Towards Exposing Cyberstalkers in Online Social Networks

Jiamou Liu[1], Yingying Tao[2], and Quan Bai[2]

[1] The University of Auckland, New Zealand
jiamou.liu@auckland.ac.nz
[2] Auckland University of Technology
{kdx8538,quan.bai}@aut.ac.nz

Abstract. This paper presents work-in-progress towards a computational approach for capturing a type of deviant behaviors, that are characterised by persistent monitoring and information gathering through online social medias. Such behaviors are commonly associated with stalking on the cyberspace. We present a network-based framework for describing online user interactions. Based on this framework we provide a description of excessive and unreciprocated attention an agent pays to another agent. We conclude with a discussion on limitation of the current work and a guideline for future extensions.

Keywords: Cyberstalking · online social network · agent behaviors

1 Introduction

This paper aims to provide a formal behavioral-oriented study on certain situations, where an online user persistently monitors and gathers information from another user through online social medias. Such deviant behaviors embody, to a large extent, the main characteristics of the so-called behaviors of *cyberstalking* and *cyberbullying*. Both types of deviant behaviors have become prevalent problems that are associated with the online social medias, and have serious social and psychological implications, which hamper the safe usage of the Internet [16,1,15,14].

This research falls into the broader area of cyber-safety and privacy in computer science: In the last decade, there are major efforts towards the detection of phising, spamming, cloning, and bots on online social medias [8]. However we have found no technical breakthrough specifically focusing on cyberstalking. This may be due to the complex nature of the type of online user interactions and the actions generally associated with cyberstalking, such as online harassment and identity theft. As pointed out in [15], despite decades of criminological research, there has not been a generally agreeable definition of cyberstalking. The goal of our result is to bridge this gap by investigating cyberstalking in a formal, computational perspective and hopefully develop useful technologies

that help with the prediction, detection, and forensics of cyberstalking. Indeed, this pilot work presents work-in-progress towards this goal.

The basic premises of this work relies on tools and techniques provided by network science in the modeling, simulation, detection and prevention of cyberstalking behaviors. Network science is an interdisciplinary area that aims to explore properties of complex relational structures such as information networks, social networks and biological networks. The main object of investigation consists of a crowd of autonomous agents that are interconnected through ties. The actions of individual agents are affected by actions of other agents as well as the interactions between them. The field focuses on developing models of such networks and explores structural properties of complex networks through algorithmic analysis [3,10,13]. When viewing the Internet as a large information network, one can extract and analyze traits and trends of Internet users using various statistical and graph-theoretical tools.

The main goal of the work is to introduce a formal framework for agents in a social network, that could be used to describe a range of user behaviors. We then would like to develop tools that are able to detect abnormalities in such simulated environment, that is, the tool should help to distinguish benevolent behaviors and abnormal behaviors. We provide a mathematical model, which has the potential to be developed into an automatic mechanism for analyzing and detecting cyberstalkers through analyzing network logs.

Organization. The rest of the paper is organized as follows: Section 2 discusses existing literatures of cyberstalking in the social sciences. In particular, we revisit existing descriptions of cyberstalking behaviors and key challenges. Section 3 presents a general framework for describing user behaviors in an online social network. A crucial component of our model aims to capture the amount of attention a user pays to another. Based on this, Section 4 provides a brief discussion on our behavioral model of cyberstalking. Finally, Section 5 concludes the paper with discussion of limitation of the model and future directions.

2 Understanding Cyberstalking

The conventional, physical act of *stalking* refers to the stealthy and persistent pursuit of a specific person with unwanted and obsessive attention which results in harassment. The person who carries out stalking is referred to as a *stalker* or the *perpetrator*, and the person being stalked is the *victim*. *Cyberstalking* is the form of stalking that occurs in the cyberspace. It is believed that cyberstalking may result in a similar, if not higher, level of threatening as conventional stalking; indeed, these two types of behaviors share a number of common traits [15]:

(a) Both behaviors are characterized by obsessive attention from the perpetrator to the victim, which include monitoring and information gathering.
(b) Both behaviors are mainly driven by the perpetrator's desire and need to gain power, control and influence over the victim.

