Trust-Aware Privacy Evaluation in Online Social Networks

Yongbo Zeng[1], Yan (Lindsay) Sun[1], Liudong Xing[2], and Vinod Vokkarane[3]
[1]University of Rhode Island, Kingston, RI 02881, Email: {yongbozeng,yansun}@ele.uri.edu
[2]Univ. of Massachusetts Dartmouth, Dartmouth, MA, Email: lxing@umassd.edu
[3]Univ. of Massachusetts Lowell, Lowell, MA, Email: vinod_vokkarane@uml.edu

***Abstract*—While personal data privacy is threatened by online social networks, researchers are seeking for privacy protection tools and methods to assist online social network providers and users. In this paper, we aim to address this problem by investigating how to quantitatively evaluate the privacy risk, as a function of people's awareness of privacy risks as well as whether their friends can be trusted to protect their personal data. We present a trust-aware privacy evaluation framework, called TAPE. Simulations are performed to illustrate the key concepts and calculations in TAPE, as well as demonstrate the advantages of TAPE.**

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

I. Introduction

With the emergence of online social networks (OSNs), people are facing increasing privacy risks. Real life stories of sensitive information leakage in OSNs happen frequently [1], [2]. Most employers begin to collect potential employees' information using social networks, and information leakage through OSNs has put people's careers on risk [3].

The current research addresses privacy protection in OSNs from several angles: discussing privacy issues and protection recommendations [4], managing privacy setting [5], and adopting new architectures to build OSN [6]. On the other hand, the current methods for quantitatively evaluating the privacy level of individual users of OSNs are still not mature, and many current privacy protection approaches can greatly benefit from better privacy quantification approaches.

Quantitatively evaluating privacy level in OSNs is a challenging task. *First*, quantitative user privacy level is not a well defined concept in OSNs. *Second*, human users, whose behaviors are hard to quantify, play an important role in information diffusion. *Third*, personal information can be leaked through both online and offline media by many ways.

In this paper, we address the *first challenge* by proposing quantitative privacy definition based on privacy risk and probabilities. This quantitative measure will lead to the privacy level calculation tools, originally proposed in the reliability analysis field. To address the *second and third challenge*, we have to consider the availability of data. Since nobody can monitor all communication behaviors (online and offline) of users, researchers have to work on limited data, which can be obtained with reasonable cost. In this work, Facebook privacy setting is used as the primary data source. We also focus on the 'word-of-mouth' [7] on OSNs as the channel of information diffusion. Although other scenarios of information diffusion are not considered in this work, the proposed concepts and framework can be extended to the other scenarios.

We propose TAPE (Trust-Aware Privacy Evaluation) framework for quantitatively evaluating user privacy risk in OSNs. TAPE framework contains several novel aspects.

- It finds the similarity between the reliability analysis in wireless sensor networks (WSNs) and the privacy risk estimation in OSNs. It sets up the stage for utilizing reliability analysis tools for privacy evaluation.
- It considers the information diffusion through nodes and through links separately. Here, the information diffusion through nodes (i.e. users) mainly depends on the users' behavior, such as whether they respect other users' privacy. Such behavior is described by *Privacy Trust* (PT) and *Privacy Awareness* (PA), two new concepts introduced in this paper. The information diffusion through links (i.e. friend relationship) mainly depends on the closeness between friends in terms of whether one paying attention to the other's personal information.
- It proposes the desirable properties of PA and PT metrics, as well as specific ways to calculate PA and PT under the guidance of trust management theory. It is the first time that privacy trust concept has been used in evaluating privacy level in OSNs.

The proposed TAPE framework and algorithms are illustrated and tested based on real Facebook user data. The proposed PA algorithm is also compared with the known algorithm called IRT [8].

II. Related Work

Researchers studied privacy protection from two directions. Along the *first direction*, fundamental changes to social networking sites were suggested to enhance user privacy. For example, Baden et al. [6] proposed a new type of OSN using encryption to hide user data and allowing user to define privacy policies. The *second direction* is developing privacy tools based on existing OSNs. For example, the tradeoff between social network utility and personal privacy was studied in [8]. Fang et al. [5] developed privacy wizards to give user recommendation for privacy setting. Gundecha et al. [4] proposed an approach to identify a user's vulnerable friends. In this paper, we propose to assist user privacy protection by providing quantitative evaluation of privacy risk. Our work belongs to the second category.

