Online Social Networks Privacy Study Through TAPE Framework

Yongbo Zeng, *Student Member, IEEE,* Yan (Lindsay) Sun, *Senior Member, IEEE,* Liudong Xing, *Senior Member, IEEE,* and Vinod Vokkarane, *Senior Member, IEEE*

***Abstract*—While personal information privacy is threatened by online social networks, researchers are seeking for privacy protection tools and methods to assist online social network users. In this paper, we propose a Trust-Aware Privacy Evaluation framework, called TAPE, aiming to address this problem. Under the TAPE framework we investigate how to quantitatively evaluate the privacy risk, as a function of people's awarenesses of privacy risks as well as whether people can be trusted by their friends to protect others' personal information. Simulations are performed to illustrate the key concepts and calculations, as well as the advantages of TAPE. Based on the TAPE framework, we also propose an unfriending strategy in terms of privacy protection, which outperforms other existing unfriending strategies.**

***Index Terms*—Privacy, Trust-Awareness, Online Social Networks**

I. Introduction

WITH the emergence of **O**nline **S**ocial **N**etworks (OSN), people are facing critical privacy risks. In OSN, personal information can be abused, which will put users into risks. Researchers identified OSN privacy issues as two categories, inadvertent disclosure of personal information, and stalking or backtracking [1], [2]. Krishnamurthy et al. studied the problem of personal identity information leakage and how it can be misused by third parties [3]. This kind of information is able to distinguish an individual's identity either alone or when being combined with other information that is linked to a specific individual, and its leakage will lead to identity theft. Livingstone [4] demonstrated the risks when young people make friends and share personal information to express themselves online. Real life stories of sensitive information leakage in OSNs happen frequently. For example, most employers began to collect potential employees' information using social networks. According to a survey released on the EU Data Protection Day [5], privacy leakage had put people's careers on risk.

In the current commercial OSN design, privacy risk is unavoidable, due to the publicity of OSN [6], [7]. In order to benefit from the convenience of OSNs, people share personal information with friends, which makes privacy leakage possible. When privacy risk is unavoidable, we assume a risk and attempt to reduce the likelihood of harmful events. Under the assumption of unavoidable risk, risk analysis becomes extremely important. According to the National Institute of Standards and Technology (NIST), risk analysis is defined as "the process of identifying risk, assessing risk, and taking steps to reduce risk to an acceptable level" [8]. In [9], In et al. discussed the advantages of risk analysis, including designing secure privacy management, monitoring and protecting critical data, making privacy policies, etc.

Risk analysis can be performed either quantitatively or qualitatively. Quantitative risk analysis plays a critical role. The advantages of quantitative risk analysis are discussed in [10], such as using metrics to evaluate risk parameters, analyzing risk events, making sophisticated decisions.

Solutions for addressing privacy issue include educational aspect and technical aspect. For example, in [11], Gundecha et al. studied privacy issues and protection recommendations, for the purpose of educating OSN users and raising their privacy awareness. The technical aspect includes managing privacy setting [12] and adopting new architectures to build OSNs [13] [14]. There are also some other categorizations in the literature. For example, in [15], Jeckmans et al. categorized privacy research into 5 categories, including raising user awareness, law and regulations, personal information anonymization, perturbing user information, and data encryption. We attempt to solve privacy issue from another perspective – providing a quantitative privacy risk analysis framework for OSN users and researchers. In the OSN privacy research literatures, quantitatively analyzing privacy risk is still not mature, as we discuss in Section II. Therefore, we propose a framework in this paper to quantitatively evaluate the privacy level of OSN users. We believe that the proposed framework can help people to understand their privacy situations, raise their privacy awareness and thus reduce privacy risks.

Quantitatively evaluating privacy level in OSN is a challenging task. First, quantitative user privacy level is not a well defined concept in OSNs. Second, human users play an important role in the personal information leakage. It is complicated to predict an individual user's behavior. Third, personal information can be propagated through both online and offline media by many ways, such as chatting, emails, instant messages, Facebook postings, picture postings, tweets on Twitter, etc. Fourth, it is extremely difficult, if not impossible, to obtain the ground truth about a user's privacy level, with which the quantitative evaluation results can be compared.

In this paper, we address the *first challenge* by proposing quantitative definition of privacy risk, based on privacy hazard and its probabilities. This quantitative measurement will lead to the privacy level calculation tools, which were originally proposed in the reliability analysis field. To address the *second and third challenge*, we have to consider the availability of social data. Since nobody can monitor the users' all communication behaviors (online and offline), researchers have to work on limited data, which can be obtained with reasonable

costs. In this work, Facebook privacy setting is used as the primary data source. We also focus on the 'word-of-mouth', which is the primary drive of OSN information diffusion [16]. Although other privacy leakage scenarios, which we discuss in Section VI-G1, are not considered in this work, the proposed concepts, including privacy awareness and privacy trust, can be extended to those scenarios. For the *fourth challenge*, due to the lack of ground truth of users' privacy level, we compare the proposed scheme with some existing approaches, such as privacy concern model in [17] and vulnerability analysis in [11]. In addition, Monte Carlo simulation is employed to verify the results of privacy risk evaluation.

In this paper, we propose a TAPE (Trust-Aware Privacy Evaluation) framework for quantitatively evaluating users' privacy level in OSNs. The TAPE framework contains several novel aspects.

- It finds the similarity between the reliability analysis in wireless sensor networks (WSN) and the privacy risk estimation in OSNs. It sets up the stage for utilizing reliability analysis tools for privacy analysis.
- It considers the privacy leakage through nodes (i.e. users) and through links (i.e. friend connections) separately. Here, the privacy leakage through nodes mainly depends on the users' behavior, and we define two metrics in TAPE to estimate it. The first one reflects whether one respects others' privacy, and it is named as *Privacy Awareness*. The other one reflects how much one's friends trust her/him in terms of not gossiping their information to others, and it is named as *Privacy Trust*. The privacy leakage through a link mainly depends on the relationship between the two users, in terms of whether one paying attention to the other' personal information.
- It proposes the desirable properties of privacy awareness and privacy trust metrics, as well as specific ways to calculate them under the guidance of trust management theory. It is the first time that the privacy trust concept has been used in evaluating privacy level in OSN.

Besides privacy risk estimation, the TAPE framework has the ability to conduct **sensitivity analysis** for friend links, which is similar to the concept of vulnerability in [11]. Through the sensitivity analysis, an OSN user can understand how much his/her privacy level is affected by a particular friend connection. The sensitivity analysis yields a practical way to improve OSN users' privacy level.

As a summary, the **contributions** of this work include: (a) the TAPE framework, which considers privacy leakage through nodes and links separately and utilizes traditional reliability analysis tools, as well as the definition of privacy risk, in a quantitative way; (b) the privacy awareness and privacy trust metrics; (c) a privacy awareness algorithm, which shows a clear advantage over the know algorithm called IRT [17] in the current literature; (d) the sensitivity analysis metric, from which we propose an efficient unfriending strategy.

This paper is organized as follows. Related work is discussed in Section II. TAPE framework is described in Section III, followed by discussion of information spreading probability algorithms and the proposed algorithms in Section IV. Privacy assessment and sensitivity analysis metric are presented in Section V. Experiment results and conclusion are presented in Section VI and Section VII respectively.

II. Related Work

Privacy in OSN have attracted many attentions. OSN service providers allow users to manage who can access which information (e.g. in Facebook and Google+), and to hide sensitive information to non-connected users (e.g. in Linkedin). Researchers studied privacy protection from two directions. Along the *first direction*, fundamental changes to the current design of OSN were suggested to enhance users' privacy. Felt et al. [14] studied and discussed the privacy concerns of social network APIs for third parties. Guha et al. [18] proposed an approach to hide user data by mapping real data to fake data. Within the first direction, "Privacy by Design" (PbD) is an important approach. In [19], Wolf et al. operationalized the concept of PbD through the process of design and development of OSN, and several social requirements of OSN were identified to optimize the privacy from a user perspective. Encryptions are usually used when adopting PbD. For example, in [13], Baden et al. proposed a new type of OSNs by using attribute-based encryption to hide user data, in which symmetric keys are used to encrypt messages and only the designated friend groups can decrypt the messages. In [20], Erkin et al. proposed to use homomorphic encryption and multi-party computation techniques to hide privacy-sensitive data from the service provider in a recommender system, without losing the significant usability of data. The *second direction* is developing privacy protection tools based on existing OSNs. For example, Fang et al. [12] developed privacy wizards to give user recommendation for privacy setting, and Gundecha et al. [11] proposed an approach to identify a user's vulnerable friends. In this paper, we propose to assist users' privacy protection by providing quantitative evaluation of privacy risk and conducting sensitivity analysis for friend links. Our work belongs to the second direction.

There have been several quantification models for privacy evaluation in OSN. Alim et al. [21] proposed a vulnerability quantification model which consists of three components: individual, relative and absolute vulnerabilities. They examined the visibility of OSN users' profiles and computed the clustering coefficient to compose individual vulnerability. Based on individual vulnerability, relative vulnerability and absolute vulnerability were calculated. Besides privacy risk evaluation, friend vulnerability analysis, also referred to as sensitivity analysis in this paper, is considered to be a good way to improve personal privacy. Abdulrahman et al. proposed a node vulnerability metric [22] and a multi-agent vulnerability analysis [23] based on the friendship graph of MySpace. Vulnerability index (V-Index) was proposed to measure how vulnerable an OSN user is based on her/his friends' privacy setting [11]. Privacy setting and its implications were considered as a primary factor in the existing models. In this paper, we consider privacy setting as one of the primary factors. The implications of privacy setting are represented as two metrics – privacy awareness and privacy trust. Besides privacy setting, the TAPE framework is able to adopt social tie analysis

approaches when implementing the module of link information spreading probability. The network topology and information diffusion patterns are also considered.

