AI-SNIPS: A Platform for Network Intelligence-Based Pharmaceutical Security

Timothy A. Burt[1,2], **Nikos Passas**[3], **Ioannis A. Kakadiaris**[1,2,4]

[1]Computational Biomedicine Lab (CBL), University of Houston, Houston, TX USA
[2]Dept. of Physics, University of Houston, Houston, TX USA
[3]School of Criminology and Criminal Justice, Northeastern University, Boston, MA USA
[4]Dept. of Computer Science, University of Houston, Houston, TX USA
ioannisk@uh.edu

Abstract

This paper presents **AI-SNIPS** (**AI S**upport for **N**etwork **I**ntelligence-based **P**harmaceutical **S**ecurity), a production-ready platform that enables stakeholder decision-making, secure data sharing, and interdisciplinary research in the fight against Illicit, Substandard, and Falsified Medical Products (ISFMP). AI-SNIPS takes as input cases: a case consists of one or more URLs suspected of ISFMP activity. Cases can be supplemented with ground-truth structured data (labeled keywords) such as seller PII or case notes. First, AI-SNIPS scrapes and stores relevant images and text from the provided URLs without any user intervention. Salient features for predicting case similarity are extracted from the aggregated data using a combination of rule-based and machine-learning techniques and used to construct a seller network, with the nodes representing cases (sellers) and the edges representing the similarity between two sellers. Network analysis and community detection techniques are applied to extract seller clusters ranked by profitability and their potential to harm society. Lastly, AI-SNIPS provides interpretability by distilling common word/image similarities for each cluster into signature vectors. We validate the importance of AI-SNIPS's features for distinguishing large pharmaceutical affiliate networks from small ISFMP operations using an actual ISFMP lead sheet.

Introduction

Illicit, Substandard, and Falsified Medical Products (ISFMP) damage public health & welfare on a global scale: lower bound estimates predict ISFMP induce fatal pneumonia between 72,000 – 169,000 children each year, and falsified anti-malarial medicines are responsible for an additional 116,000 deaths (OECD/EUIPO 2020). ISFMP include stolen, diverted, price-gouged, unregistered, unlicensed, and counterfeit medical products; examples include toxic or ineffective prescription drugs, dietary supplements, face masks, vaccines, and testing kits. ISFMP damage companies' brands, undermine competition and the rule of law, and support other illicit activities such as terrorism, human trafficking, and money laundering (Zaman 2018; Passas 2000; Cesareo 2016; Accri 2018). The increased complexity and volume of ISFMP sales resulting from the COVID-19 pandemic continues to challenge the resources of many teams investigating these crimes, necessitating new platforms for AI-enabled stakeholder decision support which do not act as "black boxes."

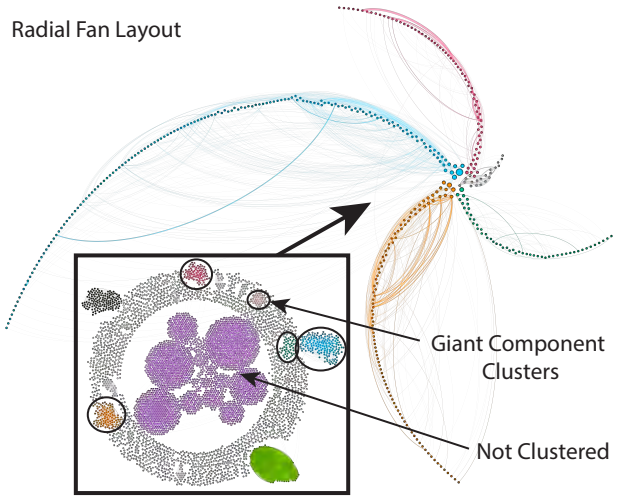

Figure 1: Visualization showing layouts for the discovered Company X Network. The largest connected component in the network is arranged as a spiral, in which each of the five fans represents a cluster. Nodes (leads) are ordered such that the ones with the highest betweenness centrality are at the center of the spiral (betweenness centrality quantifies how important a node is to the flow of the network).

AI-SNIPS is designed to help risk/fraud control analysts, law enforcement/FDA analysts, and platform risk managers identify, prioritize, and solve online crimes related to the sale of ISFMP. To our knowledge, AI-SNIPS is the first non-proprietary platform in the literature that provides a scalable network intelligence tool for research and analysis of ISFMP activities. The AI-SNIPS data analytics pipeline is shown in Figure 2.

