Plohmann, Christian J Dietrich and Herbert Bos. *SoK: P2PWNED-Modeling and Evaluating the Resilience of Peer-to-Peer Botnets*. In Security and Privacy (SP), 2013 IEEE Symposium on, pages 97–111. IEEE, 2013. 3

[Saad *et al.* 2011] Sherif Saad, Issa Traore, Ali Ghorbani, Bassam Sayed, David Zhao, Wei Lu, John Felix and Payman Hakimian. *Detecting P2P botnets through network behavior analysis and machine learning*. In Privacy, Security and Trust (PST), 2011 Ninth Annual International Conference on, pages 174–180. IEEE, 2011. 4, 8

[Sandvine 2014] Sandvine. *Sandvine Global Internet Phenomena Report 2013*. https://www.sandvine.com/downloads/general/global-internet-phenomena/2014/1h-2014-global-internet-phenomena-report.pdf/, 2014. Accessed on 4th July 2015. 2

[Schoof & Koning 2007] Reinier Schoof and Ralph Koning. *Detecting peer-to-peer botnets*. University of Amsterdam, 2007. Technical report. 3

[Sen *et al.* 2004] Subhabrata Sen, Oliver Spatscheck and Dongmei Wang. *Accurate, Scalable In-network Identification of P2P Traffic Using Application Signatures*. In Proceedings of the 13th International Conference on World Wide Web, WWW '04, pages 512–521, New York, NY, USA, 2004. ACM. 4

[Singh *et al.* 2014] Kamaldeep Singh, Sharath Chandra Guntuku, Abhishek Thakur and Chittaranjan Hota. *Big Data Analytics framework for Peer-to-Peer Botnet detection using Random Forests*. Information Sciences, vol. 278, pages 488–497, 2014. 4

[Tegeler *et al.* 2012] Florian Tegeler, Xiaoming Fu, Giovanni Vigna and Christopher Kruegel. *Botfinder: Finding bots in network traffic without deep packet*

inspection. In Proceedings of the 8th international conference on Emerging networking experiments and technologies, pages 349–360. ACM, 2012. 8

[Vaněk *et al.* 2012] Ondřej Vaněk, Zhengyu Yin, Manish Jain, Branislav Bošanskỳ, Milind Tambe and Michal Pěchouček. *Game-theoretic resource allocation for malicious packet detection in computer networks*. In Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2, pages 905–912. International Foundation for Autonomous Agents and Multiagent Systems, 2012. 5

[Ye *et al.* 2004] Song Ye, Fillia Makedon and James Ford. *Collaborative automated trust negotiation in peer-to-peer systems*. In Peer-to-Peer Computing, 2004. Proceedings. Proceedings. Fourth International Conference on, pages 108–115. IEEE, 2004. 5

[Yen & Reiter 2010] Ting-Fang Yen and Michael K. Reiter. *Are Your Hosts Trading or Plotting? Telling P2P File-Sharing and Bots Apart*. In Proceedings of the 2010 30th International Conference on Distributed Computing Systems, ICDCS '10, pages 241–252. IEEE, 2010. 4, 10

[Zeng *et al.* 2010] Yuanyuan Zeng, Xin Hu and Kang G Shin. *Detection of botnets using combined host-and network-level information*. In Dependable Systems and Networks (DSN), 2010 IEEE/IFIP International Conference on, pages 291–300. IEEE, 2010. 10

[Zhang *et al.* 2011] Junjie Zhang, Roberto Perdisci, Wenke Lee, Unum Sarfraz and Xiapu Luo. *Detecting Stealthy P2P Botnets Using Statistical Traffic Fingerprints*. In Proceedings of the 2011 IEEE/IFIP 41st International Conference on Dependable Systems&Networks, DSN '11, pages 121–132, Washington, DC, USA, 2011. IEEE Computer Society. 4

[Zhang *et al.* 2014] Junjie Zhang, Roberto Perdisci, Wenke Lee, Xiapu Luo and Unum Sarfraz. *Building a Scalable System for Stealthy P2P-Botnet Detection*. Information Forensics and Security, IEEE Transactions on, vol. 9, no. 1, pages 27–38, 2014. 4, 8, 9

[Zhang 2013] Shaojun Zhang. Conversation-based p2p botnet detection with decision fusion. Master's thesis, Fredericton: University of New Brunswick, 2013. 5

Publications

Journal Papers

1. Narang, P., Reddy, J. M., & Hota, C. IDPT: Integrated Detection of P2P Threats in Parallel. *EURASIP Journal on Information Security*, Springer. Under review since 24th September 2015.
2. Narang, P., Hota, C., & Sencar, H. T. Noise-resistant Mechanisms for the Detection of Stealthy Peer-to-Peer Botnets. *Computer Communications*, Elsevier. Under review since 2nd April 2015.
3. Narang, P. & Hota, C. (2015). Game-theoretic Strategies for IDS Deployment in Peer-to-Peer Networks. *Information Systems Frontiers*, Springer, 2015, 17(5), 1017-1028.
4. Narang, P., Hota, C., & Venkatakrishnan, V. N. (2014). PeerShark: flow-clustering and conversation-generation for malicious peer-to-peer traffic identification. *EURASIP Journal on Information Security,* Springer, 2014(1), 1-12.