The proposed work is also related to information diffusion in OSN. Gruhl et al. [24] studied the dynamics of information spreading in weblogs. Adar et al. [25] demonstrated a technique for inferring information propagation through a blog network by applying epidemic models of information spreading. Cha et al. [16] studied social cascades over Flickr social network. Researchers also attempt to build mathematical model to solve problems of information diffusion in online social network, such as [26]–[28]. In addition, there are literatures discussing the relationship between tie strength and information propagation [29], which is related to the information spreading probability that is discussed in this paper. Different from the previous information diffusion work, the proposed TAPE framework considers information diffusion in the context of privacy protection, which requires different set of features and considerations.

III. Trust-aware Privacy Evaluation Framework

In this section, the TAPE framework is discussed in details. We first define privacy risk from the perspective of information diffusion. The binary decision diagram (BDD) which was commonly used for system reliability analysis is employed to calculate privacy risks. The concepts of node information spreading probability and link information spreading probability are proposed.

A. Acronyms

OSN	Online Social Network
PIO	Personal Information Owner
UD	Undesirable Destination
UG	Undesirable Group
PA	Privacy Awareness
IPA	Individual Information Privacy Awareness
PT	Privacy Trust
ISP	Information Spreading Probability
NISP	Node Information Spreading Probability
LISP	Link Information Spreading Probability
WSN	Wireless Sensor Network
BDD	Binary Decision Diagram
BM	Birnbaum's Measure
IRT	Item Response Theory

B. Notations

I_j	Alice's type j personal information
UG	Alice's undesirable group related to I_j
UD_i	Alice's ith undesirable destination in UG
V_j	Privacy risk for I_j
Z_j	Privacy leakage hazard for I_j
L_j	Privacy leakage probability for I_j
$ISP(.)$	Calculation of ISP
S	The set of Alice's privacy setting
s_j	Alice's privacy setting for I_j
PA_u	User u's PA
$G_{PA}(.)$	Calculation of PA
$rank^+_{u,j}$	Proportion of users whose privacy setting for I_j is looser than user u

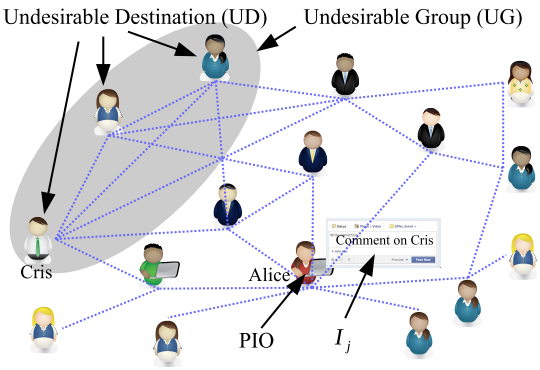


Fig. 1: Online social network of Example 1.

$rank^-_{u,j}$	Proportion of users whose privacy setting for I_j is tighter than user u
$\langle A, B\rangle$	Friend link between A and B
PT_u	User u's PT
PT_{u,f_i}	PT evaluation based on friend f_i's recommendation
$G_{PT}(.)$	Calculation of PT
$T_{A,B}$	Evaluation of how much A trusts B
R^+_u	Positive recommendations for user u

C. Online Social Network Privacy

Some OSNs (e.g. Facebook and Linkedin) encourage people to use real names and upload personal information onto a page known as 'Profile'. Such personal information is often accessible by friends directly, and can even flow to thousands of other people through retweet (e.g. on Twitter), sharing (e.g. on Facebook) and online communication (e.g. chatting). The privacy concern in OSNs is well known, but how can we define the privacy risk in a quantitative way?

Before discussing quantification of privacy risk, we first look at two examples.

Example 1. *Alice is a student, and she posted a piece of comment complaining her teacher Cris on her social network site. Alice does not want Cris and other teachers in the same department to know the comment. Figure 1 shows the example social network.*

Example 2. *Alice posted a photo, and she does not want anyone, except her friends, to see this photo.*

Generally, in some scenarios, we want some personal information to be known only by friends, and in some other scenarios we don't want certain personal information to be viewed by specific people [2]. In Example 1, the personal information concerned by Alice is her comment on Cris, and in Example 2, the personal information is her photo. It is clear that an user has different types of personal information, and that the privacy concerns depend on the particular type of personal information. We introduce the notation I^u_j to denote user u's type j personal information. Without loss of generality, we present the framework in the context of protecting Alice's privacy, i.e. u="Alice". Alice is also referred to as the **personal information owner (PIO)**. In the rest of the paper, for simplicity, we use I_j to represent I^{Alice}_j.

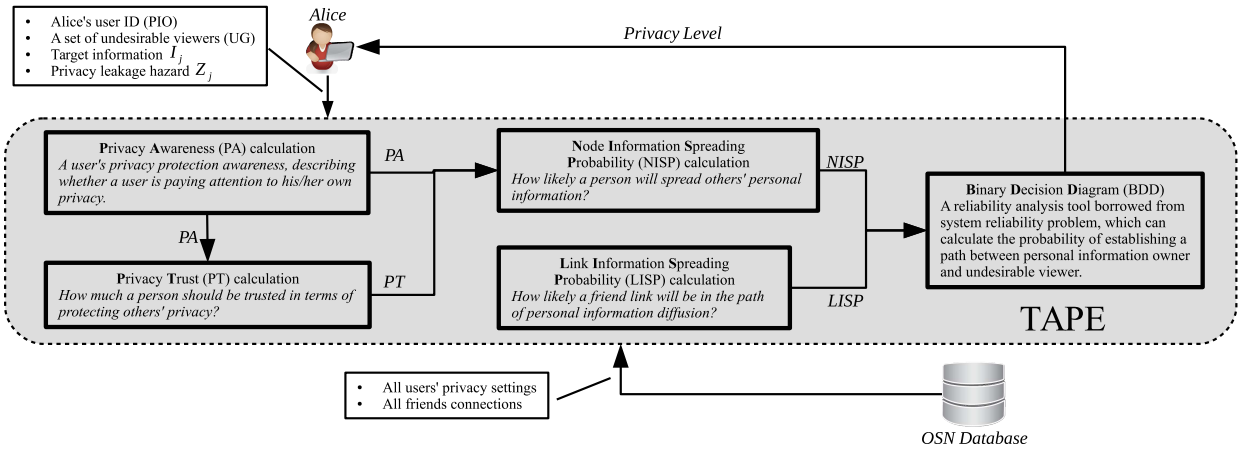


Fig. 2: Core structure of the TAPE framework

It is noted that privacy concerns are related to the "undesirable viewers". We define the concept of **Undesirable Group (UG)** of I_j, denoted by UG_j, as follows. If Alice does not want her information I_j to be seen by user u_i, then u_i is put into UG_j, where u_i is also called **Undesirable Destination (UD)** of I_j. In Example 1, Alice's UG is $\{\text{Cris}\}$. In Example 2, Alice's UG contains all users except her friends.

In other words, if I_j flows to any UD, Alice considers her privacy of I_j being violated and **privacy leakage** happens. In the rest of the paper, for simplicity, we use UG to represent UG_j^{Alice}.

Information leaking to different persons has different potential risks to the PIO. Without defining UD, this difference cannot be captured. The privacy definition based on UD is a more generalized definition. In this definition, users are classified into 4 types:

1) personal information owner (e.g. Alice),
2) users who are allowed to access personal information according to the privacy setting (e.g. Alice's friends),
3) users to whom the exposure of personal information causes damage (i.e. the undesirable group)
4) users not belonging to the above three types.

In the existing work, people usually assume that there are no type 4 users, such as in [11]. The definition of privacy leakage in this paper becomes the traditional definition as long as the UD is defined as the complement set of the type 1 and type 2 users. Our definition can also handle the cases that Alice only concerns that the privacy leaks to a specific set of users, as seen in Example 1. In other words, our definition can capture the fact that privacy leaking to different persons has different damage to the PIO. Such difference is usually not captured by the privacy setting alone.

D. Privacy Risk and Related Concepts

With the proposed TAPE framework, we aim to answer two questions: 1) Can we measure the probability of personal information leakage as a measurement of privacy risk in OSN? 2) How is the personal information leakage related to privacy risk? In this subsection, we first introduce the key concepts of the TAPE framework.

In [15], privacy is considered as keeping a piece of information in its intended scope. In TAPE, the leakage of personal information I_j occurs when any users in the undesirable group UG_j view I_j. Here the undesirable group is the same as the unintended scope in [15]. We assume that I_j can only be obtained through online information diffusion, which only occurs through friend connections. This assumption is a result of the limitation of data, as discussed in Section I. In the future, if more data are available, such as cell phone contact data, this assumption can be revised. Due to this assumption, the UG in Example 2 can be simplified as $\{$*all of Alice's 2-hop neighbors*$\}$. We define **privacy leakage probability** of I_j, denoted by L_j, as the probability that at least one UD views I_j through information diffusion in the OSN.

$$L_j = Pr\{privacy\ \ leakage\ \ happens\} \tag{1}$$

In statistics, the notion of risk is often modeled as the expected value of an undesired outcome [30]. That is

$$\begin{aligned} Risk = &(probability\ \ of\ \ the\ \ accident\ \ occurring) \\ &\times (expected\ \ loss\ \ in\ \ case\ \ of\ \ the\ \ accident) \end{aligned} \tag{2}$$

In the context of OSN, we argue that privacy risk of information I_j, denoted by V_j can be computed as

$$V_j = L_j \cdot Z_j \tag{3}$$

where L_j is privacy leakage probability as defined in (1) and Z_j describes the expected loss/damage in case of privacy leakage. In this paper, we also use another term "privacy level" to describe an individual's privacy, and obviously, the lower privacy risk is, the higher the privacy level is. In TAPE, Z_j is referred to as **privacy leakage hazard** and is normalized within interval $[0, 1]$. We argue that Z_j should be determined by the PIO (e.g. Alice) when the damage of the privacy leakage is subjective. For instance, in Example 1, Alice may be the best person who determines the damage if Cris saw her complain on Facebook? In many cases, PIO is often the best person to estimate the damage/loss of privacy leakage. In some other cases, PIO may not have the knowledge to determine the damage. For example, in Example 2, Alice may not understand the consequence of revealing her photo to strangers. In such case, Alice needs assistance to determine the damage, such as the privacy leakage problem study in [3]. In this paper, we simply assume that Z_j can be provided by PIO. In the rest of

this paper, when we compare privacy risks, Z_j is considered as constant 1. Based on this assumption, the privacy leakage probability L_j is equivalent to privacy risk V_j. *The core task in TAPE is to estimate the privacy leakage probability* L_j.

