Covert Online Ethnography and Machine Learning for Detecting Individuals at Risk of Being Drawn into Online Sex Work

Panos Kostakos
Center for Ubiquitous Computing
University of Oulu
Oulu, Finland
panos.kostakos@oulu.fi

Lucie Špráchalová
Department of Sociology
Charles University, Prague
Prague, Czech Republic
lucie.sprachalova@gmail.com

Abhinay Pandya
Center for Ubiquitous Computing
University of Oulu
Oulu, Finland
abhinay.pandya@oulu.fi

Mohamed Aboeleinen
Center for Ubiquitous Computing
University of Oulu
Oulu, Finland
mohamed.aboeleinen@student.oulu.fi

Mourad Oussalah
Center for Ubiquitous Computing
University of Oulu
mourad.oussalah@oulu.fi

Abstract - How can we identify individuals at risk of being drawn into online sex work? The spread of online communication removes transaction costs and enables a greater number of people to be involved in illicit activities, including online sex trade. As a result, social media platforms often work as springboard for criminal careers posing a significant risk to the economy, public health and trust. Detecting deviant behaviors online is limited by the poor availability of ground-truth data and machine learning tools. Unlike prior work which focuses exclusively on either qualitative or quantitative methods, in this paper we combine covert online ethnography with semi-supervised learning methodologies, using data from a popular European adult forum. We obtained risk assessment results of 78 users using covert online ethnography, and set out to build a machine learning model that can predict the risk factor in other 28,832 users. Results show that a combination-based approach in which all features are used yields the most accurate results.

Keywords— Cover online ethnography, semi-supervised machine learning, crime detection, online sex work.

I. INTRODUCTION

How can we identify individuals at risk of being drawn into online sex work? This paper explores the use of semi-supervised learning methodologies, covert online ethnographies, social network analysis, and social data mining to detect illicit sex services in online adult forums. The proliferation of online communication reduces transaction costs and enables a greater number of ordinary people to turn into *indoor sex* workers. Prior research has identified specialized adult chat forums as market aggregators [1-2]. The spread of unregistered sex workers has far-reaching implications for economic development, social cohesion, and public health. We present a proof-of-concept tool that uses semi-supervised learning for detecting individuals at risk of being drawn into sex work online. We scraped over 28,000 profiles from a public adult forum which is suspected to enable unregistered sex transactions. The database is freely available in csv format [3]. To account for the lack of labeled data, we conducted covert online ethnography for over twelve months and carried out non-participant observations and semi-structured interviews with a convenient sample of users (N=78). Based on self-reported data, chat transcripts, and notes from lurking, we labeled a small number of users based

the risk of their being involved in illicit transactions. Subsequently, we trained a semi-supervised learner on labeled and unlabeled data using a mixture of textual and numeric features. Results demonstrate that textual features like personal descriptions improve the overall performance, while node-based metrics add very little to the overall score.

II. RELATED WORKS

Scientific understanding of sex-work has increased by an order of magnitude in the past few decades (Fig. 1). This trend is in line with other technological, economic, and demographic changes that are taking place in modern metropolitan cities [4]. In this section, we review the negative impact of illicit sex work and outline the main influential factors cited in the literature. We then move on to explore the application of online ethnography as a tool for studying deviant online communities.

A. Impact of Illicit Sex Markets

Sex work is often perceived as a challenge to traditional societal norms, values, and institutions. Consequently, illicit sex work such as prostitution is very often studied in the context of declining social, economic, and health indicators. For brevity, in the section below, we highlight the major themes and topics discussed in the literature.

Economic impact. The negative economic impact of illicit sex markets is by far the strongest and most commonly cited topic in academic and grey literatures [5 - 11]. For one thing, prostitution is often associated with tax evasion, a legal consideration. The research emphasizes that sex workers do not "declare their actual working hours or income correctly, in order to avoid paying taxes and thus to retain a higher net income" [12]. Another commonly discussed topic is the link between sex work and organised crime, with protection, money laundering, and human trafficking for prostitution being some ways in which organized crime is a part of sex trafficking [11].

