A Survey of Game-Theoretic Approaches to Modeling Honeypots

Andrey Vishnevsky
Information Security Department
Bauman Moscow State Technical University
Moscow, Russia
andreyryu@yandex.ru

Petr Klyucharev
Information Security department
Bauman Moscow State Technical University
Moscow, Russia
pk.iu8@yandex.ru

***Abstract* — Honeypots are fake information resources that authorized users never connect with and which are under permanent control of information security specialists. Honeypots are widely used traps for hackers, which gather features of attacks. Collected features then are accumulated in anti-virus databases which serve as evidences in cyber forensics or as reference samples in machine learning systems. The quality of security tools depends on the ability to gather representative information about actual cyber-attacks.**

During the past twenty years, honeypots have evolved from standalone tools emulating one or two network services to systems of many highly interactive traps. Modern honeypots emulate a large scale of services from FTP and SSH to VoIP and industrial systems. They can monitor web-attacks, client-side exploitations, targeted attacks in corporate networks and intruder's activity. The weakest point occurs when hackers are aware of traps and often avoid honeypots by comparing them to real systems.

To solve this problem, researchers introduced game theoretic models to adapt behavior of honeypots to be undetected by hackers. In addition, they embed machine-learning techniques to improve the performance of honeypots if only a few stages of hacker attacks are executed. This paper is a review of game-theoretic models on which adaptive honeypots function.

Keywords—survey, review, honeypots, machine learning, game theory, deception games, intrusion detection, information security.

I. INTRODUCTION

Honeypots are traps disguised as information resources, which capture the details of computer attacks aimed at them. The collected data is added to signature bases of antiviruses, blacklists of firewalls and serve as reliable evidences in computer forensics. All trapped files, links, ip-addresses and other artifacts are clearly malicious software fragments, because authorized users do not interact with honeypots.

Modern honeypots emulate a wide range of vulnerable programs from web-applications and database management systems to VoIP-services, Internet-of-Things firmware and industrial information systems. Honeypots have started emulating not only server-side software but client-side applications: web-browsers and plugins for them, office software and even entire operating systems.

In the recent scientific literature, the mechanics of honeypots and mathematical models determining the behavior of traps were described.

The main issues of the deception mechanism have been the attraction of attackers into traps, revealing their intention and collecting as much evidence as possible. From the other side, attackers use their methods and targets in interacting with the honeypot system. This suggested system's way helps to achieve increased interactions and adaptation.

The rest of this paper is organized as follows. Section II contains summaries of the literatures that focus on honeypot reviews. Section III contains summaries of the articles that outlines the recent implementations of game-theoretic model in deception systems. Section IV contains description of self-adaptive honeypots which mechanics, in our opinion, can be improved by using the game theory. Section V is the conclusion of this survey, followed by the list of references in this survey. Each literature is referenced by the list of the authors.

II. OTHER HONEYPOT REVIEWS

There are a number of different types of honeypots with various technical options. Such deceptive technologies are widely mentioned in scientific literature.

In 2012 Bringer M.L., Chelmecki C., and Fujinoki H. published the research in the field of honeypots aimed at the invention of new types of honeypots, improvement of their creation and configuration, optimization of output data processing methods, and modernization of traps camouflage [1].

The history of honeypots' evolution from 1997 to 2016 is described by Nawrocki M. et al. in the survey [2]. Their review summarizes that modern honeypots can emulate vulnerabilities in files transfer services: SMB, FTP, TFTP [24]; remote access services: SSH [21], Telnet; email protocols: SMTP, POP3, IMAP; database management systems: Elasticsearch, MSSQL, MySQL [20]; wireless protocols: IEEE 802.11 (WiFi), Bluetooth; entire workstations with operating systems: Microsoft Windows XP, Windows 7, Linux, Android [22]; web-applications: Apache, php-BB, php-MyAdmin, web-servers [23]; instant messaging services: IRC; applications for

The reported study was funded by RFBR according to the research project № 16-29-09517

voice communications: VoIP [19], emulate DNS vulnerabilities, IoT devices [18] and Supervisory Control and Data Acquisition (SCADA) systems [22].

