Unmasking sex trafficking supply chains with machine learning

P. Ramchandani[*,1], H. Bastani[*,1], E. Wyatt[2]

[1]Department of Operations, Information and Decisions, Wharton School, University of Pennsylvania

[2]TellFinder Alliance

[*]Corresponding authors. Email: piar2@wharton.upenn.edu, hamsab@wharton.upenn.edu

According to the FBI, sex trafficking is the fastest growing organized crime business and the third largest criminal enterprise in the world [1]. The International Labor Organization (ILO) estimates there were 4.8 million sex trafficking victims in 2017 alone [2]. Consequently, there is high demand for a *large-scale* and *data-driven* view of the underlying supply chain dynamics [3] of trafficking that can inform law enforcement, policy and social work [4]. For instance, understanding how victims are recruited in different regions can enable *preventative* interventions at the source of the supply chain (recruitment) in contrast to prevalent mitigation strategies that target the end of the supply chain (sales). Furthermore, identifying recruitment-to-sales trafficking routes of individual criminal organizations can enhance coordination strategies between relevant agencies and task forces to increase efficiency of counter-trafficking efforts [3].

However, the covert nature of trafficking provides a significant barrier to generating such insights. For example, less than one third of the existing research literature on sex trafficking uses any data, and those that do primarily leverage qualitative interviews with trafficking survivors [5]. It is hard to generate quantitative and generalizable insights from such interviews, because they are qualitative in nature and severely limited in scale; moreover, they can be traumatic for victims and can result in unreliable information.

Through a collaboration with the TellFinder Alliance Global Counter-Human Trafficking Network, we leverage unstructured, massive *deep web data* from leading adult-services websites using a novel machine learning framework to construct the first global network view of sex trafficking supply chains. The deep web represents portions of the World Wide Web that are not indexed by traditional search engines, e.g., search results from private websites, which contain relevant but temporary information. A significant portion of sex trafficking activity occurs online, making the deep web a rich and relevant data source. For example, consider the following recruitment ad targeting Covenant House (a shelter for runaway kids) that was found online by New York City law enforcement:

"Want to stop living in Covenant House shelter? Need Free room & board? Employment & Financial assistance? Room available immediately! Should have an interest in escort business making $1000+ weekly."

The person who posted the ad was later found to be trafficking respondents and selling them for prostitution [6]. This behavior is common, e.g., 1 out of 5 homeless girls in the United States end up being sex trafficked [6]. Indeed, the majority of trafficking victims are recruited through a deceptive employment offer made online [7].

Recent work has developed techniques to scrape deep web data, extract relevant meta data (e.g., phone numbers, email addresses) and convert it into databases that support trafficking investigation inquiries by law enforcement agencies [8]. However, such data has not been used for large-scale analysis of sex trafficking supply chains, primarily due to the difficulty in extracting meaningful supply chain relationships from unstructured text. The key challenge involves distinguishing content that seeks to recruit unsuspecting victims: while sex sales are prevalent and convey clear intent to consumers, recruitment efforts are sparse and designed to trick potential victims into being trafficked, making the latter hard to identify.

Broadly speaking, our goal is to understand the structure of the underlying supply chain network. We apply our approach to 14 million English-language posts scraped from the deep web using the Tellfinder tool over a 9-month period from leading adult-services websites with global geographic coverage. Our first step is to classify posts based on whether they are attempting to recruit victims or sell sex using the unstructured text in the post. We face two challenges:

1. **Extreme Data Imbalance:** We estimate that only 0.06% of posts are related to recruitment, while the rest are sales, i.e., one would have to manually label nearly 2000 randomly chosen posts to find a single instance of recruitment with reasonable likelihood. Thus, traditional supervised or active learning techniques, which rely on an initial well-balanced training set, are infeasible. We resolve this issue using weak learning in conjunction with active learning.
2. **Objective Mismatch:** We seek to identify different recruitment approaches across many locations. For instance, one auxiliary task is to identify pairs of posts (one recruitment and one sales) in different locations that are linked to the same organization by their meta data; such a pair corresponds to a potential edge in the supply chain network. Thus, traditional learning techniques that focus purely on overall accuracy may be insufficient. We resolve this issue by adjusting the active learning procedure to incorporate metrics that are relevant to our analysis.

