information

MDPI

Editorial

Dark-Web Cyber Threat Intelligence: From Data to Intelligence to Prediction

Paulo Shakarian

School of Computing, Informatics, and Decision Support Engineering, Arizona State University, Tempe, AZ 85281, USA; shak@asu.edu

Received: 29 November 2018; Accepted: 29 November 2018; Published: 1 December 2018

Scientific work that leverages information about communities on the deep and dark web has opened up new angles in the field of security informatics. The presence of online communities operating with relative impunity allows for data-driven approaches to various forms of adversarial reasoning. Outside of this space, such techniques would require data that are either classified or law-enforcement sensitive.

The pioneering work on dark-web mining by Hsinchun Chen and his group [1] laid the foundations for how dark-web data could impact cyber threat intelligence in a very broad way. We laid out a vision in early 2016 [2] on how this type of data could be leveraged to impact cyber threat intelligence in a variety of ways—from adversarial models, to understanding hacker communities, risk assessment, and data-driven prediction of cyberattacks. We were thrilled at the initial response to some of our early work in this area (i.e., [3]) that coincided with government grants, new scientific studies, and commercial efforts that have only served to help the field.

As the title of this volume suggests, there is an evolution in how the dark web can be used to impact cyber threat intelligence. Simply put, the information must be obtained, analyzed, and potentially used for prediction purposes, all of which poses significant challenges.

First, gathering information from dark-web communities poses a unique set of challenges. Implementing crawlers to gather such information is a complex process. Furthermore, the adversarial nature of the communities from which such data are collected poses a conundrum to researchers: how much detail do they publish? They also run the risk of such techniques ceasing to be viable as they become exposed to potential malicious hackers. These conversations have often taken place at conferences such as ASONAM and IEEE Intelligence and Security Informatics (ISI). For example, Richard Frank's seminal work on dark-web mining [4]—which was named best paper at ASONAM/FOSINT-SI in 2015—led me to engage in a series of conversations with him on many of the challenges he had faced while conducting that research. In this Special Issue, 'A Framework for More Effective Dark Web Marketplace Investigations' provides perhaps the most detailed description of scraping dark-web sites available to-date, offering a detailed case-study that previously researchers could only obtain through offline conversations.

While gathering information is important, data alone cannot address real-world cybersecurity problems. Current threat intelligence organizations at major companies worldwide sift through this data on a regular basis. They map out threat actors, conduct searches relevant to their organization, and synthesize the information across multiple sources. Criminologist Tom Holt was a pioneer in this area (i.e., [5]), which has gained importance recently as Chief Information Security Officers (CISO) are increasingly hiring intelligence processionals. This has led to widespread use of counter-terrorism and law enforcement techniques within operational cybersecurity elements. Techniques such as link analysis are now commonplace within cyber threat intelligence organizations. Research that applies data mining techniques to data obtained from the dark web will enable these threat intelligence teams to create an accurate picture of the threat more quickly. One key challenge is the reconciliation of threat actor identities across multiple sources. In this Special Issue, 'First Steps towards Data-Driven Adversarial Deduplication' addresses this problem head-on.

The current use of dark-web information to support real-world cybersecurity practices has been focused on augmenting intelligence practices. However, with the significant advances in the industry in technology for security information and event management (SIEM), recent work has shown that dark-web indicators can be correlated with event data and used for the prediction of cyberattacks [6]. The paper 'Predicting Cyber-Events by Leveraging Hacker Sentiment' included in this Issue takes the next step in prediction—adding sentiment mining as a prediction element (originally introduced as a way to identify interesting hacker conversations in Reference [4]).

The use of information from hacker communities such as the dark web has great promise in leading to a more threat-focused cybersecurity. The key to further progress in this area is continued evolution and automation so that such threat intelligence can be made available to a wide variety of organizations to drive security decisions and protect their infrastructure more effectively.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Chen, H. *Dark Web: Exploring and Data Mining the Dark Side of the Web*; Springer Science & Business Media: New York, NY, USA, 2011; Volume 30.
2. Shakarian, P.; Shakarian, J. Socio-Cultural Modeling for Cyber Threat Actors. In Proceedings of the AAAI Workshop: Artificial Intelligence for Cyber Security, Phoenix, AR, USA, 12–13 February 2016.
3. Nunes, E.; Diab, A.; Gunn, A.; Marin, E.; Mishra, V.; Paliath, V.; Robertson, J.; Shakarian, J.; Thart, A.; Shakarian, P. Darknet and deepnet mining for proactive cybersecurity threat intelligence. In Proceedings of the 2016 IEEE International Conference on Intelligence and Security Informatics (ISI), The University of Arizona, Tucson, AR, USA, 27–30 September 2016; pp. 7–12.
4. Macdonald, M.; Frank, R.; Mei, J.; Monk, B. Identifying Digital Threats in a Hacker Web Forum. In Proceedings of the 2015 International Symposium on Foundations of Open Source Intelligence and Security Informatics (FOSINT), Paris, France, 26–27 August 2015.
5. Holt, T.J.; Lampke, E. Exploring stolen data markets online: Products and market forces. *Crim. Justice Stud.* **2010**, *23*, 33–50. [CrossRef]
6. Almukaynizi, M.; Paliath, V.; Shah, M.; Shah, M.; Shakarian, P. Finding Cryptocurrency Attack Indicators Using Temporal Logic and Darkweb Data. In Proceedings of the 2018 IEEE Conference on Intelligence and Security Informatics (ISI-18), Florida International University, Miami, FL, USA, 8–10 November 2018.