However, the current situation of the game-theoretic models realizations in honeypots is not detailed enough [26]. We managed to find only one recent review of a game theory application in honeypots that deserves attention. It was published in 2016. Sangeetha R., Mohana M. have provided a survey on game-theoretic approaches in honeypot enabled networks for the Internet of Things (IoT) [3]. They have summarized risks of IoT infrastructures and classified honeypots related to defense against attacks on IoT.

In our view, many of the articles about realizations of game-theoretic models in deception systems were not mentioned in surveys. Our paper is intended to fill this gap.

III. GAME-THEORETIC MODELS

Researchers have proposed the game-theoretic approach to make the behaviour of honeypots more like an operation of the real computer. Table I. summarizes the articles in this category. The table reflects the most essential features of proposed game-models: definitions of players and list of available actions.

In 2009, Wagener G. et al. has built and implemented a high interactive honeypot disguised as a SSH-server [4]. Honeypot behaviour was determined by a game-theoretic model and machine learning. In the restrictions of this model, the attacker can input various commands into SSH-shell. The honeypot can execute the entered commands or replace the application which has to execute the command. The honeypot payoff is defined as a number of new unknown dangerous objects, in particular, the number of malicious files, which were uploaded by the attacker. In this model, the honeypot chooses actions with probability which is defined by a predictor of a potential payoff. The predictor is learned using a set of previous decisions of the honeypot.

In 2012, Hayatle O., Otrok H., and Youssef A. proposed a game-theoretic model of interaction between an attacker and client honeypot [5]. In this model, the attacker has a botnet, i.e. a set of infected machines managed by him. The attacker doesn't know if his current target is a trap or a real host. In the restrictions of the proposed game, the attacker, to not expose his intentions, can probe the target in three ways: commands the bot to attack the sensor machine, commands the bot to attack the real target, chooses not to perform any activity. In the case of successful probing, the intruder attacks his target or otherwise retreats. The intruder can attack without preliminary probing of the infected machine. The honeypot has only two available actions: to allow the attack or to deny it. Optimal strategies were theoretically established for the attacker and the defender.

In 2016, Mohammadi A. et al. suggested the honeypot which is composed of fake avatars in social networks could distinguish hacked profiles [6]. The signal games and Nash equilibrium were used to develop the strategy of the honeypot.

In 2016, Kiekintveld C., Lisy V., and Pibil R created a game-theoretic approach to compute selection strategy of nodes in computer network which are the most optimal to be replaced by honeypots [7].

In 2017, Shi L. et al. put forward the idea of a mimicry honeypot system [8]. In this model, the defense has at its disposal real services, honeypots and pseudo honeypots (real services disguised as honeypots). The goal of the defense is to distinguish the attacker from a legitimate user. The proposed game model was realized as vulnerable FTP-server and validated by simulation.

In 2017, Du Miao et al. tried to find out novel ways of preventing DDoS attacks which were targeted at social networks [9]. The difference with the previous works is the ability to see a new type of attackers as rational and adapting to the strategy of the defender. They offered a new pseudo honeypot game model based on a Bayesian game and showed how to find Bayesian Nash equilibrium in the restrictions of the proposed model. It was shown empirically that computed optimal strategies make the defense more effective.

In 2017, Ziad Ismail et al. formulated a game-theoretic model for intrusion detection in computer networks [10]. In the restrictions of this model, the resources of the defense are limited. The goal of the defender is to optimize the allocation of intrusion detection systems (IDS) in the network. Additionally, interdependencies between equipment vulnerabilities were taken to improve the quality of game-theoretic analysis. The proposed model was tested in a real world scenario.

In 2016, Hayreddin Ceker et al. proposed a game-theoretic model of interaction between the defender and the attacker [11]. The goal of the defender is to optimize the network configuration for DoS attack prevention. In this model the defender can camouflage honeypots as real services and vice versa. The proposed model is based on signaling game with incomplete information. The existence of perfect Bayesian equilibrium was proved and used for finding optimal strategies. It is expected that the proposed deception strategy could be used to develop high quality and cheap security solutions for preventing DoS-attacks.

In 2017, Wang K. et al. formulated the interaction between Advanced Metering Infrastructure (AMI) network and the intruder as a game-theoretic model [12]. The defender's challenge is to embed honeypots to an AMI network for DDoS attack detection. To explore the optimal strategies of the defender and the attacker Bayesian Nash Equilibriums were used. The proposed game-theoretic model was validated empirically in the smart grid and its efficiency was proved.