Our machine learning pipeline proceeds roughly as follows. First, we train domain-specific word embeddings and collect expert-identified terms (e.g., “escort” or “sugar daddy” based on past trafficking investigations and conversations with domain experts) to develop an initial ‘recruitment vocabulary.’ This informs a weak learning heuristic to identify an initial well-balanced training set. However, this training set is clearly biased by the purview of domain experts and does not provide a complete view of the numerous templates of recruiting posts on the deep web. Thus, we use pool-based active learning, which is known to improve classifiers with significantly reduced manual labeling effort. Specifically, we iteratively label additional posts and update our deep learning predictive model until its performance converges. We address objective mismatch by tailoring the prioritization function to additionally incorporate geographic diversity and the likelihood of identifying new network connections. Through this process, we manually labeled ~60,000 posts, and identified ~8000 recruiting posts. Note that this corresponds to 13% of the labels being positive, compared to only 0.06% of the labels being positive on a random subsample of the data; furthermore, the active learning process allowed us to uncover far more recruiting templates beyond the initial expert-identified vocabulary. Finally, we connect identified recruitment and sales content via shared meta data to identify supply chain networks.

Our results yield the first view into sex trafficking supply chains of this scale that can be used for policy-relevant insights, including identifying common trafficking routes, pinpointing popular approaches used to recruit victims in specific locales, and discerning regional variations in recruitment vs. sales pressure. For instance, Fig. 1 shows common recruitment-to-sales trafficking routes identified by our approach. Note that coverage in some parts of the world is limited by our restriction to posts in the English language; future work can adopt our approach to other languages to improve global coverage.

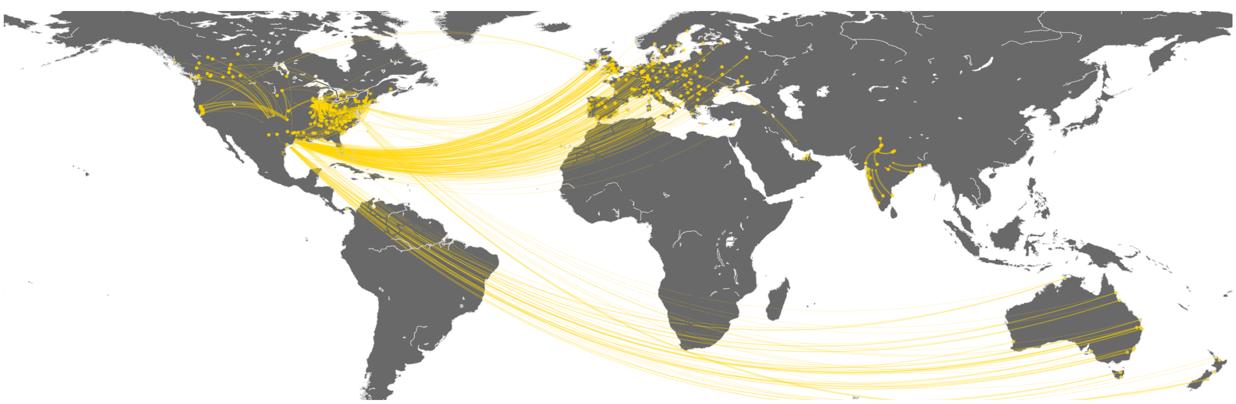

Figure 1 *We provide the first large-scale, global view of sex-trafficking supply chain networks using our novel machine learning framework. Edges are based on aggregating regional connections between recruitment and sales activity for individual criminal organizations. Note that coverage in some parts of the world is limited by our restriction to posts in the English language.*

Bibliography

[1] A. Walker-Rodriquez and R. Hill, "Human Sex Trafficking," FBI: Law Enforcement Bulletin, 2011.

[2] "Human trafficking by the numbers," International Labor Organization, 2017.

[3] J. Roby and M. Vincent, "Federal and state responses to domestic minor sex trafficking: the evolution of policy," Social Work, 2017.

[4] J. Orme and F. Ross-Sheriff, "Sex trafficking: Policies, programs, and services," Social Work, 2015.

[5] D. Okech, Y. Joon Choi, J. Elkins and A. Burns, "Seventeen years of human trafficking research in social work: a review of the literature.," Journal of Evidence-Informed Social Work, 2018.

[6] L. Brody, "OneZero: How Artificial Intelligence is Tracking Sex Traffickers," May 2019. [Online]. Available: https://onezero.medium.com/amp/p/276dcc025ecd.

[7] K. Kotrla, "Domestic minor sex trafficking in the United States," Social Work, 2010.

[8] Ce Zhang, et. al., "DeepDive: declarative knowledge base construction," *Communications of the ACM,* 2017.