Dynamic Security Risk Evaluation via Hybrid Bayesian Risk Graph in Cyber-Physical Social Systems

Shancang Li, Shanshan Zhao, Yong Yuan, Qindong Sun, Kewang Zhang

***Abstract*—Cyber-physical social system (CPSS) plays an important role in both the modern lifestyle and business areas, which significantly changes the way we interact with the world. The increasing influence of cyber systems and social networks is also a high risk for security threats. In this paper, we investigate the potential risks in social networks using a hybrid Bayesian risk graph (HBRG) model to analyse the temporal attack activity patterns in dynamic cyber-physical social networks. In this model, a hidden Markov Model (HMM) is proposed to model the dynamic influence of activities, which then be mapped into a Bayesian risk graphical (BRG) model that can evaluate the risk propagation in a layered risk architecture. Our numerical studies demonstrate that the framework can model and evaluate risks of user activity patterns that expose to cyber-physical social systems.**

.

***Index Terms*—Activity Profile Modelling, Risk Analysis, Hidden Markov Model, Bayesian Risk Graph, Cyber-physical social system.**

I. Introduction

The cyber-physical social networks, such as `Facebook, Twitter, Youtube, Google+, etc.`, are playing an important role in our modern lifestyle and business model, which are significantly changing the way we access to information, interact to others, and even change the business model across the world [1], [2], [3], [4]. In recent, it is reported that more than 69% of the public uses some type of social media/networks in American and nearly 80% of businesses are getting invovled in developing social network resources, such as social videos, social media advertising, and social messages [1], [5]. These social networks provide incredible opportunities and resources for online users, however, there are also a high risk for the online security threats or attacks. The social network also makes cyber attackers easier to exploit vulnerabilities and it is being weaponized by the attackers [6].

With the increasing usage of social networks and the emergence of new technologies, such as the Internet of Things (IoT), Big Data, cloud computing, the security issues in social networks face novel research problems and challenges [2], [5], [6]. The hundreds of millions of users are facing security threats, such as, cyber crimes, identity stolen, device/social profile hacked, overconfidence, etc. Effective risk evaluation should be provided to help the social network platforms and users well understand the security situation they are facing, and accordingly security and privacy protection solutions should be developed to help the users stay safe online [7], [8]. In the past decade, lots of research efforts have been done on security of social networks, including security risk analysis, abnormal activities detection, cyber crimes, terrorist attacks, malicious users or device detections, shortened or hidden URL, etc., to reduce the potential threats that the users facing. Actually, most users often woefully unaware of most security threats or attacks they are facing [8]. It is reported in [9] that there is 70% of increase in scams (such as hidden URLs, phishing requests, etc.) in social networks, which spreads rapidly since most users often more like or re-share links or information posted by their friends. However, when a user profile is compromised, the attacker can also spread the scams rapidly. In some cases, the scammers embraced some popular dating applications (`app`) or some adult-themed contents or links to attract more clicks from users [10].

Many social system security vulnerabilities are exposed to attackers without knowing by the users. The most common vulnerabilities can be summarized into following five categories: (1) careless profile leakage; (2) dumpster diving, the attacker can compromise user identity with information provided by the user self; (3) information that can be used to attack your profile, such as hints to help guess your password; (4) links to malware, (5) corporate spies and activist stalkers. Meanwhile, in commercial areas, many methods have been developed to collect user messages that can break the defenses, the emerging technologies, such as deep learning, cloud computing, big data, etc., can be effective to do this. The social networks have become new source of risks and poor security protection can put the users at serious risks. In [11], the FBI highlighted that the social network `Facebook` scamming has become the most common form of malware distributed in 2016. The social systems, such as `Facebook, Twitter, etc.`, are increasingly an effective tool for cyber criminals, terrorist groups like ISIS. The social systems security are facing severe' security challenges:

1) Most social systems are unable to secure the environment for users. It is reported that 2% of the `Facebook`

Dr. Shancang Li is with Department of Computer Science and Creative Technology, University of the West of England, Bristol, UK. Email: `shancang.li@uwe.ac.uk`, Tel: +44 (0)117 32 86693.

Dr. Shanshan Zhao is with Department of Engineering, Design and mathematics, University of the West of England, Bristol, UK. Email: `shanshan.zhao@uwe.ac.uk`.

Prof. Qindong Sun is with School of Computer Science, Xi'an University of Technology, Xi'an 710048, China.

Prof. Yong Yuan is with Institute of Automation, Chinese Academy of Science, Beijing, China. Email: `yong.yuan@ia.ac.cn`.

Dr. Kewang Zhang is with School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China.

and 5% of Twitter monthly average users are "false accounts" [11], [12], which means that the social network platforms are unable to provide reliable system to identify duplicated or fraudulent accounts.

2) Social network scamming are highly-effective and lucrative. Only less than 20% of scams can be spotted by the professional users [11], [13], [14]. The social networks are becoming the new cyber-weapons of choice for cyber criminals. The social networks provide unreliable "trustiness" for users which may result in rampant scams spread rapidly. The hight volume of visible social network scams makes it difficult for security expert to deal with.
3) Social networks and its data are weaponizing by the cyber criminals. Cyber attackers/criminals can easily target specific victims through social networks. For example, the LinkedIn was a key recon tool for the cyber criminals who executed the Anthem data breach and its 80 million stolen records. The Twitter was the target of an innovative malware exploit dubbed "hammertoss", which is rumored to be connected to Pentagon's data breach last summer that took down the security agencys' 4,200 employee email server for two weeks [15], [16].