Public Health. One of the most discussed issues concerning prostitution and sex work is sexually transmitted disease (STDs) and the risk of the spread of these diseases. Other illnesses are associated with the practice, such as a higher risk of cervical cancer; unwanted pregnancies are another risk of sex work [13,10]. Prostitutes are at a higher

risk than the general population for violence (injuries and batteries) [13], assault, verbal threats of physical assault [7], rape, unwanted sex, and murder [6,10]. The eroticization of violence and humiliation is another emerging health risk [5]. As pointed out in [7], there is an increasing market for more extreme sexual acts in prostitution transactions (beating, torture or humiliation during sexual activities), all of which put the workers at a higher risk of physical injury or health problems.

Psychology. In the field of mental health, sex work is positively associated with depression, drug use, post-traumatic stress disorders [7], and other types of trauma [5] as well as problems such as anxiety, anhedonia, impairment of sexual pleasure [10], depersonalization [5], fear of recognition, and self-hatred [5,10]. A large number of sex workers are under the control of third parties such as pimps [7], and this means that they are not free agents and unable to keep all of their pay, as they have to pay a part of their money for protection, marketing, and housing. Also, sex-workers are exposed to cyberbullying, theft of personal data online, the posting of private pictures of them online or cruel comments made about them in social media, all of which can have an adverse effect on the worker [5].
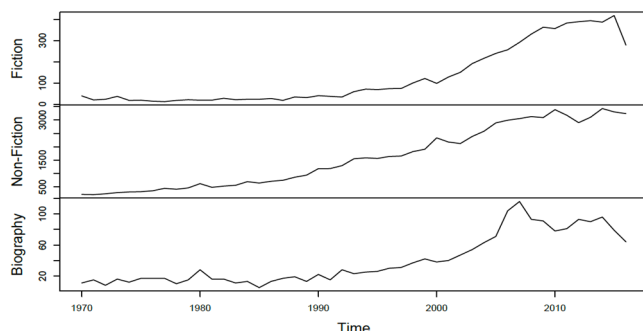

Fig. 1. Number of publications on prostitution per genre (1970-2016). Source: WorldCat.org

B. Online ethnography

With human activity increasingly taking place in online environments, there is an increasing need to study human behaviours with computational tools. Online ethnography is often defined as the adaptation of "ethnographic research techniques to the study of communities emerging through cyber mediated communication" [14]. Online ethnography (covert and overt) has been used widely to study chat forums and websites with sexual and erotic content, to conduct sensitive research [15; 16], and to identify deviant online behaviours [17].

A number of studies have used online ethnography as a method to investigate online sexual behaviors and market dynamics [18, 19]. It has been pointed out [1] that "internet-mediated prostitution is a suitable model system to study emergent properties of socioeconomic feedback". Some researchers also emphasize that non-sexual websites are often used by sex workers to meet new customers [2]. Below we list some key advantages and disadvantages in using an online ethnography methodology for studying online sexual behaviors and networks.

Advantages: flexibility and adaptability, public availability, broader cohort of respondents, ability to reach geographically unreachable respondents, wide range of topics,

naturalistic data, particularized and specific understanding, immersive data that provide local knowledge, detailed/rich data which are easy to obtain, simple research, research of meanings and patterns, symbolism, community trends, longitudinal access and real-time trends, archived data, time saved, faster and more economically viable, suitable technique to study online cultures and communities, application oriented and contextually oriented research, the ability to orient to experienced and highly involved users [14].

Disadvantages: Extremely difficult measurement (numbers of postings, messages, websites, boundaries of studied field or group); constant movement of online space; research affected by researcher's skills, interest, and insight; cultural diversity; inability to generalize results; only text-based research; authenticity and missing cues (non-verbal cues such as emotions) [21].

C. Covert online ethnography.

While overt and covert online ethnographies differ in various ways from each other, the latter is used for gaining more naturalistic insights into online communities, and it is used especially to study subjects or groups whom are difficult to approach overtly [17]. Building on the strengths and weaknesses of the literature, the paper combines cyberstealth techniques (covert ethnography) and machine learning to enable improved detection of deviant behaviors in an online dating forum [22].

III. RESEARCH METHODS

In this section, we describe the data sources and the key concepts that inform the analysis used in the paper.