In 2017, Nguyen T., Wellman M. P., Singh S. have explored the problem of allocating detection resources (detectors) in a computer network to deter botnet attacks [13]. In our opinion, honeypots could be used as detectors in the implementation of this model. In the proposed game-theoretic model, the attacker eavesdrops on network traffic and tries to send the stolen data outside the defender's network. The defender allocates limited detectors to protect the most valuable resources of the network. The goal of the defender is to randomize the placement of the detectors so that the locations of them become unpredictable to the intruder. The algorithm of computing optimal game strategies was offered with some heuristics for approximately solving the game when the number of nodes in the network is large. Given game model was evaluated via synthetic and real-world network topologies.

TABLE I. DECEPTION GAME MODELS

Players	Actions
Players: the attacker and high interaction honeypot. [4]	The attacker can enter commands into SSH. The honeypot can execute or spoof entered command with probability defined by reinforcement learning.
Players: attacker and client honeypot [5]	The attacker can scan trap, attack another computers from the infected trap or idle. The client honeypot can permit or block the attack.
Players: fake avatar in social network, an attacker and legal user. [6]	The attacker and user can send benign or infected messages. The fake avatar can raise the alarm or idle.
Players: attacker and honeynet composed of real computers and honeypots [7]	The defender places honeypots in computer network and sets their importance level. The attacker chooses target in the network on the basis of resource importance.
Players: server, legal user and attacker. [8]	The server can response to user as real service, as honeypot or as pseudo honeypot. The attacker can attack or abandon the attack.
Players: real service, honeypot service and pseudo honeypot service, legitimate users and attackers [9]	Real service, honeypot and pseudo honeypot can provide service or not. Legitimate user and the attacker can provide access or not
Players: an attacker and a defender [10]	The attacker can attack any node in the network. The defender's can distribute monitoring resources on network nodes in order to detect attacks.
Players: an attacker and a defender [11]	The defender can deploy honeypots, disguise normal systems as honeypots and honeypots as normal systems. The attacker can successfully compromise normal systems, but not honeypots.
Players: real communications, honeypot service, anti-honeypot service, legitimate visitors and attackers [12]	Real communications, honeypot service and anti-honeypot service can provide service or not provide service. Legitimate visitors and the attackers can access or not access.
Players: an attacker and a defender [13]	The defender can deploy limited number of detection resources in network. The attacker can compromise limited number of nodes in the network.

IV. RELATED WORK

Honeypots are designed to attract intruders and collect information about attacks. The game theory provides methods which can be used to make traps more interactive and therefore indistinguishable from real resources than traditional honeypots. Consequently, the area of adaptive honeypots is relative to game theory topic as a point of novel game models realization in the future. From the viewpoint of applying game-theoretic approaches, in our opinion, are next articles about adaptive honeypots.

In 2017, *Fernandez G., Nieto A., and Lopez J.* have formulated a concept of malware-driven honeypots and developed a mechanism for the dynamic reconfiguration of honeypots [14]. The goal of the proposed honeypot management system is to create trap environments so the whole malicious activity could be captured. To fulfill malware requirements the management system uses recent Indicators of Compromise (IOS) from malware intelligence services such as Malware Information Sharing Platform (MISP) and Virus Total Intelligence (VTI). It is the first published approach of using malware intelligence platforms for the dynamic deployment of honeypots. In our opinion, malware writers choose required features of a victim machine environment to infect as many computers as they can and to be undetected as long as possible. So, fulfilling the requirements of malware in traps can be described in the future as a competitive game between the defender and the attacker.

In 2017, *Pauna A. and Bica I.* presented a changing behavior honeypot system that overlaps with some of the disadvantages in the existing deception systems [15]. The offered honeypot system is made by using Python and it emulates a SSH (Secure Shell) server. The proposed system interacts with attackers and uses means of reinforcement learning algorithms. In our opinion, reinforcement learning of honeypots can be used to make traps capable of dynamically changing their behavior. This task is naturally related to strategic decision making using game-theoretic approaches.