The main contributions of this work are: (1) to develop a hybrid risk analysis model for social network, in which a Hidden Markov Model (HMM) is introduced to model the dynamic user activities that can cause potential risks. The HMM model takes account of the aggregated influences of activities of neighbours (such as followers, friends, etc.) that can affect the activities of a user. (2) a Bayesian risk graph (BRG) model is introduced in the top layer to analyse the potential risks that the activities can cause. The BRG model is able to classify the user activities into three level (static, dynamic, behaviour) and can dynamically evaluate the potential risks that caused by user activities. (3) A node mapping scheme is proposed that can map the HMMs in the bottom layer to the Bayesian nodes in the BRG model.

The remainder of this paper is organized as follows. Section II reviews the related works and Section III details the hybrid Bayesian risk model. In Section IV, the risk evaluation framework is proposed based on the HBRG model for social networks. Section V concludes the paper with a summary and discussion.

II. Related Works

The social network platforms afford users both opportunities and risks. In 2016, social network platforms, such as Twitter, Facebook, Youtube made strenuous efforts to purge risky social networking practices to ensure users safety on social networks. More than $360,000$ malicious or inactive accounts have been suspended. In recent, a lot of research efforts have been made on investigating novel research problems and challenges in security risks that the social networks are facing. In [16], a framework is proposed that can separate the spammers and unsolicited bloggers from the genuine experts of a specific domain by using the hyperlink induced topic search (HITS). The proposed framework is able to help recommendation system and alike services to identify bloggers, however it is unable to do more deep investigation, such as *event tracking, social network forensics, and timeline matching*, which are important in forensics analysis and risk evaluations. Li et al. presented a profile matching schemes for social networks named Scalable and Privacy-preserving Friend Matching protocol (SPFM) [17]. This scheme can provide a scalable friend matching and recommendation solutions without revealing the users personal data to the cloud. In [18], a social privacy protector is proposed for preserving privacy information in social networks. It is able to identify the fake profile by analysing the user's friend list. The developed classifier utilizes a machine learning algorithm to identify the fake account by measuring connection strengths of the user with the people in friend list based on a heuristic that considers several features, such as common friends, common groups, common posts, etc. However, this method is unable to mark the users and it does not take into account the microbloggers that may affect the results.

In recent, with the research progress made in big data analysis, a number of research works have been done on analysis of forensics in social networks. In [19], a forensics analysis framework is proposed that is able to identify important data sources for automated forensic analysis on social network user data. The proposed identification-graph can visualize the identify graph based on the social networks data without the collaboration of the social network operators. The proposed framework does not take the affect of neighbors influences in their event tracking. In [20], coupled HMM model is proposed to describe the temporal activity patterns in social networks. This model is able to accurately learn models with sufficient observations. In [3], a HMM model is proposed to address the information integration problems.

In our previous works [2], we investigated the security risks in mobile systems and proposed a multi-layered hierarchical Bayesian network [2]. The system is able to dynamically analyse the potential risks in mobile systems by integrating static analysis, dynamic analysis, and behaviour analysis in a hierarchical framework. The risks and their propagation through each layer are well modeled by the Bayesian risk graph, which can quantitatively analyze risks faced to both apps and mobile systems. In [20], Vasanthan et al. investigated the user activity patterns by using a Markov model methods based on the observed data and a clustering algorithm is proposed that can group users according to the interaction behaviours. Tu *et al.* proposed a collaborative scheme based on hierarchical and hybrid Bayesian networks (HHBNs) to investigate the information integration [3], [21]. The HHBNs are used to analyse a terrorist attack scenario. However, the proposed methods are unable to dynamically analyse the activity pattern on social networks. Riek *et al.* investigated the cybercrimes in social networks based on a parsimonious model, which is able to identify risk causal factors that reduce users' intention to use online service. In [23], Ross *et al.* proposed a socio-physical approach by taking the joint interaction and integration of social and physical into a system to improve emergency response and preparedness. The proposed methods can evaluate and reduce risks by enabling

an informed-coordinated response strategy, which is effective for static social activities and physical activities, however it is unable to deal with the dynamic user on social networks.

It is clear that a good model of user activity can be very helpful for analysing the risks and security threats for the activity patterns in dynamic social networks. In the following sections, we will introduce a hybrid model by using the well-developed HMMs and the BRG we have developed previously. The HMMs are able to model complicated dynamic user activity patterns in social networks, meanwhile the BRG can model the causal factors of potential risks caused by these activity patterns.

III. Hybrid Bayesian Risk Model

A. Threats and Risks in CPSS

As mentioned above, the social network platforms are free to use for users, in which the risk management and evaluation are critical but inadequate for the incredible resources. Most social network platforms, such as `Twitter`, `Facebook`, `LinkedIn` *et al.*, are using proactive risk management and evaluation scheme to monitor users anomalous activity based on the behavioral analysis and clustering algorithms [4], [5], [6]. The attackers or potential criminals may also continuously explore the vulnerabilities by mixing various attack techniques, such as *profile attacks, scamming, fake apps, like-jacking, etc.* In many cases, the attackers can collect sensitive privacy information from users' profile by combining different pieces of information in many different ways. Generally, we can group the possible attacks into two categories: (1) Vertical attacks, which focus on a specific social network site or specific user, and (2) Horizontal attacks, which focus on the cross-correlation networks to mine the sensitive information that might be useful for committing attacks. The information might come from multiple different sources (such as social networks, emails, IoT, interested forums, etc.). The dumpster diving attack is a typical horizontal attack, in which an attacker can cross-correlated and complement the attributes of a user's profile by mining his different profiles in other social networks, posts, and replies for posts, etc.

Typical features of social networks include free webspace, building profiles, building conversations/content, messengers, creating pages, etc. These features introduce new threats/risks like social network worms (such as `Koobface`), botnet, hijack, phishing scams, trojans, shortened links, data links, and APTs, etc. Actually, in social networks, these threats/risks are no longer a single type of attacks. It becomes a security attack scheme by integrating data collection, processing, data mining from numerous internal and external sources in a real time way. The attacks might be systematic and occurs rapidly. Therefore, the security analysis and risk evaluation schemes for social networks should be able to provide a comprehensive and rapid response. It is necessary to develop a security risk/threat analysis model by incorporating the social network influence as perceived by the users, which should be designed to be able to real-timely identify the major factor leading the security risks.

