

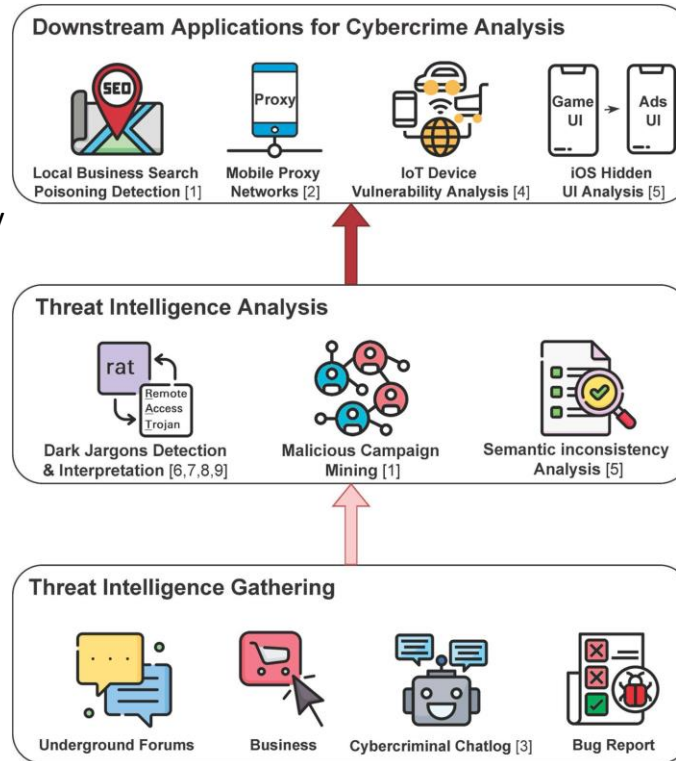
SaTC: CORE: Medium: Collaborative: Understanding and Discovering Emerging Cybercrimes through Automatic Analysis of Online Text Traces

Challenge:

- Innocent-looking deceptive content can be easily blended into legitimate traces
- Encoded words such as dark jargons are extensively used by cybercriminals
- Hard to interpret the text information and extract actionable knowledge

Solution:

- Applying NLP techniques to discover the semantic discrepancies in the cybercrime text content
- ML techniques can help discover in-depth knowledge from the text traces



Scientific Impact:

- A critical step for more intelligent and automated defense against Cybercrimes based on text content
- Result in more in-depth understanding of the dark business world, enabling identification of its key weaknesses and rapid responses to new threats

Broader Impact:

- Contribute to new interdisciplinary research on applying NLP and learning techniques to support intelligent security protection
- Outcomes of the project are widely disseminated
- Involve HBCU students
- Organize semantics-aware security challenge

Project number: 1801432, 1850725, 1801652
Contacts: XiaoFeng Wang*, Xiaojing Liao*, Chengxiang Zhai+
* Indiana University Bloomington
+ University of Illinois at Urbana-Champaign



Selected Publications

- **[1] Demystifying Local Business Search Poisoning for Illicit Drug Promotion**
Peng Wang, Zilong Lin, Xiaojing Liao, XiaoFeng
in Proceeding of ISOC Network and Distributed System Security Symposium (NDSS), 2022.
- **[2] Your Phone is My Proxy: Detecting and Understanding Mobile Proxy Networks**
Xianghang Mi, Siyuan Tang, Zhengyi Li, Xiaojing Liao, Feng Qian, XiaoFeng Wang
in Proceeding of ISOC Network and Distributed System Security Symposium (NDSS), 2021
- **[3] Into the Deep Web: Understanding E-commerce Fraud from Autonomous Chat with Cybercriminals**
Peng Wang, Xiaojing Liao, Yue Qin, XiaoFeng Wang
in Proceeding of ISOC Network and Distributed System Security Symposium (NDSS), 2020
- **[4] Understanding and Securing Device Vulnerabilities through Automated Bug Report Analysis**
Xuan Feng, Xiaojing Liao, XiaoFeng Wang, Haining Wang, Qiang Li, Kai Yang, Hongsong Zhu, Limin Sun
in Proceeding of USENIX Security Symposium (Security), 2019
- **[5] Understanding iOS-based Crowdturfing through Hidden UI Analysis**
Yeonjoon Lee, Xueqiang Wang, Kwangwuk Lee, Xiaojing Liao, XiaoFeng Wang, Tongxin Li, Xianghang Mi
in Proceeding of USENIX Security Symposium (Security), 2019
- **[6] Reading Thieves' Cant: Automatically Identifying and Understanding Dark Jargons from Cybercrime Marketplaces**
Kan Yuan, Haoran Lu, Xiaojing Liao, XiaoFeng Wang.
in Proceeding of USENIX Security Symposium (Security), 2018
- **[7] DarkJargon.net: A Platform for Understanding Underground Conversation with Latent Meaning**
Dominic Seyler, Wei Liu, Yunan Zhang, XiaoFeng Wang, and ChengXiang Zhai
in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2021.
- **[8] Towards dark jargon interpretation in underground forums**
Seyler, Dominic, Wei Liu, XiaoFeng Wang, and ChengXiang Zhai.
In *European Conference on Information Retrieval*, pp. 393-400. Springer, Cham, 2021.
- **[9] A study of methods for the generation of domain-aware word embeddings**
Seyler, Dominic, and ChengXiang Zhai.
In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2020.