(c) Both behaviors are *victim-defined*, that is, the level of seriousness of a particular stalking incident is determined by how much intimidation the victim perceives [16].

However, cyberstalking should not be regarded simply as an extension of conventional stalking, but rather a different deviant behaviors in its own form [6]:

1 Firstly, conventional stalking includes some clearly defined, detectable actions such as physically following the victim home, vandalizing victims properties and sending gifts, leaving a physical trail of evidences. Cyberstalking, on the other hand, are much harder to define: Here, a victim may purposely expose private information on online social networks, making monitoring extremely easy; any online user may derive information regarding another's occupation, age, and address with little effort. Furthermore a stalker may also use identity masks, making information gathering unnoticeable.
2 Secondly, in most conventional stalking incidents, the stalker and the victim are within each other's social or physical periphery: Either they have prior relationship (whether real or perceived), or they live or work within relatively close proximity of each other. For cyberstalking, the victim is more often chosen at random and may occur between two people with arbitrary physical distance.
3 Thirdly, cyberstalkers often employ tools and techniques that are unique to the use of the Internet. For example, a number of incidents involve stalkers carrying their deeds using Trojan softwares, email spamming or phishing techniques [8], all of which are challenging technical online threats themselves.

Due to the above reasons, the detection, prevention, and forensics of cyberstalking become considerably more challenging than for conventional stalking. Despite serious research efforts, there has not been major technical advancements on the detection and prevention of cyberstalking. The most widely used methodology is *profiling*: by gathering common characteristics of cyberstalkers and victims (such as age, gender, occupation, etc.), the method aims to identify the likelihood of an individual to be a stalker or be stalked [1]. While profiling provides certain statistical information, it is far from an effective method for prediction and detection. In view that anyone has the potential to cyberstalk, one needs a *behavioral* approach, which focuses on describing behaviors, rather than summarising personal attributes of cyberstalkers.

One of the most comprehensive definitions of cyberstalking was offered by Bocij and McFarlane in [5]:

> *"A group of behaviours in which an individual, group of individuals or organisation uses information technology to harass one or more individuals. Such behaviour may include, but are not limited to, the transmission of threats and false accusations, identity theft, data theft, damage to data or equipment, computer monitoring and the solicitation of minors for sexual purposes. "*

The definition entails a large sum of deviant behaviors, which vary by nature and require different countermeasures. As a pilot study, we can only focus on a particular aspect of cyberstalking. In an earlier study [11], Meloy provides a more "strip-down" definition, which asserts cyberstalking as consisting of two major functions

1. the stalker gathers private information of the target to further a pursuit; and
2. the stalker communicate (in real time or not) with the target to implicitly or explicitly threaten or to induce fear.

In this paper we focus on the first function above. We argue that stalkers' gathering of information is the prerequisite and precursor to all cyberstalking behaviors as laid out by Bocij and McFarlane. It is the stalker's ability to collect personal information of the victim that places the victim in the danger of threats and harassment. By appropriate monitoring and control of the gathering of information, one may develop a feasible countermeasure for cyberstalking.

3 A Model for Online Behaviorals

3.1 A Model of User Interactions

People interact in an online social network in a variety of ways, which range from posting profiles, publishing blogs and photos, forming groups, leaving comments, messaging, sending emails or simply clicking `like` button on others posts. To capture this diverse range of interactions, we need an abstraction of cyberspace that serves as a general platform for our investigation. Indeed, we provide a high level description which implicitly embodies all the mentioned activities.

Naturally, an online social network consists of ties between its members; such ties could mean mutual friendship, or the relationship when a person "follows" another which indicates a directed interest from one person to another. Such ties consist of visible links among users of the network and form the basis of the dissipation of information.