There have been several quantification models for privacy evaluation in OSNs. Alim et al. [9] 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. A set of axioms for the vulnerability models were presented in [10]. They all provide meaningful and useful quantification approaches for OSN privacy evaluation. This work approaches the privacy quantification problem from a different angle. That is, to consider how likely a friend reveals others' personal information, described by the privacy trust concept, which is a widely studied research problem [11].

The proposed work is also related to information diffusion in OSNs [7]. For example, researchers attempt to build mathematical model to solve problems of information diffusion in OSNs, such as [12]. Different from the previous work, the proposed TAPE framework considers information diffusion in the context of privacy protection, which requires different sets of features and considerations.

III. Trust-aware Privacy Evaluation Framework

A. Social Network Privacy

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

Before we discuss quantification of privacy level, let us first look at two examples.

Example 1. *Alice is a student, and she wrote a piece of comment complaining her teacher Cris on her social network site. Alice does not want Cris to know the comment.*

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

In Example 1, the personal information concerned by Alice is her comment on Cris, and in Example 2, the personal information is her new boyfriend's photo. It is clear that a user has different personal information, and the privacy concerns for different personal information can vary. We introduce the notation I_j^u to denote user u's jth 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 often use I_j to represent I_j^{Alice}.

It is noted that privacy concerns are related to the "undesirable viewers". We define the concept of **Undesirable Group (UG)** of I_j^{Alice}, denoted by UG_j^{Alice}, as follows. If Alice does not want her information I_j to be seen by user u', then u' is put into UG_j^{Alice}, where u' 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 all users except her friends.

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

B. 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 level in OSNs? 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 the context of OSNs, the leakage of personal information I_j occurs when any UDs in UG_j obtain I_j. In TAPE, we assume that I_j can only be obtained through information diffusion in OSNs, 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 **information leakage probability** of I_j, denoted as L_j, as the probability that an UD obtains I_j.

In statistics, the notion of risk is often modeled as the expected value of an undesirable outcome. That is

$$\begin{aligned} Risk = &(probability\ \ of\ \ the\ \ accident\ \ occurring) \\ &\times (expected\ \ loss\ \ of\ \ the\ \ accident). \end{aligned} \tag{1}$$

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

$$R_j = L_j \cdot Z_j, \tag{2}$$

where L_j is information leakage probability defined earlier and Z_j describes the expected loss/damage of information leakage. In this work, Z_j is called as **information leakage hazard** and is normalized within interval $[0, 1)$. We argue that Z_j can be determined by the PIO (e.g. Alice) and/or existing research on the consequence of personal information leakage [13]. Therefore, the core task of TAPE is to estimate the information leakage probability L_j .

C. Toward Information Leakage Probability Estimation

In TAPE, a social network is represented by an undirected diagram. Users are the nodes, and friend connections are the links. As discussed earlier, personal information can be diffused to unintended recipients through the friendship links. It is important to point out that the existence of a link does not mean the personal information will be transmitted through this link. For example, Alice changes her status from "engaged" to "single". This information can be seen by all of Alice's friends. Here are three typical situations.

- Alice's friend Ned does not pay attention to Alice's status at all. Alice's status information does not disseminate to Ned although the link between Alice and Ned exists.
- Alice's friend Ned pays attention to Alice's status. Alice's status information disseminates to Ned through the link between Alice and Ned. Then, Ned respects Alice's privacy and does not tell others about Alice's status change. In this case, Alice's status information does not disseminate to others through Ned.
- Ned sees Alice's status change, and adds a post, in which he guesses that Alice broke with her boyfriend

according to her status change. Such post can be seen by all Ned's friends. In this case, Alice's status information disseminates to Ned, and then to others through Ned.

We argue that solving the problem of information leakage probability estimation can be decomposed into two tasks.

1) The first task is to estimate the probability whether a user's private information will be disseminated through a particular link or a particular node. In this work, such probabilities are referred to as **information spreading probabilities** (ISP). In the above example, the ISP of the link between Alice and Ned reflects whether Ned will pay attention to Alice's information. The ISP of node Ned reflects whether Ned respects Alice's privacy.
2) The second task is to compute the information leakage probability (i.e. L_j), given the network topology, the ISP of nodes/links, the PIO (i.e. Alice), and the undesirable group (i.e. UG_j).