B. Hybrid Bayesian Risk Model

The hybrid Bayesian risk model has a two-layer and interconnected architecture as shown in Figure 1. Basically, the Hybrid Bayesian Risk Graph (HBRG) model consists of a hidden markov model (HMM) layer and an interconnected Bayesian risk graph (BRG) network. In the bottom layer, the HMM is used to model the activities of user in dynamic social network, which describes the states, observations, and the aggregation of influences of neighouring nodes. In the example in Figure 1, we have multiple HMMs and each might represent different activities that correspod to a node in BRG layer. The HMM is powerful for modelling users' activity evolution in social networks according to a Markov chain with a hidden state that is influenced by the collective activity of the neighbouring of the user. Meanwhile, the BRG network (also known as risk causal network) is a probabilistic graphical model that represents a set of featured risks and their conditional dependencies via a directed acyclic graph (DAG) [2]. In the HBRG model, the nodes represent the risks and the links between nodes, and layers represent probabilistic causal dependencies.

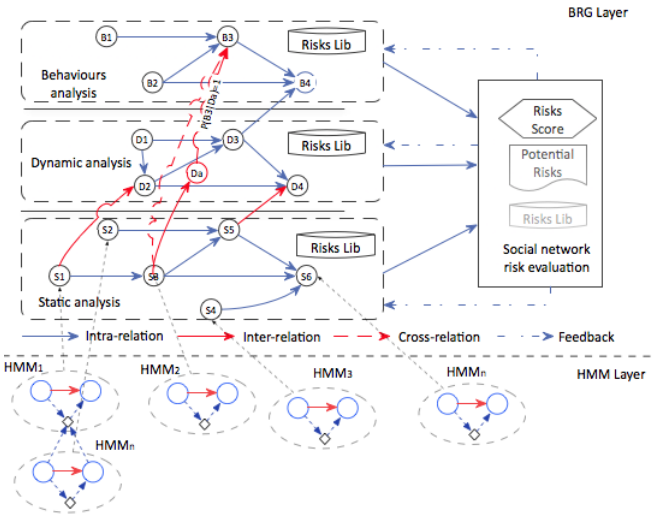

Fig. 1. Hierarchical Bayesian Risk Graph Model

1) HMMs: The HMM is powerful for modelling sequential states [3] that has been widely used in network activities modelling. Since the user activity in social networks has distinctly non-*Poissonian* characteristics [3], [20], furthermore, the activities of neighbouring users (followers, friends in social network) can significantly affect user activity, the HMM can well describe user activity and explicitly take into account the interaction between users by introducing a coupling between two stochastic processes. The works in [3], [23], [24], [25], [26] show that the coupling HMMs can well model a probabilistic process of status and the interactions between these activities. Actually, in social networks, the user activity has the following features: (1) the activity of users are dynamic; (2) the individual activity may be preferentially affected by other linked users (such as followers, friends, or even some unlinked users); (3) the states of users are unobservable. The HMM is

able to model the hidden states, which correspond to different patterns in user activity. In social neworks, the state transition can be influenced by its neighbors and it is possible to explain the observed data and predict the future activity of a user [20]. With a learning model, the HMM can cluster users and find the resulting cluster structure allowing intuitive characterization of the users in terms of the interaction dynamics between a user and his social network.

In this work, the user dynamic activity in social network can be modelled by the following three components: *User States; (2) Observation density; (3) Influence of neighbors*. We use a three - tuple $H_i = \{Q_i, \mathcal{T}_i, Z_i\}$ to represent an HMM H_i.

Let $\mathbf{t} = \{t_1, t_2, \ldots, t_N\}$ denote the time-stamps of a specific user's activities over a given period-of-interest. The time duration $\mathcal{T} = \{\tau_i = t_i - t_{i-1} | i = 1, 2, \ldots, N\}$.

Definition 1. *User State. In a HBRG, we use Q_i to denote the states of a user i, which could be active ($Q_i = 1$) or inactive ($Q_i = 0$).*

Let $Q_i \in \{0, 1\} (i = 1, \ldots, N)$ denote the states of a user, $Q_i = 0$ means it is in *inactive* state at period τ_i, and $Q_i = 1$ means it is in an *active* state. It is clear that the state of Q_i is dependent only on Q_{i-1}, and we have

$$P_0 = \begin{bmatrix} 1 - p_0 & p_0 \\ q_0 & 1 - q_0 \end{bmatrix} \tag{1}$$

and

$$P_1 = \begin{bmatrix} 1 - p_1 & P_1 \\ q_1 & 1 - q_1 \end{bmatrix} \tag{2}$$

in which p_0, q_0 denote probabilities from the inactive state to the activate state.

Definition 2. *Observation. In social networks, the user states are unobservable, the observations $\mathcal{T} = \{\tau_1, \tau_2, \ldots, \tau_N\}$ are a series of activities among suspicious users, tweets, and things with a time stamp associated with each link between two entities in social networks.*

Definition 3. *Influence of neighbours* **Z**. *In addition to Q_{i-1}, the evolution of Q_i is also influenced by other users in social networks (such as friends, followers, etc.) through activities like* `post, reply, like, retweet, comment, etc.`

In this work, we use Z_i represent the influence of the neighbours, which can be described by

$$\begin{aligned} P(Q_i | Q_{i-1}, Z_i) \quad = \quad & P_0(Q_i | Q_{i-1}) \cdot (1 - \phi(Z_i)) \\ & + P_1(Q_i | Q_{i-1}) \cdot \phi(Z_i) \end{aligned} \tag{3}$$

in which $\phi(Z) : Z \mapsto [0, 1]$ is assumed as the simplified capture of the evolution of Q_i.