A. Ethics statement

All of the data collected and used in the analysis are publicly available data; therefore, approval and informed consent were not required. The profile pages of the users from which data were extracted are publicly available and accessible to everyone. The user-generated content provided in profiles is also publicly available unless the user specifies otherwise in his/her privacy settings, and in that case, the data was not available to us. This paper does not quote any of the qualitative data extracted from lurking; therefore, permission and/or informed consent was not required. Our data extraction methods do not violate the website's terms of service. The complete anonymized dataset has been made available to the scientific community for use in crime detection research [3].

A. Data and Concepts

We selected a popular European adult dating forum as our ethnographic research site based on prior knowledge. We used social data mining to collect quantitative data for each user's profile page (N=28,832) and covert online ethnography to collect qualitative data on users' self-reported behavior regarding paid sex (N=78).

B. The adult forum

The forum we observed is the largest community-based website for adults with BDSM preferences in a European country. We have chosen this forum because of its popularity and membership. All offers and payment for sexual activities are strictly banned, and their detection in user-to-user communication leads administrators to cancel the profile of the user. While such forums have strong informal social-

control mechanisms, commodification of sexual services is expected to be high. We deployed a web crawler and collected the following data from 28,832 profiles:

- **URL**: The URL of the profile; **Name**: Nickname; **Gender**: Male/female; **Location**: Country of origin; **Picture URL**: The URL of each avatar.

- **Verification**: Whether the user was online or offline

- **Sexual orientation**: Options include heterosexual, homosexual, bisexual, or bisexual—curious. It is possible to choose nothing.; **Sexual polarity**: Options include dominant, submissive, and switch.

- **Looking for**: Options include long-term relationship, flirt, friendship, e-mailing; **About me**: The part where user can share a personalized bio.

- **Friends**: A list of id number of users added as friends.

- **Ads**: User can post public personal ads.

- **Meetings**: Number of community meetings in which the user personally participated.

- **Photos**: Number of profile photos added by the user.

- **Discussion**: Number of contributions to discussions (new threads and comments).

- **Chat time**: Shows how many hours the user has spent on the site since first login.

- **Last login**: The date when the user was last logged in. It is also a mark of activity of the user online.

- **Member since**: When the profile was created.

- **Points**: Points from administrators for activities such as new pictures, participation in discussion. Users receive negative points for breaking rules of the forum.

- **Age**: Age of user.

C. Human annotated data.

Over the course of one year, a member of our research team conducted interviews with 78 users in order to extract information about their tendencies to buy and/or sell sexual services in the forum. We had to overcome the challenge of correctly classifying a number of unlabeled profiles based on a small sample of labeled profiles.

D. Semi- supervised Classification.

Having obtained risk assessment results of 78 users using covert online ethnography, our objective is to build a machine learning model that can predict the risk factor in other 28,832 users. This is an ill-posed machine learning problem and can yield severely deteriorated performance results. However, under certain structural assumptions on the distribution of the underlying data, we can make use of semi-supervised learning setup for predicting the risk labels of unlabeled data.

Semi-supervised learning methods make the following assumptions on the underlying distribution of the data. (1) Continuity: this is the semi-supervised smoothness assumption which requires that the data distribution is such that nearby points are labeled identically. (2) Cluster property: data is distributed in discrete clusters with low intra-cluster distances and high inter-cluster distances, (3) Manifold property: this assume that the data lies on a lower dimensional

manifold and thereby the curse of dimensionality can be avoided. In this case we can attempt to learn the manifold using both the labeled and unlabeled data to avoid the curse of dimensionality. Then learning can proceed using distances and densities defined on the manifold.

Labeled data in our case refers to the 78 samples with risk labels "high risk" and "low risk" denoting high and low risk of being drawn into online sex trade, respectively. We utilize the following semi-supervised learning algorithms: Self-training [24], Semi-supervised support vector machines (S3VM) [25], and Contrastive Pessimistic Likelihood Estimation (CPLE) [23].