In 2014, *Pauna A., Patriciu V.V.* have created an autonomous honeypot system capable of learning and adapting its behaviour by interacting with the attackers [16]. The designed case adaptive SSH honeypot is based on an existing medium interaction honeypot (Kippo) and implements Case Based Reasoning and Belief-Desire-Intention agents. Case Base Reasoning system is a system which solves tasks by making decisions used for similar tasks. The Belief-Desire-Intention agent model has a view of the world (Beliefs), a number of goals (Desires) and possible actions (Intentions). The actions are planned using the accumulated experience. The practical experiments have shown that the number of captured payload s is relatively similar to the ones obtained by the standard Kippo honeypot. As in the previous work, game theory approaches can be used to improve the intellectuality *of traps.*

In 2017, *Orzel M.J. and Grzegorz K.* proposed few schemes of web-attack detection [17]. For this purpose features from web-server log files were used. The collection of events was gathered from real web-site logs and helped to find unwanted web crawlers' traces. In our opinion, the considered features from web server log files could be used for building web-based server side honeypot system. Because there was no article about implementing the game theory to web-based honeypots, we mentions this article as a paper containing features to be collected by novel server-side traps.

V. CONCLUSION

For the last decades honeypots have evolved to networks of sensors which emulate various types of devices and applications. Honeypot configurations are increasingly based on game-theoretic models. The game-theoretic approach is used for honeypot preparation before an attack and for trap behavior adaptation during an attack. However, most publications about realization of game-theoretic models in honeypots are purely theoretical. Only a few practices are related to implementation of game-theoretic models to honeypots disguised as FTP and SSH servers.

Most effectiveness is expected from implementation of the game-theoretic approach to high interaction honeypots and to social networks, but this carries the risk of additional opportunities given to the attackers. These risks are discussed in the articles related to legal and ethical issues of using honeypots [25].

The combination of the game theory and machine learning has, in our view, the greatest potential for a honeypot to develop. The honeypot experience enriched during the attacks, by our estimates, will allow honeypot strategies to adapt so the traps will be indistinguishable from real services.

REFERENCES

[1] Bringer M.L., Chelmecki C., Fujinoki H. A survey: Recent advances and futuretrends in honeypot research. 2012. V. 4. N 09. DOI: 10.5815/ijcnis.2012.10.07.

[2] Nawrocki M., Wahlisch M., Schmidt T., Keil C., Schonfelder J. A survey on honeypot software and data analysis. 2016.

[3] Sangeetha R., Mohana M. A Survey on Game Theory against Attack in Honeypot Enabled Networks for IoT. 2016. DOI: 10.17148/IJARCCE.2016.51051.

[4] Wagener G., State R., Engel T., Dulaunoy A. Adaptive and self-configurable honeypots. Integrated Network Management. 2011. P. 345-352. DOI: 10.1109/INM.2011.5990710.

[5] Hayatle O., Otrok H.,Youssef A. A game theoretic investigation for high interaction honeypots. Proceedings of the 2012 IEEE International Conference on Communications (ICC), Ottawa, Canada. 2012. P. 6662-6667. DOI:10.1109/ICC.2012.6364760.

[6] Mohammadi A., Manshaei M. H., Moghaddam M. M., Zhu Q. A Game-Theoretic Analysis of Deception over Social Networks Using Fake Avatars. Proceedings of the Decision and Game Theory for Security - 7th International Conference, GameSec 2016. 2016. V. 9996. P. 382-394. DOI:10.1007/978-3-319-47413-7_22.

[7] Kiekintveld C., Lisy V., Pibil R. Game-Theoretic Foundations for the Strategic Use of Honeypots in Network Security. // Jajodia S., Shakarian P., Subrahmanian V., Swarup V., Wang C. (Eds.) Cyber Warfare. Advances in Information Security. V. 56. P. 81-101. 2015. DOI: 10.1007/978-3-319-14039-1_5.

[8] Shi L., Zhao J., Jiang L., Xing W., Gong J., Liu X. Game theoretic simulation on the mimicry honeypot. Wuhan University Journal of Natural Sciences. 2016. V. 21. P. 69–74. DOI: 10.1007/s11859-016-1140-2.

[9] Du M., Li Y., Lu Q., Wang K. Bayesian Game Based Pseudo Honeypot Model in Social Networks. 2017. DOI: 10.1007/978-3-319-68542-7_6

[10] Ismail Z., Kiennert C., Leneutre J., Chen L. A game Theoretical Model for Optimal Distribution of Network Security Resources. 2017. DOI: 10.1007/978-3-319-68711-7_13.