Related Work

Early warning systems which utilize graph representations and community detection show promise for analyzing social media events (Lu et al. 2022) and news streams (Vine et al. 2022); both are forms of Open-Source Intelligence

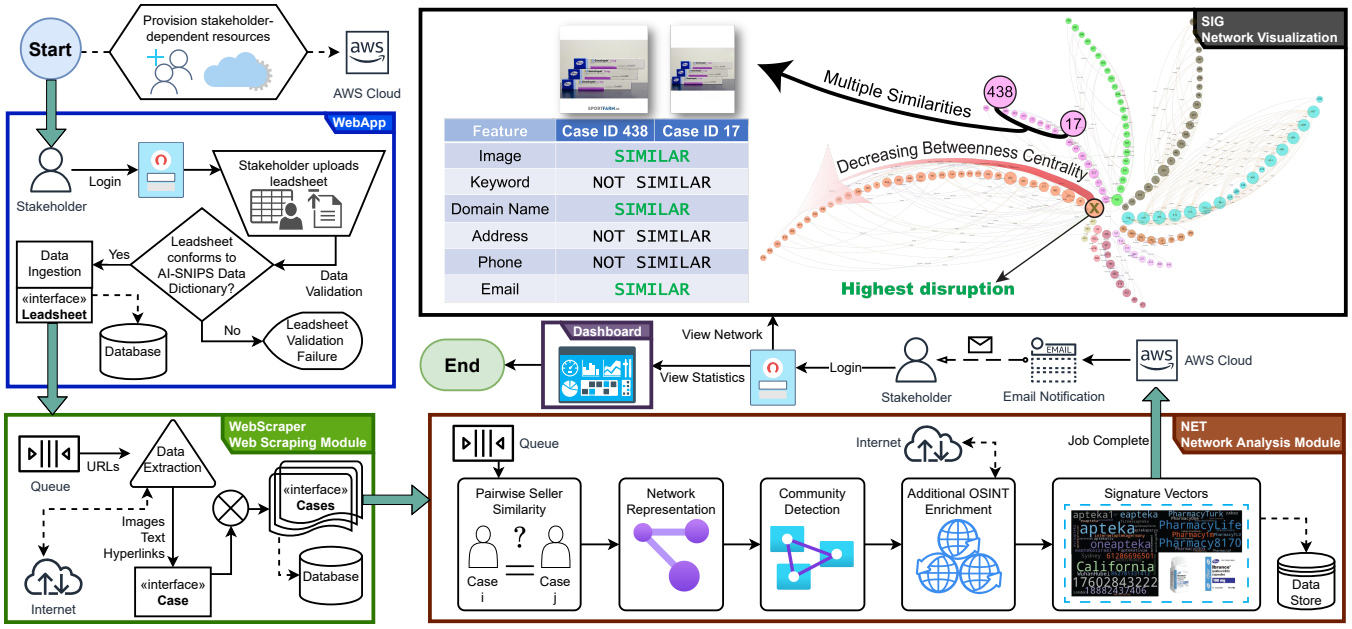

Figure 2: Prototype diagram of the AI-SNIPS platform utilizing AWS as the cloud service provider.

(OSINT). OSINT refers to data and information available to the public. Maybir and Chapman (2021) recently demonstrated the use of OSINT for monitoring drug market trends in real time. Ecstasy pill characteristics and regional usage trends within Australia were scraped from an online forum and found to be consistent with more costly and coarse-grained approaches such as drug use surveys and wastewater analysis.

A retrospective analysis of pharmaceutical affiliate programs using leaked financial transactions was published by McCoy et al. (2012). These affiliates often utilize spamming tactics and are some of the longest-running (and most profitable) illegal online businesses; examples include SpamIt and Canadian Pharmacy.

The AI-SNIPS Research Team presented the AI-SNIPS Proof of Concept at the APS March Meeting (Burt et al. 2022). Additionally, our team recently published a Q&A format newsletter article (Burt, Passas, and Kakadiaris 2022) outlining how AI-SNIPS can be utilized to solve challenges in supply chain security.

Applied Methodology

To illustrate how stakeholders use AI-SNIPS to solve online ISFMP cases, we use AI-SNIPS to process an actual ISFMP lead sheet provided to us by Company X, a large pharmaceutical company, who tasked us with finding the "big fish" in the pond.

AI-SNIPS identified five clusters in the largest connected component; each cluster predominantly consisted of known spammers/illicit pharmaceutical affiliate programs (Figure 1), despite AI-SNIPS having no prior knowledge about pharmaceutical affiliate networks. About two-thirds of the clustered leads consisted of two sellers, which were not the targets for this use case.