Self-training works by iteratively labelling the unlabeled instances using a trained classifier, and then re-training it on the resulting dataset. Semi-supervised support vector machines (S3VM) uses transduction learning to solve Overall Risk Minimization (ORM) problem rather than inductive learning of traditional support vector machines solving Statistical Risk Minimization (SRM). Contrastive Pessimistic Likelihood Estimation (CPLE) is a framework applicable for all classifiers which can yield prediction probabilities such that the model trained on both labelled and unlabeled data should not be worse than models trained only on the labelled data. We use CPLE with Logistic Regression algorithm..

E. Accuracy Evaluation.

We used various learning methods in order to automatically identify the people who were potentially at risk, using their profile information. We checked the accuracy of our methods against the validation set as prepared manually via covert ethnography. We measure how well our algorithm predicts the positive class, the category of people who are susceptible to buying/selling sex online using the semisup-learn toolkit [26].

TABLE I.   RESULTS USING DIFFERENT FEATURES

Features	PF	PF+TF	PF+TF+NF	PF+NF	TF
CPLE	61.91	62.55	64.61	60.81	58.34
Self training	78.45	76.79	77.16	75.18	56.89
S3vm Linear kernel	71.64	72.89	73.11	73.01	60.12
S3vm Polynomail kernel	73.56	72.15	74.13	74.16	66.21
S3vm Exponential kernel	77.87	76.59	78.75	77.51	69.10

PF: Profile features; **TF**: textual features; **NF**: SNA features

IV. RESULTS

We used three frameworks for semi-supervised learning and performed comparative experiments. These are (a) Contrastive Pessimistic Likelihood Estimation (CPLE) [23], (b) self-training [24], and (c) semi-supervised support vector machines [25].

Three distinct sets of features were used in our experiments:

1. Profile features: gender, age, sexual orientation, polarity, points, last_login, member_since, chat time, no. of discussions the user has taken part in, no. of ads posted by the user.

2. Profile features plus textual features extracted from "about me": unigram and bigrams extracted from textual description of about self and the user's description of "looking for".

3. Above features plus "social network" features: Based on the data we have scraped from the forum, we could build a "social network" of the users on the forum. Then for each user, we add the following indices to reflect their social position in the network: degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, Katz centrality, and PageRank centrality.

Results of our experiments with semi supervised approaches are shown the Table 1. We conduct experiments with various feature set combinations (of the three feature sets mentioned above). Table 1 shows Accuracy scores obtained from various classifiers with different feature combinations. As can be seen from Table 1, utilizing all features in S3VM with exponential kernel obtains the best results which is not surprising since it is consistent with the continuity and cluster assumptions of the data distribution. Regularization in all models comes from the unlabeled data naturally in the semi-supervised learning setup.

We also note that while deep learning models are well suitable to deal with problem of classification, they are unsuitable for semi-supervised learning.

V. Conclusion

Training a learner (human or machine) to detect people at risk of being involved into illicit activities is difficult because of the low availability of labeled data. We observe that identifying users prone to participate in online sex trade can be achieved with reasonable success when we use all available data: profile information, self-written description of biographical details, text written by the user to describe what he/she is looking for, and social network features.

Among these features, we find that a combination-based approach in which we used all features yielded the most accurate results. Using textual features alone did not yield good results, as very few users among the 28,833 users had written textual description about themselves. Similarly, using social network features also did not improve the results much, since only about 3,000 users had at least one friend in the forum. There was a large number of users with no friends. Using different kernels did not yield much improvement in accuracy, a result which we attribute to a very low-complexity decision boundary. We believe that if we have additional data on users' behavior (reflected in chatting, discussion, ads, etc.), accuracy can be further improved.

References

[1] Holme, Petter, 2012. Social, Sexual and Economic Networks of Prostitution. Leonardo, 45(1):80–8.

[2] Jonsson, L. S. C. G. and Svedin, M. H., 2014. Without the Internet, I Never Would Have Sold Sex: Young Women Selling Sex Online. Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 8(1): 1–14.

[3] Database: Risk of being drawn into online sex work, https://www.kaggle.com/panoskostakos/online-sex-work.

[4] Farley, Melissa et al., 2011. Comparing Sex Buyers with Men Who Don´t Buy Sex. Psychologists for Social Responsibility Annual Conference, Boston, Massachusetts July 15. Available from: http://www.prostitutionresearch.com/pdfs/Farleyetal2011ComparingSexBuyers.pdf.

