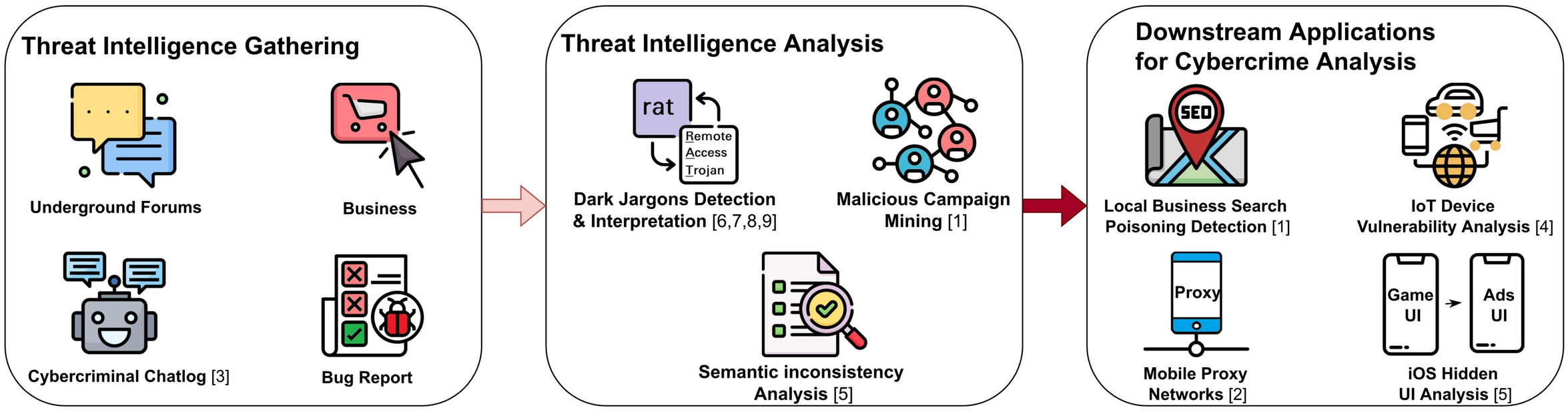
SaTC: CORE: Medium: Collaborative:

Understanding and Discovering Emerging Cybercrimes through Automatic Analysis of Online Text Traces

Project number: 1801432, 1850725, 1801652

Contacts: XiaoFeng Wang*, Xiaojing Liao*, Chengxiang Zhai+

* Indiana University Bloomington, + University of Illinois at Urbana-Champaign



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

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

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

Broader Impact:

- Will contribute to new interdisciplinary research on applying NLP and learning techniques to support intelligent security protection
- Outcomes of the project can be transferred to the industry partners
- Working closely to involve HBCU students
- Organizing an annual semantics-aware security challenge

Publications:

[1] Wang, et al. Demystifying Local Business Search Poisoning for Illicit Drug Promotion.NDSS, 2022.

[2] Mi, et al. Your Phone is My Proxy: Detecting and Understanding Mobile Proxy Networks. NDSS, 2021

[3] Wang, et al. Into the Deep Web: Understanding E-commerce Fraud from Autonomous Chat with Cybercriminals. NDSS, 2020

[4] Feng, et al. Understanding and Securing Device Vulnerabilities through Automated Bug Report Analysis. USENIX Security, 2019

[5] Lee, et al. Understanding iOS-based Crowdturfing through Hidden UI Analysis. USENIX Security, 2019

[6] Yuan, et al. Reading Thieves' Cant: Automatically Identifying and Understanding Dark Jargons from Cybercrime Marketplaces. USENIX Security, 2018

[7] Seyler, et al. DarkJargon.net: A Platform for Understanding Underground Conversation with Latent Meaning. SIGIR, 2021.

[8] Seyler, et al. Towards dark jargon interpretation in underground forums. In European Conference on Information Retrieval,2021.

[9] Seyler, et al. A study of methods for the generation of domain-aware word embeddings. SIGIR, 2020.