E. Toward Privacy Leakage Probability Estimation

In TAPE, a social network is represented by an undirected diagram. OSN users are the nodes, denoted as u_i, and their friend connections are the links, denoted as $\langle u_m, u_n \rangle$. As discussed in the previous section, personal information can be diffused to unintended recipients through friend links. It is important to point out that the existence of a link does not necessarily mean the personal information will be transmitted through it. For instance, in Example 1, Alice posted a piece of comment, but some of her friends may not read it. Here are three typical situations.

- Alice's friend Bob does not pay attention to Alice's posting at all. Alice's comment does not disseminate to Bob through the link between Alice and Bob.
- Alice's friend Bob pay attention to Alice's posting and read it. Alice's comment disseminates to Bob through the link between Alice and Bob. Then, Bob respects Alice's privacy and does not tell others about Alice's comment. In this case, Alice's comment does not disseminate to others through Bob.
- Bob reads Alice's comment, and retweets it. Such retweeting can be seen by all Bob's friends. In this case, Alice's comment disseminates to Bob, and it is possible to disseminate to others through Bob.

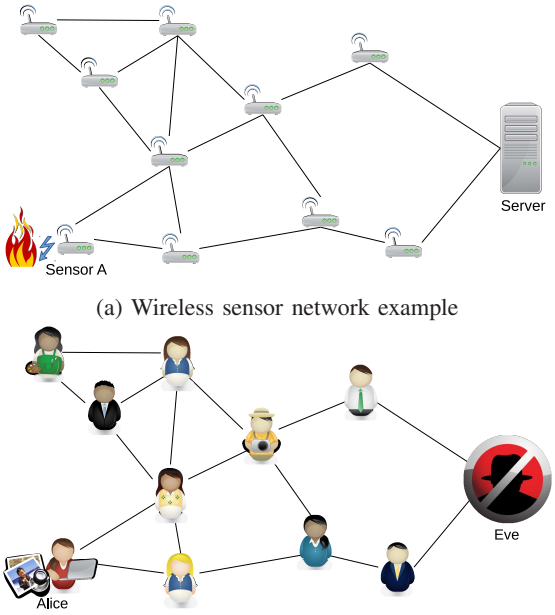
We argue that the privacy leakage probability estimation problem can be decomposed into two tasks.

1) **Task 1:** The first task is to estimate the probability whether a user's personal information will be disseminated through a particular component c (a link or a node). In this work, such probabilities are referred to as **information spreading probabilities** (ISP), denoted by $ISP(c)$. The ISP of the link between Alice and Bob depends on factors such as whether Alice and Bob are good friends, whether Bob actively communicates with Alice in OSN, and whether the information is interesting enough to catch Bob's attention. The ISP of a node is determined by complicated factors, ranging from knowledge to personality, which is extremely difficult to quantify or even understand.
2) **Task 2:** The second task is to compute the probability of privacy leakage (i.e. L_j), given the network topology, the information spreading probabilities of links and nodes, the PIO (i.e. Alice), and the UG.

In the rest of this section, we first discuss the solution to the second task (Section III-F), and then present the solution to the first task (Section IV). Figure 2 shows the core structure of the TAPE framework.

F. Privacy Analysis and Reliability Analysis

When investigating information diffusion in OSN, we found reliability graph, which has been used as one of the reliability analysis tools (e.g. WSN reliability), can be adapted to solve the problem. An example is shown in Figure 3.



(a) Wireless sensor network example

(b) Online social network example

Fig. 3: Similarity between WSN and OSN. In 3a, Sensor A detects fire, and the detection will be sent to the server through other sensor nodes. In 3b, Alice (PIO) feels her photos are improper to be viewed by Eve (UD).

TABLE I: Concepts mapping

Reliability Analysis for WSN	TAPE
System reliability	Privacy leakage probability
Reliability graph	Social graph
Source node	Personal information owner (PIO)
Destination	Undesirable destination (UD)
Node/edge failure probability	1 - node/link information spreading probability (ISP)

In a reliability analysis problem, the system is represented by a **reliability graph**, whose links and nodes are assigned **failure probabilities**. The system has a **source** node and a **sink** node that is usually a station. If there is no path from the source to the sink can be established, the system fails. For example, in a WSN, the nodes are sensors, and the links are the communication channels. A sensor's failure probability depends on its battery, environment temperature, work load, etc. A communication channel's failure probability depends on distance, environment noise, etc. In the context of reliability analysis, one often needs to estimate the probability that there is at least one path can be established from the source sensor to the destination sensor [31].

In the TAPE framework, we have defined the information spreading probability for nodes and links in the previous section. This concept is kind of "opposite" to the failure probability. For example, if node A fails to forward data to its neighboring nodes with probability x, node A's failure probability is x in the context of WSN reliability analysis,

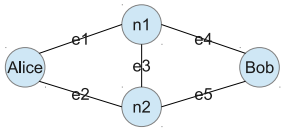


Fig. 4: An OSN information diffusion example.

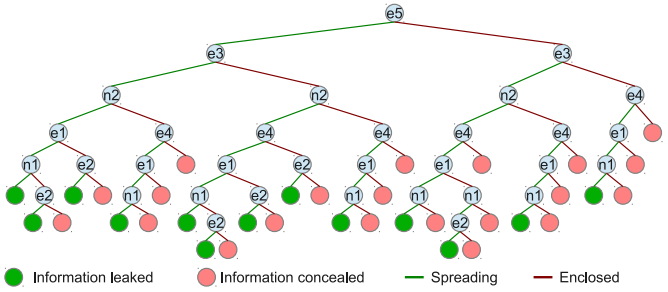


Fig. 5: BDD graph of the example in Figure 4

whereas this node's information spreading probability is $1-x$ in the context of privacy analysis. The goal of WSN is to transmit data successfully, whereas the goal of privacy protection is to prevent personal information from propagation. Therefore, in the TAPE framework, we can also define failure probability of nodes/links as $1-ISP$. We propose to use the binary decision diagram (BDD) method, which is commonly used in reliability analysis [31]–[33], to solve Task 2 described in Section III-E. Table I shows the important concepts in TAPE, as well as the concepts mapping.

A BDD is a directed acyclic graph created based on Shannon's decomposition. It is an efficient tool to manipulate boolean expressions. For example, in Figure 4, Alice is PIO and Bob is UD. All nodes and links are assigned ISPs. In order to calculate the information leakage probability L_j, we first use a boolean expression to represent L_j.

$$\begin{aligned} L_j =& ISP(e_1)ISP(n_1)(ISP(e_4)+ISP(e_3)ISP(n_2)ISP(e_5)) \\ &+ ISP(e_2)ISP(n_2)(ISP(e_5)+ISP(e_3)ISP(n_1)ISP(e_4)) \end{aligned} \tag{4}$$

Then, a BDD graph is constructed based on the reliability expression. The BDD graph is a binary tree (Figure 5), each sub-tree is considered to be a sub-expression. The left sub-tree of a BDD node represents the expression when the node successfully spreads information. The right sub-tree represents the expression when the node fails to spread information. When traversing from the root to a leaf node, if the leaf node is a left child, then it gives a information leakage case. Based on the BDD diagram, we can evaluate L_j using a recursive method. The details of the BDD approach can be found in [31].

In this work, BDD is employed to compute the probability of information diffusion after modification. Due to the large size of the social network and the high computation cost of BDD, we adopt a **reduced BDD algorithm**. In particular, we set the maximum traversing depth as k times the number of hops between PIO and UD. For example, when $k=2$ and the UD is 3 hops away from the PIO, the branches longer than 6 (3×2) are discarded from the BDD graph. In Section VI, we set $k=2$.

G. Summary

By studying the similarity between the reliability analysis in WSN and the privacy risk estimation, we modify the BDD method to evaluate information leakage probability. The concept of node ISP and link ISP are developed. The core structure of the TAPE is shown in Figure 2. As a summary, TAPE is presented as a framework to solve task 2 described in Section III-E. In Section IV, we discuss details of ISP calculation. Particularly, the metrics of privacy awareness and privacy trust are proposed for node ISP calculation.

IV. Information Spreading Probability Algorithms

While most social network information diffusion models consider the impact of nodes and links together [25], we argue that information propagation through nodes and through links should be considered separately. This is why we define **information spreading probability of node (NISP)**, also referred to as node ISP, and **information spreading probability of link (LISP)**, also referred to as link ISP, which can better describe the information diffusion process. NISP is the probability that a node will spread others' information, and LISP is the probability that a link will be in the path of information diffusion. NISP and LISP imitate the nature human communication process in the real world (i.e. offline social network).