However, establishing interpersonal ties are not the only form of interactions. A published message can be viewed by anyone on the web, regardless of whether the viewer has an established tie with the publisher. Similarly, the cyberspace enables communication between two users, who may or may not have a friendship on the online social network. It is the mechanism that allow information sharing through distant parties fundamentally distinguishes physical interaction from online interaction, and hence any model of the cyberspace should take into account not only the visible connections between users, but also interactions among arbitrary users on the network.

With this view in mind, we introduce the following model of an online social network: The carcass of the model consists of a directed network where nodes represent agents (i.e., users). Directed links allows modeling of the relationship when one agent "follows" another. Note that by restricting to directed links, we also do not rule out mutual relations as they can be represented by a pair

of directed links. Each agent in this network has the ability to carry out two actions: *posting* messages and *retrieving* messages. Once a message is posted, it can be retrieved by others in the cyberspace. For simplicitly, we assume that there is a universal set of messages that could be posted and hence retrieved over the network and all posted and retrieved messages are taken from this set. Furthermore, we require that an agent should not post and retrieve the same message at the same time.

Definition 1 (Network and agents). *A* network *is a directed graph* $G = (V, E)$ *where* V *is a set of* nodes *(or* agents*) and* $E \subseteq V^2$ *is a set of* directed edges *denoting interpersonal ties. Let* M *be a set of* messages*. An* instance *of* G *is a pair of functions* $s = (\mathsf{post}, \mathsf{retrieve})$*, where* $\mathsf{post} : V \to 2^M$ *is called the* post function *and* $\mathsf{retrieve} : V \to 2^M$ *is called the* retrieve function*. We further requires that for all* $v \in V$*,* $\mathsf{post}(v) \cap \mathsf{retrieve}(v) = \varnothing$*.*

An instance $s = (\mathsf{post}, \mathsf{retrieve})$ captures the posting and retrieving behavior of agents in a single time instance: We say that a message $m \in M$ is *posted* by an agent v if $m \in \mathsf{post}(v)$; a message $m \in M$ is *retrieved* by v if $m \in \mathsf{retrieve}(v)$. Hence $\mathsf{post}(v)$ contains all messages in M that are posted by agent v and $\mathsf{retrieve}(v)$ contains all messages that are retrieved by v.

Remark. It may seem that the actions of "posting" and "retrieving" are too restrictive as they only give users limited capabilities in an social network. Nevertheless, they can be used as a high-level abstractions of a wide range of activities. For example, when one person sends an email to another person, this can be seen as a process where the sender posts a message, while the receiver – and only the retriever – retrieves the message. It is therefore not hard to see that this general set up can be used as a simple abstraction of online social networks.

3.2 A Model of Attention

Attention refers to an invested interest from one person to another. It comes often as the result of a desire to know about and establish connection with the target person. The amount of attention an agent a pays to another agent b can be viewed from two perspectives. The first is a micro perspective: the attention a pays to b depends on the posted and retrieved messages of a and b. The second is a macro perspective: the link structure of the network affects attention from a to b. In our model we need to incorporate both types of attentions to truthfully capture the attention one pays to another in the cyberspace.

Firstly, in a micro-level, we measure attention by comparing the set $\mathsf{post}(b)$ against the set $\mathsf{retrieve}(a)$. If $\mathsf{post}(b)$ and $\mathsf{retrieve}(a)$ are similar, then the agent a retrieves mostly the posted messages of b, which naturally implies that a pays attention to the messages posted by b. To measure similarity between two sets, we adopt the well-known *Jaccard distance*, which is the ratio of the size of the intersection of the two sets over the size of their union. Thus the Jaccard distance has a range between 0 and 1 where a higher value implies a higher

level of similarity. We define the *message-based attention index* (M-index) from a to b as the Jaccard distance between $\mathsf{retrieve}(a)$ and $\mathsf{post}(b)$. A special case is when $a = b$, that is, we need to define the amount of attention a person gives to her own posted messages. We set the M-index from a to a to 1, i.e., a pays full attention to her own posted messages, which is clearly reasonable.