Since Z_i is a function of the activity of all the neighbours, assume that Z_i is dependent on Q_{i-1} from its past history, however it is related to Z_{i-1}. Then we have

$$P(Z_i | Q_i^{i-1}, Z_1^{i-1}) = P(Z_i | Q_{i-1}) \tag{4}$$

Eq. (4) presumes that user aggregation de-correlates Z_i from its past history.

Each HMM provides new information corresponding to the node in the Bayesian networks and saves the influence results back into the network. It can be seen that according to the $\mathcal{T}_i$ and Z_i, we can forecast the $\mathcal{T}_{i+1}$, it is of immense significance in tasks such as *potential attacks, advertising, anomaly detection, etc.* in social networks. A simple posteriori (MAP) predictor of the form [3]

$$\hat{\tau}_{i+1}|_{map} = \arg\max_T f(\tau_{i+1} = T | \tau_i, \mathbf{Z}_i) \tag{5}$$

More details about the HMMs optimization and prediction can be found in [27]. In the model, each HMM corresponds to a node in BRG and the observation can directly be updated using the prior probability. In social networks, the observation $\mathcal{T}$ might be imperfect and in practical some new approaches can be used (such as data mining, etc.) to improve the parametrized HMMs.

2) Bayesian Risks Graph Model: In our previous work [2], we described a multilayer Bayesian risk graph model that consists of three layers and each of them forms a directed acyclic graph based on the featured risks. The link between nodes denotes the probabilistic causal dependencies and each node maintains one or more conditional probabilities table (CPT). In this work, we applies similar scheme but the nodes can maintain dynamic CPT(s). The interconnected BRG contains three subnetworks: (1) intra-network; (2) inter-network, covers the connections between two adjacent networks; (3) cross-network, the links between behavior network and static network. Virtual nodes can be added between cross-network and inter-network to reduce the relation space and computation complexity.

In HBRG, the HMMs can estimate the transition probabilities of risks in CPTs as described above. The sophisticated HMMs can provide accurate CPTs evaluation to test whether a risk can cause other effect or make contribution on other risks. In this work, we use the `Twitter` dataset to statistically learn risk features at the static, dynamic, and behavioural networks to build accurate CPTs.

In Figure 1, a typical hierarchical Bayesian Risk graph (HBRG) is proposed, which integrates a hierarchical risk analysis architecture into a Bayesian risk graph. Figure 2 shows an example of HMM, where Z_i denotes the aggregated activity of neighbouring nodes.

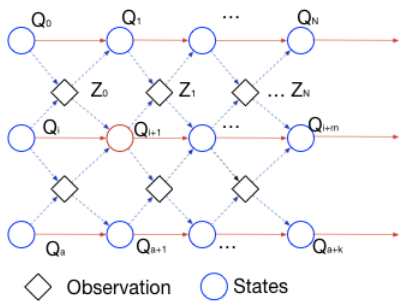

Fig. 2. HMM Model

Figure 3 shows a two-layer HBRG model, in which the top layer BRG Layer addresses a Bayesian Risks Graph (BRG)

that can provide a friendly risk evaluation framework. The HMM layer in the bottom can well describe the dynamic temporal patterns of user activity in social networks.

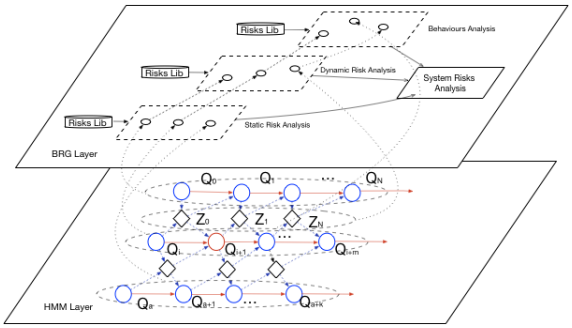

Fig. 3. Two-layer HBRG

In social network risk modelling, one or more vulnerabilities may cause more than one threats (risks). The risk states or its propagation are usually constructed as a Bayesian graph model $\mathbf{G} = \{\mathbf{V}, \mathbf{E}\}$, where $\mathbf{V}$ and $\mathbf{E}$ denote variable risks and links between risks, respectively. The edges in $\mathbf{G}$ can be modelled with local conditional probability distributions. For a node $v \in \mathbf{V}$, a conditional probability distribution $P(v|Pa(v))$ is used to describe the transition.

The process that one or more vulnerabilities propagate to one (or more) different threat(s) could be defined as a dependence graph. The risk states or its propagation are usually constructed as a DAG and the transition between nodes could be modelled with local conditional probabilities. In social networks, the BRG can model, analyse, and predict risks.

Definition 4. *Dynamic Conditional probability table (CPT). In a BRG, one needs to specify the qualitative parameters. Each node maintains a dynamic CPT to describe the conditional probability distribution for a particular combination of values of its parental nodes.*

Definition 5. *Hierarchical Risk Template. In social networks, we group the risks into following three categories:*

1) *Behavior Risk*, which includes social network behaviors that might cause potential attacks or privacy leakage, etc., such as `profile collections, tweets mining, etc.`.
2) *Dynamic Risk*, it involves potential activities that attempt to compromise security in social networks.
3) *Static Risk*, it includes the potential risks and threats based on the static analysis, such as malware, size analysis, permission analysis, virus matching, etc.

The risk template provides BRG a template that describes the basic potential risks of an activity profile in social networks.