- NISP describes the probability of speaking, i.e. talking about others.
- LISP describes the probability of listening, i.e. hearing what is said.

In this section, we focus on the algorithm of NISP, followed by a brief introduction of the LISP algorithms proposed in literatures.

A. Node Information Spreading Probability (NISP)

Evaluating NISP of a person is very challenging, because it is related to one's knowledge and personality. In the offline social network, we probably can estimate the NISP of a person based on experiences if we know this person well. Obviously, such estimation can be biased and limited, and most importantly cannot be applied in OSNs due to data limitation. Instead of resolving a challenging problem in social science, we propose to examine NISP based on the quantitative data available in OSNs.

In particular, we propose two metrics that should be used to estimate NISP – privacy awareness and privacy trust.

1) Privacy Awareness: The first metric is **privacy awareness (PA)**, which depends on a user's privacy setting. We argue that privacy setting reflects a user's privacy protection awareness, describing whether a user is paying attention to his/her own privacy. There are many different ways to compute a user's PA. In TAPE, *PA evaluation is a module*. The input is a set of the user's privacy setting, which is represented as $S_u = \{s_{u,j} | j = 1,2,\ldots,J\}$ where u is the user and $s_{u,j}$ is the privacy setting for information I_j. Privacy setting has options $\{v_1, v_2, \cdots, v_M\}$, in which v_n is a looser setting than

TABLE II: Desirable properties of PA algorithms

Special Cases	$rank_1^+$	$rank_1^-$	**Desirable PA value when normalizing to** $[0, 1]$
Case 1: Alice's privacy setting is looser than all others'	0	≈ 0	$PA_{Alice} \approx 0$
Case 2: Alice's privacy setting is tighter than all others'	≈ 1	0	$PA_{Alice} \approx 1$
Case 3: Everyone has the same privacy setting	0	0	$PA_{Alice} = 0.5$
Case 4: Many users (including Alice) have loose setting, and a few users have tight setting	0	small	$PA_{Alice} < 0.5$, and it should be small, but not too small because most people share the same opinion as Alice.
Case 5: A few users (including Alice) have tight setting, and many users have loose setting	big	0	$PA_{Alice} > 0.5$, and it should be higher than PA_{Alice} in case 7.
Case 6: A few users (including Alice) have loose setting, and many users have tight setting	0	big	PA_{Alice} should be smaller than PA_{Alice} in case 4.
Case 7: Many users (including Alice) have tight setting, and a few users have loose setting	small	0	PA_{Alice} should be high, but not too high because most people share the same opinion as Alice.

v_m for $n < m$, and correspondingly, $s_{u,j}$ can have a value in $\{1, 2, \ldots, M\}$. For example, in Facebook, v_1 = 'everyone', v_2 = 'networks and groups', v_3 = 'friends of friends', v_4 = 'friends', and v_5 ='self'. Let I_1 be the photo in Example 2. Alice's privacy setting for I_1 is 'friends', i.e. $s_{Alice,1} = 4$. Without loss of generality, in the rest of this paper, we use S and s_j to represent Alice's privacy setting set and privacy setting respectively.

Let PA_u represent PA of user u and G_{PA} represent the adopted PA calculation, then

$$PA_u = G_{PA}(S_u) \tag{5}$$

The TAPE framework can accommodate many PA algorithms. However, what are the design criteria for PA algorithms? Based on the possible distributions of privacy setting and Alice's possible adoptions, we identified seven special cases and the desirable PA values in these special cases in Table II, which serves as a guidance for the PA algorithm design. To better understand Table II, we define $rank_{u,j}^+$ as the proportion of users whose privacy setting for I_j is looser than u, and $rank_{u,j}^-$ as the proportion of users whose privacy setting for I_j is tighter than u. As long as we know the statistics of users' privacy setting for I_j and the adoption of u, we can compute $rank_{u,j}^+$ and $rank_{u,j}^-$.

Example 3. *Table III shows the statistics of birthday (I_1) privacy setting as an example. Alice allows only her friends to see her birthday.*

In Example 3, $s_1 = 4$, $rank_{Alice,1}^+ = 0.5+0.4+0.1 = 0.55$, and $rank_{Alice,1}^- = 0.1$.

In the rest of this section, we look at the insights of spacial cases in Table II. We assume people can apply privacy setting for one type of information in Table II. However, it is easy to extend it to multiple types of information.

Case 1 and **Case 2:** In fact, these two cases rarely happen in real life. We use them to demonstrate the extreme cases in PA calculations. In Case 1, Alice chooses to open her information in OSN, while others choose to hide it, which may indicate that the information is sensitive and releasing this type of information does not benefit the PIOs. In this case, Alice should get a minimum PA value due to the disclosure of sensitive information. On the other hand, if Alice chooses to hide the information, while others open it. It may means Alice is prudent when deciding to open information. Thus, Alice should get a maximum PA value in case 2.

TABLE III: An example: privacy setting statistics for birthday

Privacy setting	Proportion of users adopting this privacy setting
'everyone'	5%
'networks'	40%
'friends of friends'	10%
'friends'	35%
'self'	10%

Case 3: Without further evidence, it is difficult to interpret one's PA. Therefore, Alice has a neutral PA. In addition, if an action made by the majority, without further evidence, the action should get neither a significant negative nor positive assessment, e.g. neutral PA in TAPE.

Case 4 and **Case 5:** To see the insight of case 4 and case 5, we look at an example. Assume many people release birthday information to friends because they want to remind friends about their birthdays, even if they know the privacy risk. In this case, if Alice releases her birthday, her PA should not be largely reduced. On the other hand, if Alice hides her birthday, her PA should be relatively high, since in order to gain better privacy protection, she gives up the opportunity of receiving more birthday gifts and greetings.

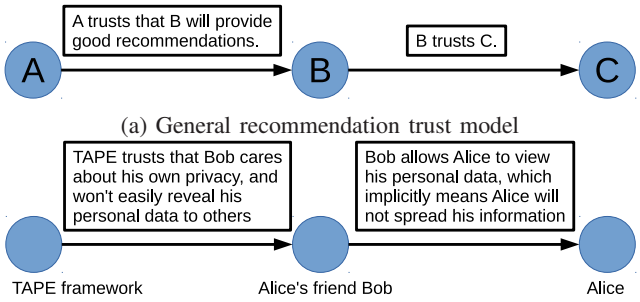
Case 6 and **Case 7:** Many people using tight privacy setting may imply that the information is sensitive. If Alice adopts a loose privacy setting for this type of information, her PA is low. On the other hand, if Alice adopts a tight setting, she should get a larger PA but relatively smaller than that in case 5.

We have to point out that Table II may not include all possible cases. For example, if Alice adopts a tight setting for birthday and a loose setting for phone number, and Bob adopts a loose setting for birthday and a tight setting for phone number, it is difficult to compare Alice's PA with Bob's. In such case, we need more data to make the PA evaluation more accurate. At current stage, we argue that those desirable properties in Table II provide a satisfying guidance for the PA algorithm design.

2) PA Algorithm: In our proposed PA algorithm, **individual information privacy awareness (IPA)** is calculated first. IPA is the PA value calculated from privacy setting of one type of information. Let $IPA_{u,j}$ denote the IPA of u for I_j.

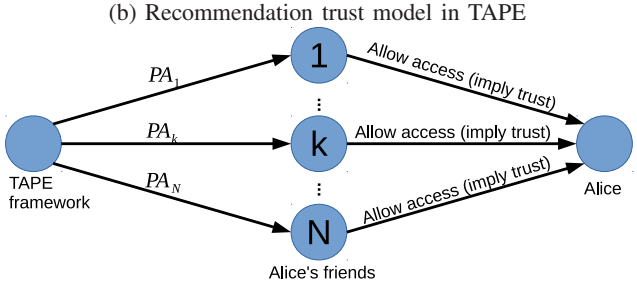
$$IPA_{u,j} = \frac{1}{2}(rank_{u,j}^+ - rank_{u,j}^-) + \frac{1}{2} \tag{6}$$

where $rank_{u,j}^+$ and $rank_{u,j}^-$ are defined in Section IV-A1. It is easy to verify that (6) satisfies the desirable properties in



(a) General recommendation trust model

(b) Recommendation trust model in TAPE



(c) Multiple recommendations in TAPE trust model

Fig. 6: Privacy Trust model

Table II. As an example, $IPA_{Alice,1} = 0.725$ in Example 3. Obviously, $0 \leq IPA_{u,i} \leq 1$. In fact, people can develop more sophisticated calculation to replace (6), according on the implementation environment and data availability. After calculating the IPAs for all types of information, the PA of u is calculated by

$$PA_u = \frac{1}{J}\sum_{j=1}^{J} IPA_{u,j} \tag{7}$$

In the literature, there are some approaches proposed to evaluate similar metrics. For example, in [17], the Item Response Theory (IRT) was used for modeling "privacy concern". In Section VI-A, we compare the proposed PA algorithm with IRT privacy concern model.

3) Privacy Trust: We propose another metric to evaluate how much a person should be trusted in terms of protecting privacy. Because this metric reflects how much one's friends trust her/him in terms of not gossiping their information to others, it is named as **privacy trust (PT)**. In fact, this type of trust is very difficult to evaluate based on direct evidence. *First*, direct evidence is rarely available, because we cannot wait someone to commit bad behaviors (e.g. gossip others) before estimating PT. *Second*, the clues that people use to determine whether a person is trustworthy in offline social networks are usually not available in OSNs. Alternatively, indirect evidence is used to predict OSN users' PT. Such indirect evidence can be established based on recommendations [34]. For example, Figure 6a shows a typical recommendation based trust model. If A trusts B, and B gives a recommendation saying that she/he trusts C, then A is able to develop certain level of trust to C.