Then, in a macro-level, the second factor for measuring attention is the link structure of the network $G = (V, E)$. Indeed, the M-index only measures an agent's interest in the *information* sent by another agent, which is not necessarily the same as the interest on the target agent herself. For example, an agent a may retrieve blogs posted by another agent b in the hope to monitor a close friend c of b; while no message posted by c is retrieved by a, a's primary intention is to get to know about c. This method of "indirect" information gathering has been shown to pose a great privacy threat on the Internet [12]. We measure this *linked-based attention* using a Markov chain model.

We briefly outline the necessary steps in defining linked-based attention as follows: Given an instance s of the network G, we define an *attention transfer Markov chain* $T_s(a)$ for every node a by taking into account the local network structure of a and neighbors of a. The linked-based attention of a is then given by the stationary distribution vector of this Markov chain.

4 Capturing Stalking-like Behaviors

As argued in [15], cyberstalkers tend to be emotionally distant loners who is eager to seek the attention and companionship of another. The problem lies in the fact that these individuals often become obsessed or infatuated with the victim, but such a feeling is not reciprocated. Hence it is reasonable to view stalking as *excessive, unreciprocated attention over a lengthy period of time*. We need to make clear of the following notions.

To capture the above notions mathematically, we use a measure for the extend of attention an agent pays to another in the network. The judgement whether an agent i pays obsessive attention to another agent j depends not only on the attention from i to j alone, but also how much attention j generally receives from all other users in the network. At the same time it should also depend on how much attention j pays to i. Furthermore, the type of watchful behavior that we want to measure here happens over a lengthy period of time. Thus, assuming that the network G does not change with time. Given a sequence of instances $s_0, s_1, \ldots$ of G, we accumulate the attention between i and j over all the instances in this sequence. The result is a numerical indication of how likely the agent i pays to the agent j in the network.

5 Discussions on Limitations

This work-in-progress amounts to our first step towards building a mathematical framework of cyberstalking. It is still a speculation that our definition leads to automated technologies for the identification, detection and prevention of

cyberstalking. Nevertheless, we believe that by proposing a precise definition, it becomes possible to tackle these problems from a computational perspective.

Given the simplicity of the current model, some natural limitations exist, which we will try to address in our future works.

Justification of the attention model The model of link-based attention using Markov chains provide a powerful measure on a latent intention from one agent to another in the network. Through our initial experiments, we identify a clear indication that cyberstalking behaviors result in a very high level of attention. It is therefore imperative to carry out more theoretical and empirical studies regarding this model to further verify its validity. To accomplish this task, one would need to design appropriate simulation models and scenarios and identify how much different factors (such as network density, level of interactions, network structural properties, etc.) affects the level of attentions among agents. One should also consider verifying this model using real-world data collected from online networks.

Message types and meanings In real life networks, any message carries a meaning, i.e., a piece of information that relates to certain topics or people. For example, an online user normally maintains a profile page which contains important personal information. A stalker in collecting information about the user, typically first attempts to access the user profiles of the victim and friends of the victim. Another possible scenario is when the stalker pays a high level of attention on the topic of interest to the victim, rather than the specific messages. Such intention cannot be precisely described by our model. To deal with this issue, one would consider not only the interpersonal relations among agents, but also semantic relations among messages and topics, as well as attributes of messages.

Harassments and threats Cyberstalking in most legal contexts entails a level of intimidation or harassment to the victim. Indeed, such harassment must come from communication between the stalker and the victim. Numerous real-world cases of cyberstalking involve the stalker making certain early contacts (such as sending friend request) to the victim. The definition in this paper, however, merely capture the process of information collection, but not communication between agents. Hence more novel approaches based on cognitive or behavioral science may be employed to provide a useful solution.