[5] Hughes, Donna M., 2003. Prostitution Online. Journal of Trauma Practice, 2(3-4):115-131.

[6] Vanwesenbeeck, Ine, 2013. Prostitution Push and Pull: Male and Female Perspectives. Journal of Sex Research, 50(1):11-16.

[7] Farley, Melissa, Franzblau, Kenneth and Kennedy, M. Alexis, 2014. Online Prostitution and Trafficking. Albany Law Review, 77(3):1039-1094.

[8] Jones, Angela, 2015. Sex Work in Digital Era. Sociology Compass, 9(7):558-570.

[9] Levey, Tania G. and Pinsky, Dina, 2015. A Constellation of Stigmas: Intersectional Stigma Management and the Professional Dominatrix. Deviant Behavior, 36:347–367.

[10] Kaye, Anders, 2017. Why Pornography is not Prostitution: Folk Theories of Sexuality in the Law of Vice. Saint Louis University School of Law Journal, 60(243):243-292.

[11] Shapiro, Melanie and Hughes, Donna M., 2017. Decriminalized Prostitution: Impunity for Violence and Exploitation. Wake Forest Law Review, 52(2):359-377.

[12] Chon, Don Soo, 2015. Gender Equality, Liberalism and Attitude Toward Prostitution: Variation in Cross-National Study. Journal of Family Violence, 30:827–838.

[13] Boels, Dominique, 2015. The Challenges of Belgian Prostitution Markets as Le-gal Informal Economies: An Empirical Look Behind the Scenes at the Oldest Profession in the World. European Journal of Criminal Policy and Research, 21:485–507.

[14] Van Hout, Marie Claire and Hearne, Evelyn, 2016. Netnography of Female Use of the Synthetic Growth Hormone CJC-1295: Pulses and Potions. Substance Use & Misuse, 51(1).

[15] Belz, Frank-Martin and Baumbach, Wenke, 2010. Netnography as a Method of Lead User Identification. Creativity and Innovation Management, 19(3):304-313.

[16] Langer, Roy and Beckman, Suzanne C., 2005. Sensitive research topics: Netnography revisited. Qualitative Market Research: An International Journal, 8(2).

[17] Calvey, David, 2013. Covert Ethnography in Criminology: A Submerged yet Creative Tradition. Current Issues in Criminal Justice, 25(1):541-550.

[18] Edlund, Lena, Engelberg, Joseph and Parsons, Christopher A., 2009. The Wages of Sin. Columbia University Economics Discussion Paper, 0809–16.

[19] Cunningham, Scott and Kendall, Todd D., 2010. Risk Behaviours among Internet-Facilitated Sexworkers: Evidence from Two New Datasets. Sexually Transmitted Infections, 86(3):100-110.

[20] Chao, Biming, 2016. Relevance and adoption of netnography in determining consumer behavior patterns on the web. Scholedge International Journal of Business Policy & Governance. 2(6):12-17.

[21] Crichton, Susan and Kinash, Shelley, 2003. Virtual Ethnography: Interactive Interviewing Online as Method. Canadian Journal of Learning and Technology, 29(2). Available online: https://www.cjlt.ca/index.php/cjlt/article/view/26548/19730

[22] Ebo, Bosah, 1998. Internet or Outernet? In B. Ebo (ed.) Cyberghetto or Cybertopia? Race, Class, and Gender on the Internet, 1–12. Westport, CT: Praeger.

[23] Loog, M. (2016). Contrastive pessimistic likelihood estimation for semi-supervised classification. IEEE transactions on pattern analysis and machine intelligence, 38(3), 462-475.

[24] Chapelle, O., Scholkopf, B., & Zien, A. (2009). Semi-supervised learning (Chapelle, o. et al., eds.; 2006) [book reviews]. IEEE Transactions on Neural Networks, 20(3), 542-542.

[25] Bennett, K. P., & Demiriz, A. (1999). Semi-supervised support vector ma-chines. In Advances in Neural Information processing systems (pp. 368-374).

[26] https://github.com/tmadl/semisup-learn