In TAPE, we propose to evaluate an individual's PT based on implicit recommendations from her/his friends. For example, an implicit recommendation for Alice from her friend Bob is established when Bob allows her to access his personal information. Moreover, if Bob has a high PA value, it implicitly tells us that Alice may be trusted not to propagate others' personal information (Figure 6b). In real life, a person working on privacy research (e.g. myself) usually has good attention to the protection of privacy, thus this person has high PA. If I choose to tell someone my personal information, it means that I trust this person not to release my personal information to others. Although such implicit recommendations have noises, it may be the best resource to compute PT in OSNs.

PT calculation in TAPE is a module whose inputs are PAs of the user's friends and trust evaluations that how much the user is trusted by friends. Let PT_u represent PT of u and G_{PT} represent PT calculation, then

$$PT_u = G_{PT}(PA_{friends\ of\ u}, T_{friends,u}) \tag{8}$$

where $T_{friends,u}$ indicates how much u is trusted by friends. Similar to PA calculation, we argue that the PT calculation should follow 3 rules.

Rule 1. *The level of PT depends on the number of positive recommendations. The recommendation from a high PA friend is considered to be a positive recommendation. The more positive recommendations a user gets, the higher her/his PT should be.*

Rule 2. *Negative recommendations should be carefully used. On the one hand, if a user with low PA trusts Alice, this should not affect Alice's PT either positively or negatively. On the other hand, if a user with high PA does not allow Alice to view personal information, it is not sufficient to indicate Alice is trustless.*

Rule 3. *Although each additional positive recommendation can increase the PT, such incremental diminishes when the number of positive recommendations is getting very large. For example, when positive recommendation number increases from 3 to 6, PT can increase a lot. However, when the number increases from 300 to 303, PT should not increase as much as the earlier case.*

4) PT Algorithm: In the literature, there are many trust models. We adopt a trust model using Beta function to address concatenation propagation and multi-path propagation of trust [34].

In the context of PT, the recommendation accuracy (arrow from A to B in Figure 6a) is replaced by PA of node B, and the trust value (arrow from B to C in Figure 6a) is the implicit trust of B towards C, represented by $T_{B,C}$ in TAPE. For simplicity, in the paper, we set $T_{B,C}$ as a constant value, by assuming that when two users are connected in OSN, they have certain chance to see each other's personal information, but it does not necessarily mean they have already read or will read that information. In the future, when more OSN data is available, such as the nuanced privacy setting, $T_{B,C}$ can be calculated more accurately.

We use R_u^+ to denote the set of positive recommendations for u, i.e. u's friends whose PA values are higher than a threshold (ϵ^+). The PT calculation we adopt is described as follows.

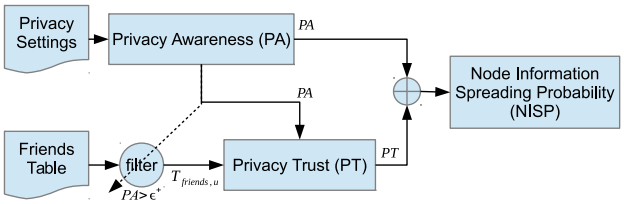


Fig. 7: Node ISP Calculation Diagram

First, we estimate PT through one recommendation path

$$PT_{u,f_i} = PA_{f_i} T_{f_i,u} + (1 - PA_{f_i})(1 - T_{f_i,u}) \tag{9}$$

where $f_i \in R_u^+$ is ith high PA friend of u. Then the variance of the estimation of PT_{u,f_i} is calculated

$$\sigma^2_{u,f_i} = \begin{cases} \frac{PT_{u,f_i}(1-PT_{u,f_i})^2}{2-PT_{u,f_i}}, & PT_{u,f_i} > 0.5 \\ \frac{PT^2_{u,f_i}(1-PT_{u,f_i})}{1+PT_{u,f_i}}, & PT_{u,f_i} < 0.5 \end{cases} \tag{10}$$

The PT is calculated using Beta trust model

$$PT_u = \frac{a}{a+b} \tag{11}$$

where $a = \sum_{i=1}^{|R_u^+|} a_i - 1$, $b = \sum_{i=1}^{|R_u^+|} b_i - 1$ and

$$a_i = PT_{u,f_i}\left(\frac{PT_{u,f_i}(1-PT_{u,f_i})}{\sigma^2_{u,f_i}} - 1\right) \tag{12}$$

$$b_i = (1 - PT_{u,f_i})\left(\frac{PT_{u,f_i}(1-PT_{u,f_i})}{\sigma^2_{u,f_i}} - 1\right) \tag{13}$$

5) Calculation of NISP: PA and PT are surely two important factors that determine NISP. However, PA and PT metrics are not probability values. No theories can be used to compute NISP from PA and PT. In this work, we adopt a heuristic approach, which estimates the NISP as a weighted average of PA and PT,

$$ISP(u) = w \cdot PA_u + (1 - w) PT_u, \tag{14}$$

where w is a weight factor between 0 and 1. Here weighted average is one of the simplest ways to combine PA and PT. In fact, People can develop more complicated calculation depending on the implementation environment and data availability. In the experiments in Section VI, we choose $w = 0.5$. Figure 7 shows the diagram of the NISP calculation. In the future, real human users must be involved (e.g. questionnaire) to understand the relationship among NISP, PA and PT.

B. Link Information Spreading Probability (LISP)

As discussed earlier, LISP of the link between Alice and Bob depends on whether Bob heard what Alice said. Furthermore, it depends on whether Alice has a strong tie with Bob and whether the information is interesting enough to catch Bob's attention. In the current literature, many works have investigated social ties [29], [35]. Note that the TAPE framework can accommodate any algorithms for LISP calculation, as long as the outcome of LISP calculation is a value between 0 and 1 indicating the probability of information spreading. In this paper, we do not propose a specific algorithm for calculating LISP. In the experiments, we adopt a constant value for LISP and focus on demonstrating the impacts of PA and PT.

V. Privacy Assessment and Privacy Improvement through TAPE

By evaluating NISP and LISP, and utilizing the reliability analysis method, TAPE has the ability to assess one's OSN privacy level. More importantly, based on the privacy assessment process, TAPE is able to tell people the strategies of improving privacy level.

A. Privacy Assessment

As discussed in Section III, by utilizing the BDD method and adopting proper NISP and LISP algorithms, TAPE is able to evaluate privacy leakage probability from the PIO to the UD. In real life, people usually want to avoid certain personal information being viewed by multiple people, which is the reason why we define undesirable group. Without further modification, TAPE can solve multiple UD case. Given an undesirable group $UG = \{UD_1, UD_2, \ldots, UD_K\}, K > 1$, the information leakage probability to UG is

$$L^{UG} = 1 - \prod_{UD_i \in UG} (1 - L^{UD_i}) \tag{15}$$

where L^i is the privacy leakage probability to UD_i. Here, privacy leakage happens if any one UD gets the information.

B. Privacy Improvement Strategies

The goal of privacy protection in TAPE is to reduce privacy risk. From a user's perspective, the most practical strategy is to block a friend to access certain personal information, also referred to as **unfriending**. In TAPE, we develop a method that can identify the friend link which contributes to the privacy leakage the most. We adopt **Birnbaum's measure (BM)** [36] to find such a friend link.

Originally, Birnbaum's measure was used to examine the sensitivity or importance of a component in reliability graph. In TAPE, we use it to evaluate the sensitivity of a friend link. Birnbaum's measure evaluates the partial derivative of the leakage probability with respect to LISP of link c.

$$BM(c) = \frac{\partial L}{\partial ISP(c)} \tag{16}$$

For the single UD case, the detailed calculation of BM, which uses the BDD graph, can be found in [36].

We derive Birnbaum's measure for multiple UD case. By rewriting (15), we get

$$\log(1 - L^{UG}) = \sum_{UD_i \in UG} \log(1 - L^{UD_i}) \tag{17}$$

For the right-hand side,

$$\begin{aligned} &\frac{\partial}{\partial ISP(c)} \sum_{UD_i \in UG} \log(1 - L^{UD_i}) \\ &= \sum_{UD_i \in UG} \left(\frac{-1}{1 - L^{UD_i}} \frac{\partial L^{UD_i}}{\partial ISP(c)}\right) \end{aligned} \tag{18}$$

For the left-hand side:

$$\frac{\partial \log(1 - L^{UG})}{\partial ISP(c)} = \frac{-1}{1 - L^{UG}} \frac{\partial L^{UG}}{\partial ISP(c)} \tag{19}$$

Plugin together, the Birnbaum's measure of c for UG is

$$\begin{aligned} BM_{UG}(c) &= \frac{\partial L^{UG}}{\partial ISP(c)} \\ &= \sum_{UD_i \in UG} \left(\frac{1-L^{UG}}{1-L^{UD_i}} \frac{\partial L^{UD_i}}{\partial ISP(c)} \right) \\ &= \sum_{UD_i \in UG} \alpha_i BM_i(c) \end{aligned} \tag{20}$$

where $\alpha_i = \frac{1-L^{UG}}{1-L^{UD_i}}$ is the contribution weight of the ith UD, and $BM_i(c)$ is the Birnbaum's measure for c when computing the privacy leakage probability to UD_i. Finally, the unfriending strategy proposed in TAPE is to find a friend link, which is

$$c^* = \arg\max_c BM_{UG}(c) \tag{21}$$

TAPE suggests to block c^* to improve privacy level.