The independent influence can be modeled as

$$\begin{aligned} \mathcal{P}(\tau_i|Q) &= \frac{P(\tau|Q)}{P(\tau)} \\ &= \frac{P(\tau|Q)P(Q)}{P(\tau|Q)P(Q) + P(\tau|\overline{Q})P(\overline{Q})} \end{aligned} \tag{6}$$

in which $P(Q)$ denotes the prior probability of states and $P(\tau)$ the posterior probability of observations. The aggregated influences can be modeled as

$$\begin{aligned} \mathcal{P}(\tau|Q) &= P(\tau|Q_1Q_2)P(Q_1)P(Q_2) \\ &\quad + P(\tau|Q_1\overline{Q}_2)P(Q_1)P(\overline{Q}_2) \\ &\quad + P(\tau|\overline{Q}_1\overline{Q}_2)P(\overline{Q}_1)P(\overline{Q}_2) \\ &= P(\tau|Q_1Q_2)P(Q_1)P(Q_2) \\ &\quad + P(\tau|Q_1\overline{Q}_2)P(Q_1)P(\overline{Q}_2) \end{aligned} \tag{7}$$

It is clear that the causal risks could be a compound of a set of attacks. To well model the joint effect of risks in social network, a *risk* classifier is effective. There are many types of attacks that can occur in a certain social network process. In HBRG, we model them as behaviour risks. We use the observation (Obsv) to determine which risk or attack occurred. In a dynamic social network, it is difficult to determine which attack occurred, since each attack can cause any value of Obsv. However, we can use the Obsv to determine the most likely event that occurred and we return that as an answer.

IV. Risk Assessment with BRGs in Social Networks

As discussed before, the proposed HBRG model is able to model the risks in social networks. In this section, a specific scenario is proposed to demonstrate the effectiveness of HBGR by analysing security issues in online social networks. In this work, we use the 'Chorus-CDT' to collect data through the `Twitter` Stream API based on a set of keywords. The keywords are derived from the 'user activity profile', and a list of 'cybersecurity terms' [2], [3], [20].

1) keywords: In practice, the security keywords can be extracted based on a trained security vulnerability keywords extractor (SVKE). The SVKE can be trained using text from security blogs, common vulnerabilities and exposures description, official security bulletins, etc.

2) Filtering and cleaning data: The tweets data should be preprocessed using a `security filter`. In this work, we collected many tweets with embedded links from 6 Jan 2017 to 17 Jan 2017. The reason that most spam and malicious messages are sent out with embedded links. The collected tweets are pre-precess before do the analysis. Two methods are used to identify malicious tweets, (1) The trend microweb reputation blacklist; (2) we use keywords developed in Section II. The tweets are analysed by extracting fields `Content, Links, Hash tags, sender data` and its `frequency`.

3) Probability: The above HMM model is used to model the dynamic user activity profile. The states of user are unobservable and the observation is based on these activities.

A. Attacks in Social Networks

Specifically, we consider following nine most common attacks in Twitter: *Fake like-jacking*; *Fake plug-in/sharing scams*; *Fake app*; *Malware attacks*; *Phishing attacks*; *Evil Twin attacks*; *Identity theft*; *Cyberbulling*, and *physical threats*.

1) Like-hijacking: the basic idea is to use fake `Like` buttons, attackers trick users into clicking website buttons that install malware, spreading attacks. This attack attempts to get user to copy and paste JavaScript or a link into their browser is a big scam-warning sign.

2) Fake plug-in/sharing/offer scams: in social networks, users can be tricked into downloading fake browser extensions that can pose like legitimate extensions but stealing data, including passwords, auto-filled form, and other information from the infected system. This kind of scams can be spotted if they offer to provide additional features, fake offers or messages, to trick users to install fake plug-in, extension, or join a fake event or group on the social network.

3) Fake apps: Fake apps have risen since 2013. The apps appear to be legitimate but often they contain some malicious payload that purported to convince them, it could used to harvest data, aggressive advertising tactics to sell the user's data, browsing habits to a third-part advertising network, etc.

4) Malware attacks: it comes in many shapes, sizes, and purposes ranging from viruses, spyware, and bots. In this attacks, the malware can either be infection or concealing, the former malware can spread and replicate itself from one users to the next. The concealment malware includes Trojan horses, rootkits, backdoors, and keylogger, etc. The XSS attack (Cross-site script) can forces a user's web browser to execute an attack's code.

5) Phishing attacks: it attempts to obtain sensitive information such as passwords, usernames, credit card details, etc. for malicious reasons. The phishing attacks always connect to account hack, spamming links, malicious url, re-tweets, etcs. The activity was tied to a hack, resulting in hundreds of identical states updates to particular band profiles. The resulting activity includes re-entering the log information (usernames, passwords), receiving volumes of spam within the accounts.

6) Evil Twin attacks: namely impersonation. This kind of attacks are increasing in social networks. It impersonates users while using that profile for financial gain, defamation, cyberbulling, physical crimes, and personal identifiable information gathering. A user can protect its account by settings or networking configurations such as Twitter has four privacy levels that.

7) Identity theft: is becoming an increasing attack in social networks since it is easier to perform but very dangerous. The identity theft is related to attacks such as dumpster diving, account hijacked, profile theft, email scams, and password re-usages, etc.

8) Cyberbulling: or cyberharassment, is on the increase in social networks, especially among child, pre-teen, or teenagers. The cyberbulling behaviour can include unlike informations, such as pages, images, links, or behaviours such as posting rumours, threats, disclose victim's privacy, pejorative labels, etc. This attack is always tied to attacks such as identity theft, fake scams, etc.

9) Physical threats: the on-line attacks can put people in physical risks. An attacker can gain access through technical means and physical means by bypassing security control, such as, leakage of proprietary information, gain access, block social network, etc.