Conference Papers

1. Narang, P., Thakur, A., & Hota, C. (2014, December). Hades: A Hadoop-based framework for detection of peer-to-peer botnets. In *Proceedings of the 20th International Conference on Management of Data* (pp. 121-124). Computer Society of India.
2. Narang, P., Mehta, K., & Hota, C. (2014, December). Game-theoretic Patrolling Strategies for Intrusion Detection in Collaborative Peer-to-Peer Networks. In *International Conference on Secure Knowledge Management in Big-data era*, 2014. DOI: 10.13140/2.1.2533.2804
3. Narang, P., Khurana, V., & Hota, C. (2014, May). Machine-learning approaches for P2P botnet detection using signal-processing techniques. In *Proceedings of the 8th ACM International Conference on Distributed Event-Based Systems* (pp. 338-341). ACM.

4. Narang, P., Ray, S., Hota, C., & Venkatakrishnan, V.N. (2014, May). Peershark: Detecting peer-to-peer botnets by tracking conversations. In *Security and Privacy Workshops (SPW)*, 2014 IEEE (pp. 108-115). IEEE.

5. Narang, P., Reddy, J. M., & Hota, C. (2013, August). Feature selection for detection of peer-to-peer botnet traffic. In *Proceedings of the 6th ACM India Computing Convention* (pp. 16:1–9). ACM.

Biographies

Brief Biography of the Candidate

Pratik Narang is a Ph.D scholar and research assistant at the Computer Science department of Birla Institute of Technology and Science – Pilani, Hyderabad campus. Prior to pursuing PhD, he completed his M.Sc.(Tech) in Information Systems from BITS Pilani in 2011. His research interests lie in the area of network security and cyber-security with applications from machine learning and game theory, and his PhD is funded by grants from Department of Electronics & Information Technology (DeitY), Government of India. He has held a visiting position with New York University, Abu Dhabi campus. He has served as a sub-reviewer for international conferences IEEE WIFS 2015 and IEEE CNS 2014, and a reviewer for IEEE Transactions on Dependable and Secure Computing, Journal of Machine Learning and Cybernetics (Springer), Information Systems Frontiers (Springer) and EURASIP Journal on Information Security (Springer).

Brief Biography of the Supervisor

Chittaranjan Hota is a Professor and Associate Dean at Birla Institute of Technology and Science – Pilani, Hyderabad Campus, Hyderabad, India. He was the founding Head of Dept. of Computer Science at BITS, Hyderabad. Prof. Hota did his PhD in Computer Science and Engineering from Birla Institute of Technology & Science, Pilani. He has been a visiting researcher and visiting professor at University of New South Wales, Sydney; University of Cagliari, Italy; Aalto University, Finland and City University, London over the past few years. His research work has been funded by University Grants Commission (UGC), New Delhi; Department of Electronics & Information Technology (DeitY), New Delhi; and Tata Consultancy Services, India. He has guided PhD students and currently guiding several in the areas of Overlay networks, Information Security, and Distributed computing. He is recipient of Australian Vice Chancellors' Committee award, recipient of Erasmus Mundus fellowship from European commission, and recipient of Certificate of Excellence from Kris Ramachandran Faculty Excellence Award from BITS Pilani. He has published extensively in peer-reviewed journals and

conferences and has also edited LNCS volumes. He is a member of IEEE, ACM, IE, and ISTE. His research interests are in the areas of Traffic Engineering in IP Networks, Security and Quality of Service issues over the Internet, Peer-to-Peer Overlay Security, Mobile Wireless Networks and Internet of Things.

Brief Biography of the Co-Supervisor

V.N. "Venkat" Venkatakrishnan's (http://www.cs.uic.edu/ venkat) broad research interests are in computer security and privacy. He is particularly interested in the security of software systems, in vulnerability analysis and automated approaches for preventing large scale attacks on computer systems. His research work derives from techniques rooted in programming languages and compilers, operating systems, software engineering and formal methods to address practical problems in computer security. Specific areas within software security that his work has touched upon in the recent past include web application security, safe execution of untrusted third party content, web malware analysis and analysis / verification of systems software. He received his Ph.D. and M.S. degrees in computer science from Stony Brook University in 2004 and M.Sc and B.E. degrees from Birla Institute of Technology and Science (BITS), Pilani, India, in 1997. He is currently a Full Professor of Computer Science at the University of Illinois at Chicago (UIC). He is recipient of the National Science Foundation CAREER award in 2009, several best paper awards including a 2010 NYU-Poly AT&T Best Applied Cybersecurity Paper Award and multiple UIC campus level awards for research as well as his teaching.