VI. Experiment Results and Discussion

TAPE framework is implemented in Matlab, and the experiments are conducted. The reduction factor of reduced BDD is 2, i.e. $k = 2$. In PT calculation, PA threshold (ϵ^+) is 0.5 and $T_{B,C}$ is 0.7. NISP values are calculated according to Figure 7. At the beginning of this section, we first do a case study to demonstrate the calculation of PA, and then we apply TAPE to two real OSN datasets. The privacy risks and friends sensitivities are calculated. Several unfriending strategies are compared.

A. Case Study

We compare two PA algorithms. The proposed PA algorithm, referred to as Rank PA, is described in Section IV-A2. The comparison algorithm is described in [17], referred to as IRT. Briefly speaking, this scheme calculates a metric called "privacy concern" based on privacy settings, by utilizing Item Response Theory. The goal is to estimate OSN users privacy concerns toward information sharing.

Since there is no ground truth on what should be the most "correct" value of PA, in order to demonstrate their major features, we compare these two schemes in special situations. Assume Alice has 3 types of information I_1,I_2 and I_3, and the related privacy settings are s_1,s_2 and s_3. The privacy setting is binary, either open (represented by 0) or hidden (represented by 1). The column index in Table IV is the possible privacy setting. We randomly generate privacy setting data as follows. For each special situation, we first specify the proportion of each privacy configuration (e.g. '000', '001', etc.), and then generate the privacy configuration realities according to the distribution. 10,000 privacy configuration realities are generated for each special situation. We conduct the case studies, and the special situations we investigate are follows.

Special Situation 1 - Most users (88.91% for I_1, 88.63% for I_2 and 89.35% for I_3) choose to open all types of information.

Special Situation 2 - Most users (89.32%) hide type 1 information. For type 2 and type 3 information, most user (88.75% for I_2 and 88.92% for I_3) open them to public.

Special Situation 3 - Most users (89.02% for I_1 and 88.82% for I_2) hide type 1 and type 2 information, and for type 3 information 89.20% users disclose it.

Special Situation 4 - Most users (88.51% for I_1, 88.48% for I_2 and 88.85% for I_3) hide all types of information.

The PA calculation are shown in Table IV, in which the proposed PA algorithm is referred to as "Rand PA" and the comparison scheme is referred to as "IRT z". We first investigate the range of each scheme. IRT z has narrow ranges for the studied situations, although the theoretical range of IRT z can be $(-\infty, \infty)$. In order to adopt IRT z in TAPE as a PA algorithm, non-trivial normalization is needed. On the other hand, the proposed Rank PA has a range from 0 to 1 as expected. In addition, the neutral value for Rank PA is 0.5, and neutral value for IRT z is 0. Then, we investigate both schemes according to desirable properties of PA in Table II and get follow observations.

1) The majority always get PA values close to neutral for Rank PA, i.e. "000" of special situation 1, "011" of special situation 2, "110" of special situation 3 and "111" of special situation 4. However, when IRT z is used, such majority behavior cannot be captured.
2) Special situation 1 corresponds to case 4 and case 5 in Table II. It is seen that both Rank PA and IRT z satisfy the desirable properties of case 4 and case 5.
3) Special situation 4 corresponds to PA case 6 and PA case 7 in Table II. Rank PA satisfies the desirable properties in PA case 6 and PA case 7. When look at "000" in special situation 4, IRT z gives a higher value than the same privacy setting in special situation 1, which violates the desirable properties.
4) When investigating column "000" in Table IV, it is expected that the PA values from top to bottom should change from neutral to small, because the more people adopting tight privacy setting may indicate that the information is more sensitive and opening it can yield lower PA values. Rank PA has such trend, while IRT z does not.

In addition, IRT z is designed for binary privacy setting. However, real privacy setting usually has more than two options, such as in Facebook. Additionally, compare to the proposed Rank PA, IRT z also suffers from higher computational cost.

B. Datasets

We use two datasets to conduct experiments.

Dataset I contains a small number of users with detailed privacy setting. In particular, the privacy setting of 514 Facebook users in the community of University of Rhode Island (URI) are collected through a survey, and their public friends relationships are obtained by a crawler. In dataset I, there are 16 types of personal information, such as 'email address', 'mobile phone number', 'education' and 'status and links'. Each type has 5 privacy setting options, including 'everyone', 'networks', 'friends of friends', 'friends' and 'self'.

Dataset II contains a large number of users with limited privacy setting information, constructed by the authors in [37]. It contains about 957,000 Facebook users. There are 4 types of

TABLE IV: Case study: Rank PA vs. IRT PA

Special Situations		Privacy setting: $s_1s_2s_3$ (0=open,1=hidden)							
		000	**001**	**010**	**100**	**011**	**101**	**110**	**111**
1	Percentage of users adopting this setting	**70.50%**	8.30%	8.92%	1.19%	8.83%	1.00%	1.10%	0.16%
	IRT z	-0.0648	0.1337	0.1337	0.3321	0.1337	0.3321	0.3321	0.5305
	Rank PA	0.4448	0.6115	0.6115	0.7782	0.6115	0.7782	0.7782	0.9448
2	Percentage of users adopting this setting	8.36%	1.18%	1.05%	0.09%	**70.53%**	8.68%	8.98%	1.13%
	IRT z	-0.0169	-0.0018	-0.0018	0.0133	-0.0018	0.0133	0.0133	0.0285
	Rank PA	0.3139	0.4806	0.4806	0.6472	0.4806	0.6472	0.6472	0.8139
3	Percentage of users adopting this setting	0.99%	0.17%	8.85%	0.97%	9.09%	0.93%	**70.27%**	8.73%
	IRT z	-0.3249	-0.1528	-0.1528	0.0194	-0.1528	0.0194	0.0194	0.1915
	Rank PA	0.1856	0.3523	0.3523	0.5189	0.3523	0.5189	0.5189	0.6856
4	Percentage of users adopting this setting	0.11%	1.20%	1.01%	9.17%	1.05%	9.16%	8.98%	**69.32%**
	IRT z	-0.0081	-0.0050	-0.0050	-0.0020	-0.0050	-0.0020	-0.0020	0.0010
	Rank PA	0.0569	0.2236	0.2236	0.3903	0.2236	0.3903	0.3903	0.5569

TABLE V: Datasets summary

	Dataset I	**Dataset II**
# of unique users	514	957,000
Average real degree	215.8	95.2
Average sampled degree	2.1	3.8
Max sampled degree	18	124
# of personal information types	16	4
Privacy setting options	'everyone' 'networks' 'friends of friends' 'friends' 'self'	'open' 'hidden'

personal information, including 'add as friend', 'photo', 'view friends' and 'send message'. Each type of information has 2 privacy setting options, 'open' and 'hidden'.

Information of the two datasets is listed in Table V.

C. Privacy Risk

It is well known that the reliability of data transmission can drop significantly as the distance (i.e. the number of hops) increases. In the context of privacy protection, does the privacy risk heavily depend on this distance? We study the relationship between the privacy risk and the distance from the PIO to UD.

We first randomly pick 100 nodes from dataset I and put them in the PIO set. In each round of simulation, we pick one node (without replacement) from the PIO set as the PIO, and pick another node from the network as the corresponding UD, which is no more than 6-hop away from the PIO. If the picked PIO is an isolated node (i.e. degree is 0), we skip it. For each pair of PIO and UD, we measure the distance, compute the privacy risk using TAPE, and plot the privacy risk in Figure 8. Each point represents one pair of PIO and UD. The x-axis indicates the distance between PIO and UD, and y-axis is the privacy risk. In this experiment, LISP is chosen from 0.5, 0.8, 0.9, and 0.95. We have the following observations

- As expected, when the distance increases, privacy risk has a decreasing trend.
- The privacy risk to 1-hop UDs (i.e. friends) can be greater than the LISP. This is because Alice's friends not only get Alice's information from Alice directly, but also through other paths. For example, Alice's friend Bob may not hear what Alice said, but he could get the message from Charlie who is another friend of Alice.
- When the distance is small, the privacy risk varies in a large range. The distance is not a dominating factor. The PA, PT and network topology jointly determine users' privacy risk. A user who is 3 hops away may be more likely to obtain Alice's person information than a user who is 2 hops away.
- As the LISP decreases, the privacy risk decreases. In the future work, incorporating the estimation of LISP will yield even a larger variation in the privacy risk values.

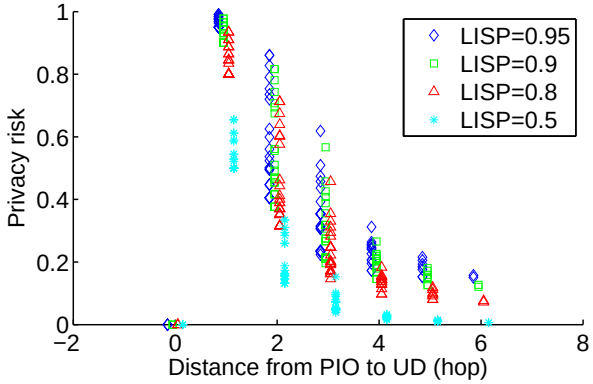
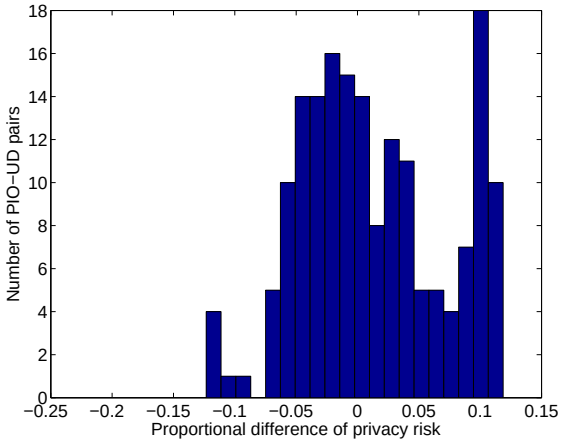


Fig. 8: Privacy risk vs. PIO UD distance.