A review of the signature vectors and seller domains unraveled the inner workings of the illicit operation: the cluster with the most cases (seen as a fan) consisted of B2C affiliate sites related to Canadian Pharmacy. The second cluster revealed product items sold on legitimate B2B platforms such as Alibaba. The third cluster was composed of a different affiliate pharmacy (exposing a possibly unknown link). The remaining two clusters revealed that these items are being sold on social media platforms (illegally).

The reason for connections between clusters is interpretable: each edge carries one or more image/text data pairs, a similarity score, and a confidence index. These metrics allowed analysts to form a Preliminary Evaluative Opinion on whether the cases (and their clusters) are "similar enough" to proceed with that line of investigation. Our stakeholders can then coordinate simultaneous disruption of seller activity across multiple vulnerable domains, inflicting the most damage to the illicit supply chain's operations.

Significance and Future Work

This work presents AI-SNIPS, a modular web platform with services for web scraping, real-time URL monitoring, network forensics & visualization, counterfeit product detection, and pharmaceutical affiliate network risk assessments. A longer format paper is currently in preparation which provides the full implementation details of AI-SNIPS for reproducibility and an in-depth evaluation of the various network quantities produced by AI-SNIPS.

Future work will add research modules and expand its data dictionary with requests from new stakeholders. In addition, we will incorporate temporal information into the graph, which is needed to validate the impact of coordinated disruption strategies.

Acknowledgments

We thank the undergraduate students Noah Alexander, Jennifer Csicsery-Ronay, and Elisa Martinez, who participated in the project. We also thank all of the stakeholders for their time and data, without which this project would not be possible. The first author was supported by the Department of Defense SMART Scholarship-for-Service Program. The work of Passas and Kakadiaris was supported by the National Science Foundation Award IIS-2039946. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors. They do not necessarily reflect the views of the National Science Foundation, other funders, the position, or the policy of the Government, and no official endorsement should be inferred.

References

Accri, K. 2018. Pharmaceutical Counterfeiting: Endangering Public Health, Society, and the Economy. Technical report, Fraser Institute.

Burt, T. A.; Passas, N.; and Kakadiaris, I. A. 2022. A New Analytical Approach and Visualization of Online Sales of Fake Products. *TAPA: The Quartely Update for the Americas*, 2022(Q2): 11–14.

Burt, T. A.; Sundaram, R.; Passas, N.; Amiji, M.; Zaman, M.; and Kakadiaris, I. A. 2022. Towards Dismantling Healing Illicit & Counterfeit Medicines Seller Networks (ICMSN) Using Percolation Theory & Machine Learning: A Simulation Study. In *Bull. Am. Phys. Soc.*, volume 67. Chicago, IL: American Physical Society.

Cesareo, L. 2016. *Counterfeiting and Piracy*. SpringerBriefs in Business. Cham: Springer International Publishing.

Lu, N.; Yang, Z.; Huang, J.; Wu, Y.; and Wang, H. 2022. Silence or Outbreak – a Real-Time Emergent Topic Identification System (RealTIS) for Social Media. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(11): 13194–13196.

Maybir, J.; and Chapman, B. 2021. Web Scraping of Ecstasy User Reports as a Novel Tool for Detecting Drug Market Trends. *Forensic Science International: Digital Investigation*, 37: 301172.

McCoy, D.; Pitsillidis, A.; Grant, J.; Weaver, N.; Kreibich, C.; Krebs, B.; Voelker, G.; Savage, S.; and Levchenko, K. 2012. PharmaLeaks: Understanding the Business of Online Pharmaceutical Affiliate Programs. In *Proc. USENIX Conf.*, 1–16. Bellevue, WA: USENIX Association.

OECD/EUIPO. 2020. *Trade in Counterfeit Pharmaceutical Products*. Illicit Trade. Paris: OECD Publishing.

Passas, N. 2000. Global Anomie, Dysnomie, and Economic Crime: Hidden Consequences of Neoliberalism and Globalization in Russia and Around the World. *Social Justice : a Journal of Crime, Conflict and World Order*, 27(2): 16–44.

Vine, N. L.; Boxer, E.; Dinani, M.; Tortora, P.; and Das, S. 2022. Identifying Early Warning Signals from News Using Network Community Detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(11): 12378–12386.

Zaman, M. H. 2018. *Bitter Pills: The Global War on Counterfeit Drugs*. Oxford University Press.