TABLE I
RISKS AND POSSIBLE CAUSALS IN TWEETS

Compound risks	Possible causal risks	Atom risks
Fake Followers	Profile attack Like-jacking Information gathering	Profile attack Like-jacking Information Gathering
Fake Plug-in/offer	DDoS Botnets Sniffing MitB API Attacks	DDoS Botnets Sniffing MitB API Attacks
Spamming	Email-based Spam Context-aware spam MitB	Email-based spam Botnets
Malware	Profile attack API Attack Password attack MitB Worms XSS Like-jacking	Email-based attack MitB MitB Like-jacking Password attack
Phishing	Identity theft Profile attack API Attack Password attack MitB Worms XSS Like-jacking	Profile attack Password attack like-jacking API Attack MitB Worms XSS
Evil Twin	Profile attack Cyberbulling Physical attack Dumpster diving Identity theft	Dumpster diving Botnets Profile attack
Identity theft	Profile attack XSS Phishing Information gathering Email-based attack Password attack	Email-based Attack Password Attack XSS Profile attack Information gathering
Cyberbulling	Harassment Identity theft Phishing Password attacks Information gathering	Information gathering Password attacks Harassment
Physical threats	Profile attack Cyberbulling Information gathering Physical access Identity theft Phishing Password attacks	Physical access Password attacks

An attacker can combine the above attacks to perform a complicate attack by behavior analytics, automation, machine learning, and other intelligent capabilities that have the ability to do really complex math in a millisecond are just a few of the layers needed to detect and prevent identity fraud.

B. Number results

Based on the collected Tweets, we summarized the possible attacks and causal in Table. I, in which the first colum address the main risks in this dataset, the second column addresses the possible causal or linked attacked, and the third column address the atom risks that cannot be divided further. Based on this table, a CPT can be easily set up.

TABLE II
RISKS AND POSSIBLE CAUSALS IN TWEETS

Risks	Possible causals	New atom risks
Fake Followers	Profile attack Like-jacking Information gathering	Profile attack Like-jacking Information Gathering

In this example, the probabilities of attacks `Profile attack, Like-jacking, information gathering` can be derived according to the model in Section III. A CPT can be set up and the probability of `Fake Follwer` happened can be derived as shown in Figure. 4, in which the probability of `Profile attack` happened is $\{high : 30.7\%, middle : 49.1\%, low : 20.7\%\}$, the probability of attack `like-jacking` found is about 56.6%, and the `Information gathering` happened with a set of probabilities $\{content : 24.8\%, features : 35.6\%, frequency : 39.5\%\}$, then the probability of attack `Fakefollower` happened is 58.9%.

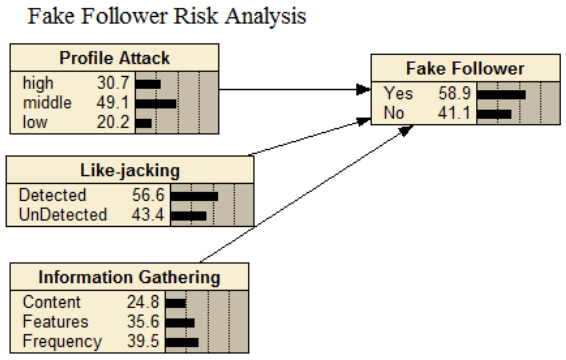

Fig. 4. Risk analysis for `Fake Follwer Attacks`

Fig. 5. shows the risk and attack analysis for the collected data, in which if the attacks *profile attack, like-jacking, information gathering* are confirmed occurred, then the probability that attack `Fake Follower` happened is 97.3%

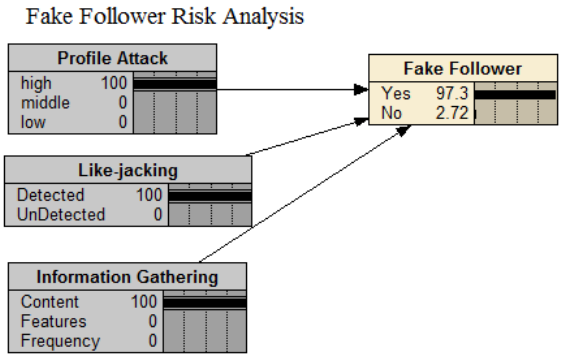

Fig. 5. When all causual attacks happened, the probability of `Fake Follwer Attacks` happens

Figure. 6 shows an example of the CPT at node `Fake Follower`, which can be dynamically updated by the HMMs.

Profile Attack	Like-jacking	Information Gathering	Yes	No
high	Detected	Content	97.282	2.718
high	Detected	Features	95.931	4.069
high	Detected	Frequency	80.395	19.605
high	UnDetected	Content	75.448	24.552
high	UnDetected	Features	71.397	28.603
high	UnDetected	Frequency	70.393	29.607
middle	Detected	Content	58.93	41.07
middle	Detected	Features	58.811	41.189
middle	Detected	Frequency	58.71	41.29
middle	UnDetected	Content	58.385	41.615
middle	UnDetected	Features	54.058	45.942
middle	UnDetected	Frequency	51.13	48.87
low	Detected	Content	48.82	51.18
low	Detected	Features	42.1	57.9
low	Detected	Frequency	40.046	59.954
low	UnDetected	Content	17.668	82.332
low	UnDetected	Features	9.495	90.505
low	UnDetected	Frequency	7.378	92.622

Fig. 6. CPT at `Fake Follwer Attacks`

Since in social network risks analysis, the proposed system is able to dynamically analyse the risk and attacks according to the real-time observations. It is also possible to anticipate the risks for next stage according to the collected social network data.

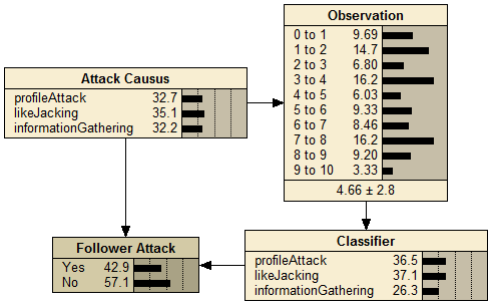

Fig. 7. Example of risk analysis for `Fake Follwer Attacks`

In Fig. 7, an example is introduced to show how to dynamically evaluate the final risks that an attack happened by applying a classifier. The `Follower Attack` could be caused by three attacks: *profile attack, like-jacking, information gathering*. In the classifier, we defined the experiential rules that can classify the causal attacks based on the observations.