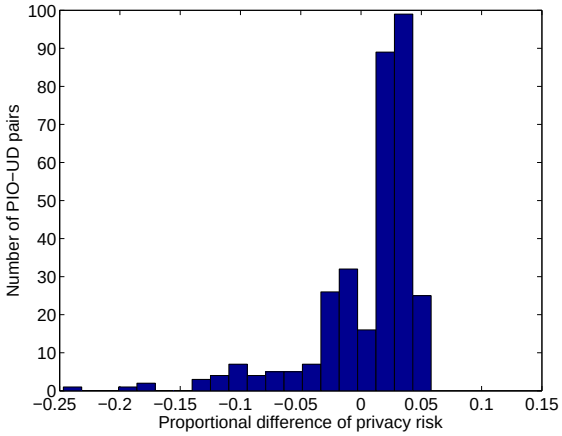
D. The impact of PA and PT

Since the lack of "ground truth" about the real privacy risk of users, it is hard to compare TAPE with other privacy evaluation methods that consider different features of the users. Instead of comparing TAPE with a specific method, we argue that a prevalent type of privacy study in OSN only focuses on network topology. We construct a comparison method, referred to as *topology-based method*, which uses the BDD to compute the privacy risk with fixed LISP and NISP. By comparing TAPE with the topology based method, we will see whether considering PA and PT metrics reveals more information that is not captured by considering the topology alone. In the experiment, we set the LISP to be 0.5, and set the NISP of the topology based method to be the average of the NISP values when considering PA and PT.

The experiment setup is similar to that in Section VI-C. We construct PIO sets for both dataset I and dataset II, and each set has 100 nodes. In each round of simulation, one node is picked up (without replacement) from the PIO sets as the PIO, and another node that is 3 hops away from the PIO is picked as UD. We calculate the privacy risk using TAPE and using the topology based method. We define proportional difference as $D = \frac{V^{Topology} - V^{TAPE}}{V^{Topology}}$, where $V^{Topology}$ is the



(a) Dataset I



(b) Dataset II

Fig. 9: Histogram of proportional difference

TABLE VI: Impact of LIST value when comparing TAPE and the topology-based method

LISP	0.2	0.35	0.5	0.65	0.8
Avg(D)	0.0129	0.0131	0.0130	0.0127	0.0123
std(D)	0.0590	0.0595	0.0592	0.0583	0.0571

privacy risk calculated based on topology, and V^{TAPE} is the privacy risk calculated by TAPE. The histograms of D for both datasets are shown in Figure 9. It is seen that the proportional difference range is from -25% to 5%. Hence PA and PT do provide additional and useful information beyond the topology. In addition, it is seen that dataset II shows more concentrated distribution around 0, and dataset I has a wider range. It is known that, dataset II has 4 types of personal information and each type has 2 privacy setting options, while dataset I has 16 types of personal information and each type has 5 privacy setting options. We argue that the comprehensiveness of privacy setting can impact the performance of TAPE.

In the topology-based method, we set the LISP to be 0.5 and the NISP to be the average NISP in TAPE. When choosing the NISP value, we argue that the NISP setting favors the topology based method. Particularly, when PA and PT are not available, it is very difficult to choose a proper NISP value for the topology based method. By choosing the average NISP

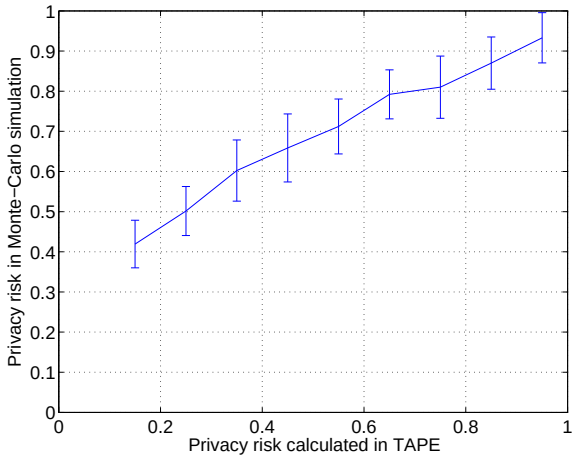


Fig. 10: TAPE vs Monte-Carlo simulation

value from TAPE, we believe that it will provide a reasonable NISP estimation for the topology based method. In the rest of this section, we conduct experiment to study how much the LISP value can impact the results when comparing TAPE and the topology-based method. The experiment setup is the same as the one using dataset I earlier in this section, and we repeat it by selecting one LISP value from $\{0.2, 0.35, 0.5, 0.65, 0.8\}$ at one time. The proportional difference D is calculated. We list the statistics of D in Table VI. The proportional difference does not change when we select different LISP. However, we have to point out that smaller LISP values will give smaller privacy risk estimations, and we already observed it in Figure 8.

E. Verification of TAPE Calculation

In the previous experiments, the privacy risks are calculated from LISP and NISP using BDD as described in Section III-F. In order to verify this calculation, Monte-Carlo simulations are used and the results are compared with the outputs of TAPE.

The simulation is conducted as follows. At the initial stage, a node is selected as PIO, and the PIO owns a token, which represents one type of personal information. During the simulation stage, we divide the time into T steps. At each step, every node with a token can pass duplicates of the token to its neighbors. The probability that node A successfully passes the token duplicate to its neighbor B is $ISP(A) \cdot ISP(\langle A, B\rangle)$, where $ISP(A)$ is the NISP of A and $ISP(\langle A, B\rangle)$ is the LISP of link $\langle A, B\rangle$. After T steps, the simulation is terminated. If there is any UD that obtains a token duplicate, this simulation is marked as 'information leakage observed'. By repeating the simulation N times, we will get N_1 'information leakage observed' simulations, and the simulated privacy risk is $\frac{N_1}{N}$.

In the experiment, we randomly select 1,000 PIOs from dataset II, and those whose degrees are less than 2 are skipped. The 3-hop privacy risks are computed using TAPE. Corresponding Monte-Carlo simulations are conducted, in which $N = 100$ and $T = 5$. Figure 10 shows the results of our experiments. It is seen that the simulation justifies the calculation of TAPE, because a strong linear relationship between simulated results and TAPE results is observed. It is noticed that the slop of the curve depends on the number of simulation steps, i.e. T.

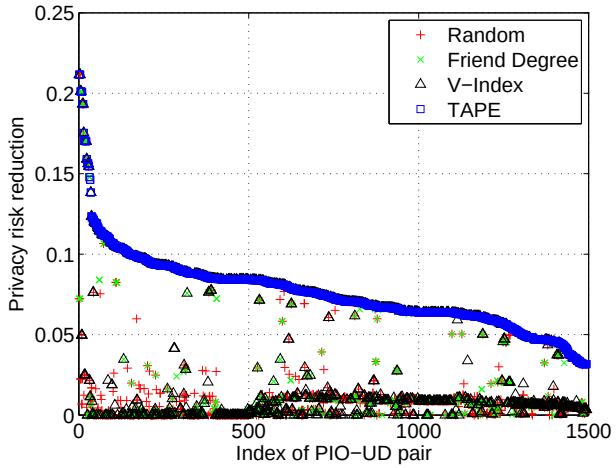


Fig. 11: Privacy improvement of unfriending

F. Sensitivity Analysis and Unfriending Strategy

Unfriending is suggested in [22], [38]. We propose an unfriending strategy based on Birnbaum's Measure, referred to as TAPE unfriending, which evaluates the partial derivative of the leakage probability with respect to the LISP of a given friend connection. In this section, we conduct experiments to compare TAPE unfriending with 3 unfriending approaches.

1) **TAPE:** In this approach, the friend link that has the largest Birnbaum's measure is blocked and the privacy improvement is calculated.
2) **Friend Degree:** Usually, those who are popular in OSNs are considered to be the critical points in information diffusion. Therefore, we examine the privacy improvement by blocking the friend with largest degree.
3) **V-Index:** Vulnerability index was proposed by Gundecha et. al. [11], which is based on privacy setting of friends. We use this approach for unfriending, by blocking the friend with the largest V-Index.
4) **Random:** We also calculate the privacy improvement by randomly removing a friend link. This approach helps us to understand the average case when no friend sensitivity indicators are available.

The experiment setup is the same as that in Section VI-E. For each PIO-UD pair, we use above approaches to remove one friend link and calculate the privacy risk reductions. The experiment results are shown in Figure 11, in which the x-axis is the index of PIO-UD pair and y-axis is the privacy risk reduction. The statistic summary is shown in Table VII. We can see that TAPE gives the best performance. It is important to point out that the privacy risk reductions are calculated using the TAPE framework. It is not surprising that the Birnbaum's measure, which is based on TAPE, performs the best. On the other hand, we show that the other unfriending strategies, which consider less information, are not as promising as the proposed TAPE framework.

G. Discussion

We define a probability based definition for privacy risk (level). Starting from the quantitative definition, the reliability evaluation method is utilized to calculate one's privacy risk in terms of information diffusion. In TAPE, PA and PT are used to capture OSN users' privacy protection behaviors. TAPE can also compute the sensitivity of one's friend links, which can assist the user to adopt unfriending strategies. TAPE can be a starting point of enhance OSN users' privacy level. Since it highly depends on data sufficiency, the OSN service providers who control the most user data could be the best candidates to implement TAPE, and their users can really benefit from it. In addition, it is expected that social links (LISP) can also impact the calculation of privacy level. Real applications should adopt an LISP algorithm while being implemented.