[11] Ceker H., Shambhu J., Quang U., Soong D. Deception-Based Game Theoretical Approach to Mitigate DoS Attacks. 2016. DOI: 10.1007/978-3-319-47413-7_2.

[12] Wang K., Du M., Maharjan S., Sun Y. Strategic Honeypot Game Model for Distributed Denial of Service Attacks in the Smart Grid. In IEEE Transactions on Smart Grid. V. 8. N 5. P. 2474-2482. 2017. DOI: 10.1109/TSG.2017.2670144.

[13] Nguyen T., Wellman M. P., Singh S. A Stackelberg Game Model for Botnet Data Exfiltration. In Decision and Game Theory for Security - 8th International Conference, GameSec 2017, Proceedings. 2017. V. 10575. P. 151-170. Springer Verlag. DOI: 10.1007/978-3-319-68711-7_9.

[14] Fernandez G., Nieto A., Lopez J. Modeling Malware-driven Honeypots. 14th International Conference On Trust, Privacy & Security In Digital Business (TrustBus 2017). 2017. V. 10442. P. 130-144. DOI 10.1007/978-3-319-64483-7_9.

[15] Pauna A., Bica I. RASSH - Reinforced adaptive SSH honeypot. 2014 10th International Conference on Communications (COMM). Bucharest. 2014. P. 1-6. DOI: 10.1109/ICComm.2014.6866707.

[16] Pauna A., Patriciu V.V. CASSHH – Case Adaptive SSH Honeypot. In Recent Trends in Computer Networks and Distributed Systems Security. SNDS 2014. Communications in Computer and Information Science, Springer, Berlin, Heidelberg. 2014. V. 420. DOI: 10.1007/978-3-642-54525-2_29.

[17] Orzel M.J., Grzegorz K. Detection of Security Incidents in a Context of Unwelcome or Dangerous Activity of Web Robots. 2017. P. 215-225. DOI: 10.1007/978-3-319-43982-2_19.

[18] Dowling S., Schukat M., Melvin H. A ZigBee honeypot to assess IoT cyberattack behaviour. 2017. P. 1-6. DOI: 10.1109/ISSC.2017.7983603.

[19] Jordao R., Vargas S., Kleinschmidt H. Capture and Analysis of Malicious Traffic in VoIP Environments Using a Low Interaction Honeypot. In *IEEE Latin America Transactions.* 2015. V. 13. N. 3. P. 777-783. DOI: 10.1109/TLA.2015.7069104.

[20] Djanali S., Arunanto F., Pratomo B. A., Studiawan H., Nugraha S. G. SQL injection detection and prevention system with raspberry Pi honeypot cluster for trapping attacker. *2014 International Symposium on Technology Management and Emerging Technologies.* Bandung. 2014. P. 163-166. DOI: 10.1109/ISTMET.2014.6936499.

[21] Koniaris I., Papadimitriou G., Nicopolitidis P. Analysis and visualization of SSH attacks using honeypots. Eurocon 2013. Zagreb. 2013. P. 65-72. DOI: 10.1109/EUROCON.2013.6624967.

[22] Jicha A., Patton M., Chen H. SCADA honeypots: An in-depth analysis of Conpot. 2016 IEEE Conference on Intelligence and Security Informatics (ISI). Tucson, AZ. 2016. P. 196-198. DOI: 10.1109/ISI.2016.7745468.

[23] Rahmatullah D. K., Nasution S. M., Azmi F. Implementation of low interaction web server honeypot using cubieboard. 2016 International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC). Bandung. 2016. P. 127-131. DOI: 10.1109/ICCEREC.2016.7814970.

[24] Perevozchikov V. A., Shaymardanov T. A., Chugunkov I. V. New techniques of malware detection using FTP Honeypot systems. 2017 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus). St. Petersburg. 2017. P. 204-207. DOI: 10.1109/EIConRus.2017.7910529.

[25] Sokol P. Legal issues of honeynet's generations. Proceedings of the 2014 6th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). Bucharest. 2014. P. 63-69. DOI: 10.1109/ECAI.2014.7090212.

[26] Shmatova E. The Choice of Strategy for the Spurious Information System on the Basis of the Game Theory Model. Voprosy kiberbezopasnosti *[Cybersecurity issues]*, 2015. N 5 (13). P. 36-40. DOI: 10.21681/2311-3456-2015-5-36-40.