Multiple online identities One important property of real life social networks is that a user can hide behind multiple anonymous identities. When stalking a victim, each online identify of the stalker may engage in a collective effort in collecting information, hence making the problem much harder. This also requires enriching the current model with mechanisms that identify user accounts that are associated with the same identity.

Incorporating historical records When stalking a victim, a perpetrator usually not only collect current post of the victim but also retrieve historical messages (old blogs, diaries, photos, etc.). However, the current definition of message-based attention only take into account the current retrieval and posts. This requires enriching the current model with a knowledge base which

stores logs to all historical messages posted by users, and updating the attention model by taking into account all historical data.

As discussed above, an automated mechanism for the prediction, detections and forensics of cyberstalking behaviors relies on resolving all the issues above by enriching and validating our model. Nevertheless, we believe that the final outcome will be an important technical breakthrough towards solving the increasingly-significant problem of online stalking and bullying. The current work-in-progress provides us with a clear guideline of how this could be achieved.

References

1. Al-Khateeb, H., Alhaboby, Z., Barnes, J., Brown, A., Brown, R., Cobley, P., Gilbert, J., McNamara, N., Short, E., Shukla, M., (2015). A Practical Guide To Coping With Cyberstalking, 1st ed.. National Centre for Cyberstalking Research (NCCR), University of Bedfordshire. Luton: Andrews UK Limited, 2015.
2. Barabási, A; Albert, R., (1999). Emergence of scaling in random networks. Science 286 (5439): 509–512.
3. Barrat, A., Barthelemy, M., Vespignani, A., (2008). Dynamical processes on complex networks, Cambridge University Press.
4. Barabási, A., Bonabeau, E., (2003). Scale-Free Networks. Scientific American 288 (5): 50-–9.
5. Bocij, P., McFarlane, L., (2002). Online harassment:towards a definition of cyberstalking. Prison service journal, 139, pp. 31–38.
6. Bocij, P., McFarlane, L., (2002). Cyber stalking: Genuine problem or public hysteria? Prison Services Journal, 140(1), 32–35.
7. Erdös, P., Rényi, A. (1959). On Random Graphs. I. Publicationes Mathematicae 6: 290—297.
8. Fire, M., Goldschmidt, R., Elovici, Y., (2014). Online social networks: threats and solutions, Communications Surveys & Tutorials, IEEE 16 (4), 2019–2036.
9. Johnson, N., Kotz, S., (1977). Urn Models and Their Application: An Approach to Modern Discrete Probability Theory, Wiley.
10. Liu, J., Moskvina, A., *Hierarchies, Ties and Power in Organisational Networks: Model and Analysis*, in Proc. of ASONAM 2015, 202–209, 2015.
11. Meloy, J, (1999). Stalking. An old behavior, a new crime. Psychiatric Clinics of North America, 199, Mar;22(1):85–99.
12. Mislove, A., Viswanath, B., Gummadi, K., Druschel, P., (2010). You are who you know: inferring user profiles in online social networks. In: Proceedings of the third ACM international conference on Web search and data mining (WSDM 2010) 251–260.
13. Moskvina, A., Liu, J., *How to build your network? A structural analysis*, to appear in IJCAI 2016.
14. Parsons-Pollard, N., Moriarty, L., (2009). Cyberstalking: Utilizing What We Do Know. Victims & Offenders: An International Journal of Evidence-based Research, Policy, and Practice 4.4 (2009): 435–441.
15. Pittaro, M. (2007). Cyber stalking: An Analysis of Online Harassment and Intimidation, International Journal of Cyber Criminology (IJCC) Vol 1 (2): 180—197.
16. Reno, J. (1999). Cyber stalking: A new challenge for law enforcement and industry. United States Department of Justice.