One of the goals of this work is to obtain a high-level understanding of various types of attacks on Twitter. The proposed model can analyse the risks or attacks from different viewpoints, for example, time ranges, keywords, users, tweets, etc. Figure 8 shows the BRG model for risks described in Table I.

Table III shows the test results, in which $P(pa)$ denotes the probability of 'Profile Attack', $P(lj)$ denotes the probability of *Like-jacking*, and $P(ig)$ denotes the probability of *information gathering*. The $P(FF)$ denotes the probability that 'Fake Follower' happened. It can be seen, in round 1, when all three attacks happened, the probability of 'Fake Follower' happened is 97.3%. In round 2 to round 6, we tested how the $P(FF)$ changed when $P(pa)$, $P(lj)$, and $P(ig)$ changed.

V. DISCUSSION & CONCLUSION

This paper proposed a HRGB model for risk analysis in social networks, which can take risk analysis based on the dynamic activity patterns. The proposed model integrates the

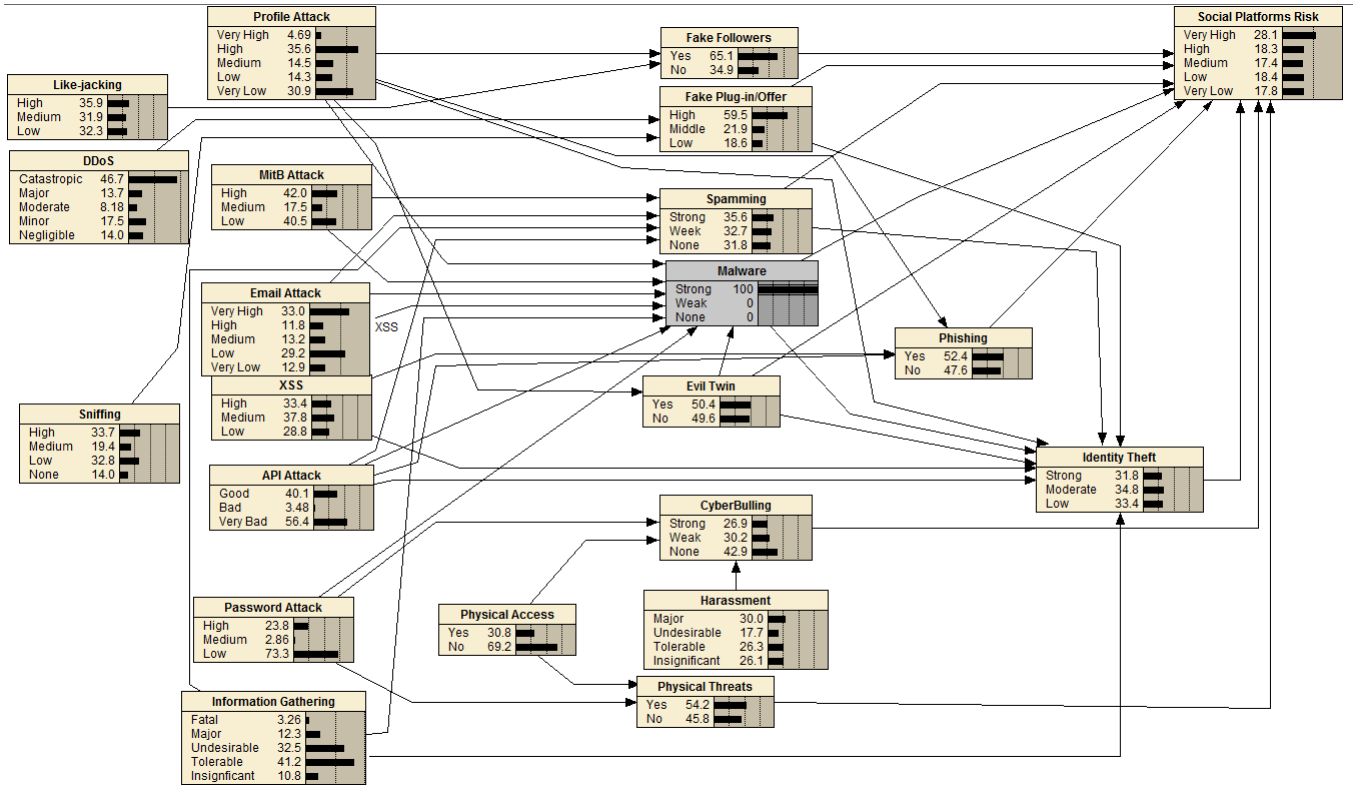

Fig. 8. BRG Model for `Twitter` Attacks

TABLE III
TEST RESULTS FOR 'FAKE FOLLOWER' ATTACK

	$P(pa)$	$P(lj)$	$P(ig)$	$P(FF)$
r_1	100	100	100	97.3
r_2	30.7	56.6	24.8	58.9
r_3	30.7	60.4	24.8	59.5
r_4	30.7	56.6	39.5	60.0
r_5	20.2	56.6	24.8	53.4
r_6	49.1	56.6	24.8	63.6

HMMs and Bayesian Risk Graph model to provide real-time risk evaluation. It can also retrieve the causes of attack or foresee potential attacks in the dynamic networks. We are continuing to refine the HBRG model to further improve its performance.

REFERENCES

[1] Christian Von Der Weth, Ashraf M. Abdul, Mohan Kankanhalli, "Cyber-Physical Social Networks", *ACM Transactions on Internet Technology*, vol.17, no.2, 2017.

[2] Shancang Li ; Theo Tryfonas ; Gordon Russell ; Panagiotis Andriotis, "Risk Assessment for Mobile Systems Through a Multilayered Hierarchical Bayesian Network", *IEEE Transactions on Cybernetics*, vol.46, no.8, pp.1749-1759, 2016.

