Using deceptive systems as data sources for a cyber-threat analysis platform

E R Gruzeva¹, E A Guseva²

¹Chechen State University named after A.H. Kadyrov, 364015, 17a Dudaev Boulevard, Groznyy, Russian Federation
²Irkutsk National Research Technical University, Lermontov Street, 83, Irkutsk, 664074, Russian Federation

E-mail: el.guseva@rambler.ru

Abstract. This article provides an example of the use of deceptive systems as data sources for a cyber threat intelligence platform. The main result is the preliminary verification and enrichment of indicators of compromise (thereby trying to solve the problems of verification of indicators of compromise, IoC).

1. Introduction
In today’s world, almost every organization uses IT technologies. Naturally, the number of various kinds of cyber threats, which negatively affect the stability of business processes, is also growing. There is a constant race between threats and information protection tools (IPS), in which IPS tends to play the role of the catching-up party. One of the technologies of early detection of new threats is threat hunting, the second, realizing the exchange and further research – threat intelligence. These technologies are implemented within the framework of cyber threat analysis platforms – the so-called Threat Intelligence Platform, TIP. They help to collect, process, transmit and store data about existing threats in the form of indicators of compromise (Indicators of compromise, IoC). The formalization and transmission of indicators are described in the STIX/TAXII standards. There is a problem with the legitimacy of data sources and the verification of such information [1–3].

The paper discusses the architecture of a TIP that uses a deceptive system as a source of indicators.

2. Materials and Methods

2.1. Deceptive systems
To obtain more accurate data about threats, including zero-day [4], deceptive systems (honeypots) [5] can be used in the information system. They are traps for attackers, working on the principle of simulating real services, applications, protocols in an information system. In a honeypot, vulnerabilities are deliberately embedded to attract the attacker, collect data and form his portrait, thus we can get the actual information about new threats, in fact, from “trusted” sources (feed) and apply in TIP [10, 11].

2.2. Approach to the use of deceptive systems and TIP
A schematic diagram of a threat analysis system using the honeypot is shown below Figure 1.
The first stage is the deployment of deceptive systems, using virtualization technology or dedicated hardware devices. All cheating systems are based on Unix-like systems, which allow collecting and sending log files or data in syslog format. Further (see Figure 1) it is necessary to aggregate and formalize the collected data. To do this, you can use the STIX standard [3] and its language principles [12].

To store the obtained files it is best suited NoSQL-database, since it is required to index the obtained data groups and to store one event as an object, in which the field and message correspond to the type key-value, and since such data are not required to update or somehow change [14-16]. But simply to collect, index, store, and overtime to remove, then document-oriented databases such as MongoDB, Redis, ElasticSearch are suitable for this purpose Figure 2.

**Figure 1.** Use of TIP and deceptive systems.

**Figure 2.** Example of a document-oriented database.
At this stage, we get indexed documents with data in the form of key – values inside. Then we have a situation where we need to use the received data both within the system and send it for processing to the “cloud” (see Figure 1).

Inside the information system, you need to deploy a web service for visualization and analysis of the data obtained, such as Kibana, Graylog 2, Grafana.

Then, using the principles and approaches of the TAXII standard [6, 13], we can transfer the collected data to the “cloud” for further processing and filtering Figure 3.

![Image](image.png)

**Figure 3.** The use of obtained data from deceptive TIP systems is usually part of a deployed SOC (security operation center).

3. Results and Discussion

It is worth noting that the SOC can be provided as a SaaS-solution and solve only the tasks of analysis, monitoring. You can use the principles of SOC systems to analyze and process already received indicators. However, this can improve the process of verification. Then the platform used transmits indicators to other participants (information systems).

The use of cloud storage will allow more flexibility in working with indicators, processing and analyzing them before using them in the platform because the data obtained in the information system itself can aggregate according to other principles and standards. Therefore, in the "cloud" indicators should also be formed from the provided information, if required [17-19]. The process of obtaining indicators from other systems, as well as the enrichment of existing ones, should be organized.

4. Conclusion

It is worth noting that the use of deceptive systems as information sources for indicators will increase the accuracy of the IOC because the information comes already after verification in the sandbox. However, this is not enough, because the data can also be received from other means of information protection, and the risk of false “positives” remains, so it is necessary to perform processing of IOC outside the information system and in the cloud storage, as well as to maintain communication between similar systems. It is also worth noting that system architectures can be different, indicating that platforms for cyber threat analysis must be as flexible as possible. Open-source systems can be used to meet the requirements [7-9].