TABLE VII: Statistic summary of unfriending strategies

Strategy	Average risk reduction	std.
TAPE	0.076	0.024
Friend degree	0.052	0.039
V-Index	0.046	0.041
Random	0.039	0.039

1) Privacy Leakage beyond one OSN: In TAPE, we assume that information can only be obtained through information diffusion within OSN. In practice, information diffusion is a much more complex process. There are several scenarios of information diffusion in social networks.

1) **Cross-OSN diffusion:** People can be active in multiple OSN platforms. For example, Alice is a friend of Bob on Twitter. She sees news about Bob on Twitter, and then she posts some words about the news on Facebook.
2) **Offline diffusion:** This is the traditional way we spread information through face-to-face conversation, phone calls etc.
3) **Online-Offline diffusion:** Information is propagated through both online and offline channels. This is the most common way we spread information in the information era.

Whereas scenario 2 is well studied in social science, scenarios 1 and 3 are challenging. In all the three scenarios, the concepts of privacy awareness and privacy trust are still valid. They have a great potential to be adopted in these scenarios and contribute to a broader study on personal privacy leakage in a hybrid online-offline world.

VII. Conclusion and Future Work

In this paper, we present a TAPE framework for the quantitative evaluation of users' privacy risk in OSNs. Mathematical tools (e.g. statistics, modeling techniques) are used to process online social network data, and signal processing tools are utilized in this work. The concepts of privacy awareness and privacy trust are introduced. Simulations are performed to illustrate the computation of privacy leakage probability, as well as to demonstrate that TAPE can capture useful information which was not captured previously. Several unfriending strategies are compared with the Birnbaum's method of TAPE, and TAPE gives the best performance. More importantly, TAPE sets up the stage for utilizing reliability analysis, which is a well-developed field, to solve privacy risk analysis problems. Besides BDD, other tools such as sensitivity analysis can surely benefit privacy research.

Future work includes developing better PA and PT algorithms, implementation of TAPE as a Facebook application and performing real user testing.

REFERENCES

[1] d. m. boyd and N. B. Ellison, "Social network sites: Definition, history, and scholarship," *Journal of Computer-Mediated Communication*, vol. 13, no. 1, pp. 210–230, 2007.

[2] R. Gross and A. Acquisti, "Information revelation and privacy in online social networks," in *Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*, Alexandria, VA, USA, 2005.

[3] B. Krishnamurthy and C. E. Wills, "On the leakage of personally identifiable information via online social networks," in *Proceedings of the 2nd ACM workshop on Online social networks*, Barcelona, Spain, 2009.

[4] S. Livingstone, "Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression," *New media & society*, vol. 10, no. 3, pp. 393–411, 2008.

[5] EurActiv. (2010) Social networks put careers at risk, survey finds. [Online]. Available: http://www.euractiv.com/Social-networks-careers-risk

[6] S. StrauB and M. Nentwich, "Social network sites, privacy and the blurring boundary between public and private spaces," *Science and Public Policy*, vol. 40, no. 6, pp. 724–732, 2013.

[7] S. E. Schechter, "Computer security strength and risk: A quantitative approach," Ph.D. dissertation, Harvard University, Cambridge, MA, USA, 2004.

[8] G. Stoneburner, A. Goguen, and A. Feringa, "Risk management guide for information technology systems," National Institute of Standards and Technology (NIST), Tech. Rep. 30, 2002.

[9] H. In, Y.-G. Kim, T. Lee, C.-J. Moon, Y. Jung, and I. Kim, "A security risk analysis model for information systems," in *Systems Modeling and Simulation: Theory and Applications*. Springer Berlin Heidelberg, 2005, vol. 3398, pp. 505–513.

[10] M. Hamdi and N. Boudriga, "Computer and network security risk management: theory, challenges, and countermeasures," *International Journal of Communication Systems*, vol. 18, no. 8, pp. 763–793, 2005.

[11] P. Gundecha, G. Barbier, and H. Liu, "Exploiting vulnerability to secure user privacy on a social networking site," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, San Diego, CA, USA, 2011.

[12] L. Fang and K. LeFevre, "Privacy wizards for social networking sites," in *Proceedings of the 19th international conference on World Wide Web*, Raleigh, NC, USA, 2010.

[13] R. Baden, A. Bender, N. Spring, B. Bhattacharjee, and D. Starin, "Persona: an online social network with user-defined privacy," in *Proceedings of the ACM SIGCOMM 2009 conference on Data communication*, Barcelona, Spain, 2009.

[14] A. Felt and D. Evans, "Privacy protection for social networking APIs," in *Web 2.0 Security and Privacy*, Oakland, CA, USA, 2008.

[15] A. Jeckmans, M. Beye, Z. Erkin, P. Hartel, R. Lagendijk, and Q. Tang, "Privacy in recommender systems," in *Social Media Retrieval. Computer Communications and Networks.* Springer Verlag, London, 2013, pp. 263–281.

[16] M. Cha, A. Mislove, and K. P. Gummadi, "A measurement-driven analysis of information propagation in the flickr social network," in *Proceedings of the 18th international conference on World wide web*, Madrid, Spain, 2009.

[17] S. Guo and K. Chen, "Mining privacy settings to find optimal privacy-utility tradeoffs for social network services," in *International Conference on Privacy, Security, Risk and Trust and International Confernece on Social Computing*, 2012.

[18] S. Guha, K. Tang, and P. Francis, "NOYB: privacy in online social networks," in *Proceedings of the first workshop on Online social networks*, Seattle, WA, USA, 2008.

[19] R. De Wolf, R. Heyman, and J. Pierson, "Privacy by design through a social requirements analysis of social network sites form a user perspective," in *European Data Protection: Coming of Age*. Springer Netherlands, 2013, pp. 241–265.

[20] Z. Erkin, T. Veugen, and R. Lagendijk, "Generating private recommendations in a social trust network," in *International Conference on Computational Aspects of Social Networks (CASoN '11)*, Salamanca, Spain, 2011.

[21] S. Alim, D. Neagu, and M. Ridley, "A vulnerability evaluation framework for online social network profiles: axioms and propositions," *International Journal of Internet Technology and Secured Transactions*, vol. 4, no. 2, pp. 178–206, 2012.

[22] R. Abdulrahman, S. Alim, D. Neagu, and M. Ridley, "Algorithms for data retrieval from online social network graphs," in *IEEE 10th International Conference on Computer and Information Technology*, Bradford, UK, 2010.

[23] R. Abdulrahman, S. Alim, D. Neagu, D. R. W. Holton, and M. Ridley, "Multi agent system approach for vulnerability analysis of online social network profiles over time," *International Journal of Knowledge and Web Intelligence*, vol. 3, no. 3, pp. 256–286, 2012.

[24] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins, "Information diffusion through blogspace," in *Proceedings of the 13th international conference on World Wide Web*, New York, NY, USA, 2004.

[25] E. Adar and L. Adamic, "Tracking information epidemics in blogspace," in *IEEE/WIC/ACM International Conference on Web Intelligence*, Compigne, France, 2005.

[26] F. Wang, H. Wang, and K. Xu, "Diffusive logistic model towards predicting information diffusion in online social networks," in *32nd International Conference on Distributed Computing Systems Workshops*, Macau, China, 2012.

[27] M. Z. Shafiq and A. X. Liu, "Modeling morphology of social network cascades," *Computing Research Repository*, vol. abs/1302.2376, 2013.

[28] E. Sadikov, M. Medina, J. Leskovec, and H. Garcia-Molina, "Correcting for missing data in information cascades," in *Proceedings of the fourth ACM international conference on Web search and data mining*, Hong Kong, China, 2011.

[29] J. Zhao, J. Wu, X. Feng, H. Xiong, and K. Xu, "Information propagation in online social networks: a tie-strength perspective," *Knowledge and Information Systems*, vol. 32, no. 3, pp. 589–608, 2012.

[30] Y.-G. Kim and J. Lim, "Quantitative risk analysis and evaluation in information systems: A case study," in *Computational Science ICCS 2007*. Springer Berlin Heidelberg, 2007, vol. 4489, pp. 1040–1047.

[31] X. Zang, H. Sun, and K. S. Trivedi, "A bdd-based algorithm for reliability graph analysis," *Technical report, Department of Electrical Engineering, Duke University*, 2000.

[32] C. Wang, L. Xing, V. Vokkarane, and Y. Sun, "Reliability analysis of wireless sensor networks using different network topology characteristics," in *International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, Chengdu, China, 2012.

[33] C. Wang, L. Xing, V. M. Vokkarane, and Y. Sun, "Manycast and anycast-based infrastructure communication reliability for wireless sensor networks," in *The 18th ISSAT International Conference on Reliability and Quality in Design*, Boston, MA, USA, 2012.

[34] Y. Sun, Z. Han, W. Yu, and K. Liu, "A trust evaluation framework in distributed networks: Vulnerability analysis and defense against attacks," in *25th IEEE International Conference on Computer Communications. Proceedings*, Barcelona, Spain, 2006.

[35] R. Xiang, J. Neville, and M. Rogati, "Modeling relationship strength in online social networks," in *Proceedings of the 19th international conference on World wide web*, Raleigh, NC, USA, 2010.

[36] J. Andrews and S. Beeson, "Birnbaum's measure of component importance for noncoherent systems," *IEEE Transactions on Reliability*, vol. 52, no. 2, pp. 213–219, 2003.

[37] M. Kurant, M. Gjoka, C. T. Butts, and A. Markopoulou, "Walking on a graph with a magnifying glass: Stratified sampling via weighted random walks," in *Proceedings of ACM SIGMETRICS*, San Jose, CA, 2011.

[38] C. Sibona and S. Walczak, "Unfriending on facebook: Friend request and online/offline behavior analysis," in *44th Hawaii International Conference on System Sciences*, Kauai, HI, USA, 2011.