In the rest of this section, we first discuss the solution to the second task (Section III-D), and then present the metrics for solving the first task (Section III-E and Section III-F).

D. Privacy Analysis and Reliability Analysis

When investigating information diffusion in OSNs, we found that one of reliability estimation methods called reliability graph [14] aims to solve the similar problem.

In a reliability analysis problem, the system is represented by a **reliability graph**, whose edges and nodes are components of the system and are assigned certain **failure probabilities**. The system has **source** nodes and **sink** nodes. If there is no path from the source(s) to the sink(s), the system fails. In the context of WSN reliability analysis, one often needs to estimate the probability that there is at least one communication path between source node(s) and destination node(s) [15].

In the TAPE framework, we have defined the **information spreading probability (ISP)** for nodes and links. This concepts is kind of "opposite" to the failure probability. For example, if a node fails to forward data to the destination with probability p, this node's failure probability is p in the context of WSN reliability analysis, whereas this node's information spreading probability is $1-p$ 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 data 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 [14], [15] to solve Task 2 described in Section III-B.

In order to utilize the BDD method for social network that is usually of large size, we revise BDD as follows. When generating the BDD graph, the maximum traversing depth is set by a factor κ, according to the number of hops between PIO and UD. For example, if $\kappa = 4$ and the UD is 3 hops away from PIO, then the BDD branches that are longer than 12 (3×4) are pruned. This revised BDD is refered to as **reduced BDD algorithm**. Table I shows the important concepts in TAPE, as well as the concept mapping.

TABLE I: Concepts mapping

Reliability Analysis in WSN	Privacy Analysis in TAPE
Communication reliability	Information leakage prob.
Reliability graph	Social graph
Source node	PIO
Destination	UD
Node/edge failure probability	1 - node/link ISP

E. Calculation of Node Information Spreading Probability

While most social network information diffusion models consider the impact of nodes and links together [16], we argue that information propagation through nodes and through link should be considered separately. This is why we define node ISP and link ISP separately, which can better describe the information propagation process.

In this subsection, we discuss the **Information spreading probability of node**, which is the probability that a node will spread others' information. We use ISP_u to represent the ISP of node u. One person is assumed to have a consistent ISP within a certain period of time.

Evaluating information spreading probability of a person is very challenging because it is related to one's knowledge and personality. In the real (i.e. offline) world, we probably can estimate the ISP of a person based on experiences if we know this person for a long time. Obviously, such estimation can be biased and limited, and cannot be applied in OSNs. Instead of resolving a challenging problem in social science, we propose to examine ISP of a person based on the quantitative information available in OSNs.

In particular, we propose two metrics that should be used to estimate node ISP.

*1) **Privacy Awareness**:* The first one is **privacy awareness (PA)**, which purely 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*. In the TAPE framework, PA evaluation is a module whose inputs are privacy settings of the given OSN user. In fact, implementations based on TAPE can use any reasonable PA algorithms. We propose a PA algorithm in Section IV.

*2) **Privacy Trust**:* We propose **Privacy Trust (PT)** to represent how much a person should be trusted in terms of protecting others' privacy. Trust can be established through recommendations [11]. For example, in Fig 2a, node B trusts node C, and node B gives a recommendation to node A saying that he/she trusts node C, node A can develop certain level of trust in node C. In TAPE framework, we propose to evaluate PT based on implicit recommendation through friendship links. Similar to PA, PT evaluation is a module of TAPE, and the inputs are the PA values of the friends of the given user. The details will be presented in Section V.

PA_u and PT_u are used to represent the PA and PT of node u respectively. Both PA and PT affect the ISP of node u. In this paper, we compute the ISP of node u as:

$$ISP_u = w \cdot PA_u + (1 - w)PT_u, \tag{3}$$

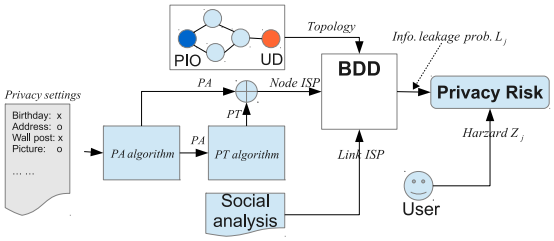


Fig. 1: Core structure of the TAPE framework

TABLE II: 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%

where w is the weight between 0 and 1. In the experiments in Section VI, we choose $w = 0.5$.