References

[1] Bhardwaj A, Goundars S 2019 A framework for effective threat hunting *Network Security* 6 15–19

[2] Bennis M 2013 When cellular meets WiFi in wireless small cell networks *IEEE Communications Magazine* 51 (6) 44–50

[3] Astely D 2013 LTE Release 12 and beyond *IEEE Communications Magazine* 51 (7) 154–160
[4] Kireeva N V and Karaulova O A 2018 Estimation of network traffic anomalies based on cyclic analysis T-Comm. 12 (11) 28–33
[5] Tumanov D A and Abramov E S 2019 Review of cyber threat analysis platforms Perspektiva-2019 300–305
[6] Ometov A Ya, Andreev S D and Kucheryavyy E A 2015 Algorithm clustering of mobile stations based on spatial and social metrics Information technologies and telecommunications 11 (3) 45–53
[7] Kucheryavyy E A and Samuylov K E 2008 Allocation methodology resources in heterogeneous networks Electrosyvaz 4 34-40
[8] Martyushev N V, Egorov Y P 2003 Determination of the signal strength with the computer analysis of the material structure. Proceedings of the 9th International Scientific and Practical Conference of Students, Post-graduates and Young Scientists - Modern Techniques and Technologies, MTT 2003, 192–194, 1438190
[9] Pashkov E N, Martyushev N V, Masson, I A 2014 Vessel ellipticity and eccentricity effect on automatic balancing accuracy. IOP Conference Series: Materials Science and Engineering 66(1) 012011
[10] Kondrat’ev V V, Govorkov A S, Lavrent’eva M V, Sysoev I A, Karlina A I 2016 Description of the heat exchanger unit construction, created in IRNITU. International Journal of Applied Engineering Research 11(19) 9979–9983
[11] Daus Y V, Yudaev I V 2017 Estimation of solar energy potential under conditions of urban development. Advances in Engineering Research 133 156-161 DOI.org/10.2991/AIME-17.2017.26/
[12] Balanovskiy A E, Astafyeva N A, Kondratyev V V, Karlina A I 2021 Study of mechanical properties of C-Mn-Si composition metal after wire-arc additive manufacturing (WAAM). CIS Iron and Steel Review 22 66–71
[13] Konstantinova M V, Olentsevich A A, Konyukhov V Y, Guseva E A, Olentsevich V A 2021 Automation of failure forecasting on the subsystems of the railway transport complex in order to optimize the transportation process as a whole. IOP Conference Series: Materials Science and Engineering 1064 012020
[14] Konyukhov V Yu, Permyakova D N, Oparina T A 2021 Numerical simulation of the size, quantity and shape of non-metallic inclusions in rails. Journal of Physics: Conference Series 2032 012071
[15] Suslov K, Shushpanov I, Buryanina N, Ilyushin P 2020 Flexible power distribution networks: New opportunities and applications. SMARTGREENS 2020 - Proceedings of the 9th International Conference on Smart Cities and Green ICT Systems, pp. 57–64
[16] Konyukhov V Y, Permyakova D N, Oparina T A 2021 Perspective for the use of industrial waste in lubricating compositions to reduce wear in friction pairs. Journal of Physics: Conference Series 2061 012046
[17] Suslov K, Piskunova V, Gerasimov D, Akhmetshin A, Lombardi P, Komarnicki P 2019 Development of the methodological basis of the simulation modelling of the multi-energy systems. E3S Web of Conferences 124 01049
[18] Konyuuhov V Y, Gorban A V, Gladikh A M 2019 Investment in the improvement of maintenance service efficiency of processing equipment of an industrial enterprise. Journal of Physics: Conference Series 1353 012102
[19] Voropai N, Ukolova E, Gerasimov D, Suslov K, Lombardi P, Komarnicki P. 2019 A Study on Cost-Effectiveness of Energy Supply Based on the Energy Hub Concept. Proceedings of 2019 IEEE PES Innovative Smart Grid Technologies Europe, ISGT-Europe 2019, 2019, 8905736
[20] Ilyushin Y V, Afanaseva O V 2019 Development of scada-model for trunk gas pipeline's compressor station. Journal of Mining Institute 240 686–693
[21] Rapatskaya L A, Buglov N A, Hao D 2013 *The ontogenesis of hydrocarbons in the North-Chinese oil-and gas-bearing basin. Geological Engineering and Mining Exploration in Central Asia*, Proceeding Source: The XVIII Kerulien International Conference on Geology, 2013, pp. 407–410