Research Proposal Transcript

Project Title: Implementing Cyber Threat Intelligence in Dark Web Cybercrime Prediction

By: Xue Ling Teh

[Opening Slide]

Speaker: Good day, thank you for tuning in. My name is Xue Ling, and I will be presenting about implementing Cyber Threat Intelligence, also known as CTI, in predicting cybercrime in the dark web. The cybersecurity industry has its sights on cybercriminal activities, which is especially prevalent in the dark web due to its obscure nature.

[Slide 2: Current Landscape]

Speaker: Cybercrimes are at the forefront of forensics and criminology research, hence the ability to forecast and predict cybercrime by profiling cybercriminals and identifying individual or group motives allows cybersecurity organisations to uphold their reputation, secure the welfare of netizens and protect future victims.

[Slide 3: Research Question]

Speaker: The main research question identified is: How has the implementation of cybersecurity techniques – particularly CTI – contributed towards cybercrime prediction in the dark web and what are potential gaps in mitigating future cybercrime?

[Slide 4: Aims and Objectives]

Speaker: There are three main aims and objectives in this project. Firstly, evaluating the existing literature that closely explores the implementation of CTI in various settings that

relate to predicting cybercrime in the dark web. Secondly, determining the strengths, limitations, and gaps in current research. Thirdly, identifying potential CTI integration challenges in predictive models for combatting cybercrime in the dark web.

[Slide 5: Literature Design]

Speaker: The first section is about existing state-of-the-art literature within the cybersecurity discipline, research design, strategy and approaches.

[Slide 6: Key Literature]

Speaker: Key literature in terms of. In a conference by Schäfer et al. in 2019, cyber intelligence is sourced using BlackWidow as a framework to automate real-time monitoring for the exploitation of leaked information. Through a five-phase process, requirements are manually gathered via scraping dark web forums, with the rest being automated. This includes anonymously collecting raw data, processing via parsing HTML, foreign language translation using Google's API, and extracting relevant data to form a connected knowledge graph consisting of topics, messages, threads and actors. Through threads analysis, BlackWidow infers actor relationships, topics and cybersecurity trends using machine learning techniques such as unsupervised text clustering for message classification.

[Slide 7: Methodology]

Speaker: A mixed approach of quantitative and qualitative in data analysis will be applied depending on the types of data collected. Secondary quantitative and qualitative analysis are gathered from literature review (Tate & Happ, 2018), whereas primary research is sourced from conducting surveys and expert interviews. Quantitative analysis will involve

evaluating historical cybercrime data and CTI impact metrics. On the other hand, qualitative analysis will involve expert interviews to explore the effectiveness of criminal profiling in cybercrime prediction. Any data collected in the form of unstructured text, including interview transcripts would be coded using the qualitative analysis coding technique written by Adu in 2019. By applying these data analysis techniques, meaningful insights can be extracted.

[Slide 8: Security Compliance]

Speaker: This section highlights the aspects of security compliance, including ethics approval and risk management.

[Slide 9: Ethical Considerations]

As part of the ethical approval application process, the method of collecting and storing participants' data are specifically listed as well as indicating their purposes. The statement in the Guide to the General Data Protection Regulation (GDPR) by the Information Commissioner's Office, any data collected from European Union citizens or residents for this project research purposes must adhere to GDPR.

[Slide 10: Deliverables]

Speaker: Lastly, this section describes the deliverables and outcomes of the project, which include the proposed applicable artefacts and project timeline.

[Slide 11: Applicable Artefacts]

Speaker: The two major applicable artefacts are the coded qualitative analysis data which contains theme and keyword analysis from unstructured data and a CTI knowledge graph compilation containing relationships between actors and topics.

[Slide 12: Proposed Timeline]

Speaker: According to Dawson (2015), a project plan is best represented by a Gantt chart. The proposed timeline details the following three main phases, which are project initiation and planning, data collection and data analysis, as well as report writing. The duration of work days is an estimation and may vary according to the situation such as availability of interviewees. The first phase would take approximately one month, second phase and third phase about two months each, with overlapping timeline between certain activities. In the first phase, the activities are defining project objectives and research question, conducting literature review, and applying for approval and ethical clearance. For the second phase, the activities are as follows: collecting historical dark web cybercrime data, gathering relevant CTI sources, preprocessing and cleaning quantitative data, and coding qualitative data. Finally, there is identifying gaps in current research, compiling research findings and report writing.

[Slides 13-15: References]

Speaker: Thank you for your time. Here is the list of references.

References

Adu, P. (2019) *A step-by-step guide to qualitative data coding*. Routledge.

Basheer, R. & Alkhatib, B. (2021) Threats from the dark: A review over Dark Web Investigation Research for Cyber Threat Intelligence, Journal of Computer Networks and Communications. *Journal of Computer Networks and Communications* 2021: 1-21. DOI: https://doi.org/10.1155/2021/1302999

Bezerra, J., César, C., de Souza, N., & Hirata, C. (2020) Extending STPA with STRIDE to identify cybersecurity loss scenarios. *Journal of Information Security and Applications*, 55: 1-13.

Dawson, C. (2015) *Projects in Computing and Information Systems: A Student's Guide*. Harlow: Pearson.

Horan, C. & Saiedian, H. (2021) Cyber crime investigation: Landscape, challenges, and future research directions. *Journal of Cybersecurity and Privacy* 1(4): 580-596. DOI: https://doi.org/10.3390/jcp1040029

Information Commissioner's Office (ICO) (N.D.) Guide to the General Data Protection Regulation (GDPR).

Open Web Application Security Project (OWASP). (2021) OWASP Top Ten. Available from: https://owasp.org/www-project-top-ten/ [Accessed 10 August 2023].

Schäfer, M., Fuchs, M., Strohmeier, M., Engel, M., Liechti, M. & Lenders, V. (2019) May. BlackWidow: Monitoring the dark web for cyber security information. *2019 11th International Conference on Cyber Conflict (CyCon).* 28-31 May. IEEE. 1-21.

Tate, J.A. & Happ, M.B. (2018) *Qualitative Secondary Analysis: A case exemplar*, *Journal of pediatric health care : official publication of National Association of Pediatric Nurse Associates & Practitioners*. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5911239/ [Accessed 18 August 2023].