[3] Haiying Tu, Jeffrey Allanach, Sanam Singh, Krishna R. Pattipati, and Peter Willett, " Information integration via hierarchical and hybrid Bayesian networks", *IEEE Transactions on Systems*, vol.36, no. 1, pp.19-33, 2006.

[4] David Rosenblum, What Anyone Can Know: The Privacy Risks of Social Networking Sites, *IEEE Security & Privacy*, vol.5, no.3, pp.40-49, 2007.

[5] Zheng Yan; Mingjun Wang, Protect Pervasive Social Networking Based on Two-Dimensional Trust Levels, *IEEE Systems Journal*, vol.11, no.1, pp.207-218, 2017.

[6] Panagiotis Andriotis; George Oikonomou; Theo Tryfonas; Shancang Li, Highlighting Relationships of a Smartphones Social Ecosystem in Potentially Large Investigations, *IEEE Transactions on Cybernetics*, vol.46, no.9, pp.1974-1985, 2016.

[7] Markus Riek; Rainer Bohme; Tyler Moore, Measuring the Influence of Perceived Cybercrime Risk on Online Service Avoidance, *IEEE Transactions on Dependable and Secure Computing*, vol.13, no.2, pp.261-273, 2016.

[8] Yanan Guo; Dapeng Tao; Weifeng Liu; Jun Cheng, Multiview Cauchy Estimator Feature Embedding for Depth and Inertial Sensor-Based Human Action Recognition, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol.47, no.4, pp.617-627, 2017.

[9] Beatrice Lazzerini; Francesco Pistolesi, Multiobjective Personnel Assignment Exploiting Workers' Sensitivity to Risk, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol.pp, no.99, 2017.

[10] Martin Mulazzani, Markus Huber and Edgar Weippl, Social Network Forensics: Tapping the Data Pool of Social Networks

[11] Yanling Chang; Alan L. Erera; Chelsea C. White, Risk Assessment of Deliberate Contamination of Food Production Facilities, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol.47, no.3, pp.432-440, 2017.
http://tech.co/5-surprising-cybersecurity-risks-on-social-media-2016-09

[12] Symantec, Cybercriminals Target Social Networks to Do Their Dirty Work, Available on 31 Mar 2017, https://www.symantec.com/connect/blogs/cybercriminals-target-social-networks-do-their-dirty-work

[13] Stal Atl Imx, Cyber-Crime and Social Networking, Avaialble on 31 Mar 2017, http://www.stlatlimxpolice.ca/cyber-crime–social-networking.html

[14] http://www.defenseone.com/technology/2016/07/how-putin-weaponized-wikileaks-influence-election-american-president/130163/?oref=search-social

[15] Amy J. C. Trappey; David W. Hsiao; Lin Ma, Maintenance Chain Integration Using Petri-Net Enabled Multiagent System Modeling and Implementation Approach, *IEEE Transactions on Systems, Man, and Cybernetics: Cybernetics*, vol.41, no.43, pp.306-315, 2011.

[16] Bruna Freitas ; Ashraf Matrawy ; Robert Biddle , Online Neighborhood Watch: The Impact of Social Network Advice on Software Security Decisions, *Canadian Journal of Electrical and Computer Engineering*, vol.39, no.4, 2016.

[17] Mengyuan Li; Ruan Na; QiYang Qian; Haojin Zhu; Xiaohui Liang; Le Yu, SPFM: Scalable and Privacy-preserving Friend Matching in Mobile Cloud, *IEEE Internet of Things Journal*, vol.pp, no.99, 2016.

[18] M. Fire, D. Kagan, A. Elyashar, and Y. Elovici, Friend or Foe? Fake Profile Identification in Online Social Networks, Social Network Analysis and Mining, 2014, Vol. 4, No. 1, pp.1-23.

[19] Martin Mulazzani, Markus Huber, and Edgar Weippl, "Social Network Forensics: Tapping the Data Pool of Social Networks", Available on 31 Mar 2017, http://www.sba-research.org/wp-content/uploads/publications/socialForensics-preprint.pdf.

[20] Vasanthan Raghavan; Greg Ver Steeg; Aram Galstyan; Alexander G. Tartakovsky, Modeling Temporal Activity Patterns in Dynamic social networks, *IEEE Transactions on Computational Social Systems*, vol.1, no.1, pp.89-107, 2014.

[21] Haiying Tu; Y. N. Levchuk; K. R. Pattipati, Robust action strategies to induce desired effects, *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol.34, no.5, pp.664-680, 2004.

[22] Highlighting Relationships of a Smartphones Social Ecosystem in Potentially Large Investigations, IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol.45, no.8, pp.1125-1137, 2015.

[23] William Ross; Alex Gorod; Mihaela Ulieru, A Socio-Physical Approach to Systemic Risk Reduction in Emergency Response and Preparedness, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol.45, no.8, pp.1125-1137, 2015.

[24] James M. Tien, An SMC Perspective on Big Data: A Disruptive Innovation to Embrace, *IEEE Systems, Man, and Cybernetics Magazine*, vol.1, no.2, pp.27-29, 2015.

[25] Robert Westervelt, Social Networking Attacks You Should Dodge, Available on 31 Mar 2017, http://www.crn.com/slide-shows/security/240163136/top-5-social-networking-attacks-you-should-dodge.htm

[26] Chao Gao; Jiming Liu, Network-Based Modeling for Characterizing Human Collective Behaviors During Extreme Events, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol.47, no.1, pp.171-183, 2017.

[27] J. Allanach, H. Tu, S. Singh, P. Willett, and K. R. Pattipati, Detecting, tracking, and counteracting terrorist networks via hidden Markov models, *Proc. IEEE Aerospace Conf.*, Big Sky, Mt, Mar. 2004.