F. Link Information Spreading Probability

As discussed in Section III-E, the ISP of the Alice-Bob link depends on whether Bob heard what Alice said. It depends on whether Alice and Bob have a strong tie between them in OSNs. In the current literature, many works have investigated this problem [17]. Note that the TAPE framework can accommodate any algorithms for link ISP calculation. In this paper, we do not propose a specific algorithm for calculating link ISP. In the experiments, we adopt a constant value for link ISP and focus on the demonstrating the impact of PA and PT.

As a summary, the structure and key components of TAPE are illustrated in Fig 1.

IV. Privacy Awareness Algorithm

A. The proposed algorithm: Rank PA

We define $rank^+_{u,j}$ as the proportion of users whose privacy setting for information I_j is looser than user u. As long as we know the statistics of users' privacy setting for information I_j, we can compute $rank^+_{u,j}$. For example Table II shows the statistics of birthday privacy setting of 10000 OSN users. We assume the order of privacy settings from loose to tight is {'everyone', 'networks', 'friends of friends', 'friends', 'self'}. If Alice allows only her friends to see her birthday, i.e. $s_{Alice,j} = friends$, then $rank^+_{Alice,j} = 0.05+0.4+0.1 = 0.55$. Similarly, we define $rank^-_{u,j}$ as the proportion of users whose privacy setting for information I_j is tighter than user u. In the above example, $rank^-_{Alice,j} = 0.1$.

Next, we compute the **individual information privacy awareness (IPA)** for each privacy setting. Let $IPA_{u,j}$ denote IPA of user u for information I_j, we propose

$$IPA_{u,j} = \frac{1}{2}(rank^+_{u,j} - rank^-_{u,j}) + \frac{1}{2}. \quad (4)$$

$IPA_{u,j}$ is normalized into $[0, 1]$, since in the context of TAPE, PA is evaluated in a probability meaning, and moreover, we consider $PA = 0.5$ as neutral, $PA < 0.5$ as unawareness, and $PA > 0.5$ as awareness. In the above example, $IPA_{Alice,1} = 0.725$.

We then calculate the IPA for all types of information: $I_1, I_2, I_3, \cdots, I_J$, and compute the overall PA of user u as

$$PA_u = \frac{1}{J}\sum_{j=1}^{J} IPA_{u,j}. \quad (5)$$

B. PA Algorithm Criteria

The TAPE framework can accommodate many PA algorithms. However, what are the design criteria for PA algorithms? We identified seven special cases and the desirable PA values in such special cases in Table III, which serves as a guidance for PA algorithm design.

To see the insight of case 4, 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 of doing so. In this case, if Alice releases her birthday, her PA should not be largely reduced. On the other hand, if Alice releases a particular type of information which most people choose not to release, Alice's PA should be reduced more. This is the reason why we consider special cases 4, 5, 6 and 7. It is easy to verify that the proposed PA algorithm does have the desirable features listed in Table III.

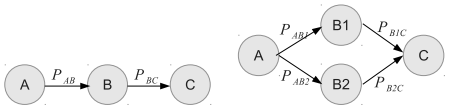
In Section VI-B, case study of Rank PA algorithm is presented, and it is compared with an item response theory (IRT) base privacy concern model proposed in [8].

V. Privacy Trust Algorithm

In TAPE, we propose to evaluate privacy trust (PT) through recommendations. If a user with high PA releases his/her information to Alice, this user implicitly tells us he/she trusts Alice not to propagate his/her personal information. This can be an implicit recommendation. In real life, if a student working on privacy research (e.g. myself) chooses to tell someone my birthday and address, it means that I trust this person not releasing my personal information to others. Although such implicit recommendations have noises and can be biased, it may be the best resource to compute PT in OSNs.

TABLE III: Desirable properties of PA calculation

Special Cases	**Desirable PA value**
1: Alice's privacy settings are looser than all others'	$PA_{Alice,j} = 0$
2: Alice's privacy settings are tighter than all others'	$PA_{Alice,j} = 1$
3: Everyone has the same privacy settings	$PA_{Alice,j} = 0.5$
4: Many users (including Alice) have loose settings, and a few users have tight settings	$PA_{Alice,j}$ should be small, but not too small because most people share the same opinion as Alice.
5: A few users (including Alice) have loose settings, and many users have tight settings	$PA_{Alice,j}$ should be smaller than $PA_{Alice,j}$ in case 4.
6: Many users (including Alice) have tight settings, and a few users have loose settings	$PA_{Alice,j}$ should be high, but not too high because most people share the same opinion as Alice.
7: A few users (including Alice) have tight settings, and many users have loose settings	$PA_{Alice,j}$ should be much higher than $PA_{Alice,j}$ in case 6.



(a) One-path trust (b) Two-path trust

Fig. 2: Trust and privacy trust propagation

Similar as in the PA calculation, we argue that the PT calculation should have two desirable properties.

First, the level of privacy trust of user u, denoted by PT_u, largely depends on the number of recommendations from the friends with high PA. If a user with low PA trusts Alice (i.e. allowing Alice to view his/her personal information), this should not affect Alice's PT either positively or negatively. We use F_u^+ to denote the number of *high quality recommendations*, i.e. the number of u's friends who have PA higher than a threshold (ϵ^+) and allow u to view a substantial amount of personal information. Note that the calculation of PT may consider other factors beyond F_u^+, such as the specific PA value of each friend. In this paper, we use a simple PT calculation, in which only the F_u^+ value is considered. That is, PT_u is a function of F_u^+, i.e. $PT_u = f(F_u^+)$.

Second, although each additional recommendation can increase PT_u, such increase diminishes when F_u^+ is very large. For example, when F_u^+ increases from 3 to 6, PT_u can increase a lot. However, when F_u^+ increases from 300 to 303, PT_u should not increase much. Specifically, when $X < Y$, we should have $f(X + \Delta) - f(X) > f(Y + \Delta) - f(Y)$.

In [11], Sun et al proposed a probability trust model that uses Beta function to address concatenation propagation and multipath propagation of trust.

Fig 2a shows an example of one path trust propagation, in which B trusts C with a trust measurement P_{BC} (direct trust value) and A has an judgement P_{AB} (recommendation accuracy) when B recommends her/his trust of C to A. The trust between A and C can also be established through multiple paths, as shown in Fig 2b, in which both $B1$ and $B2$ give recommendation of C.

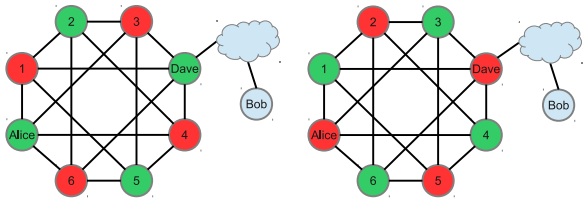
In the context of privacy trust, the recommendation accuracy P_{AB} is replaced by PA of A, which is PA_A, and the trust value P_{BC} is the implicit trust B towards C. This implicit trust, represented by $T_{B,C}$ in TAPE, is established when B allows C to view a substantial amount of personal information. For simplicity, in the current work, we set $T_{B,C}$ as a constant.

VI. Experiment Results and Discussion

We implement TAPE framework in Matlab. In PT algorithm, PA threshold (ε^+) is set to be 0.5 and $T_{B,C}$ is 0.7. Besides PA and PT, the link ISP is set to be 0.9 for all links. In this section, we first do a case study which demonstrates the features of TAPE framework and then apply TAPE to two real user datasets.

A. Case study

As shown in Fig 3, Alice, Dave and node 1 to 6 form a group. Each node has 4 social network friends within the group. Suppose Alice is considering her privacy, and she



(a) (b)

Fig. 3: Case study (PA of red nodes is 0.2, and PA of green nodes is 0.7)

defines the UD to be {Bob}, who is outside the group. We assume the PA distribution within the group is fixed. There are 4 nodes with PA=0.2 and 4 nodes with PA=0.7. In Fig 3a, all of Alice's friends have low PA values, which means their privacy settings are quite loose. In Fig 3b, all of Alice's friends have high PA values, which means their privacy settings are quite tight. After applying TAPE framework, we obtain that the information leakage probability for Alice is 0.08 in Fig 3a and is 0.12 in Fig 3b.

We make the following observations. First, only considering the Alice' friends privacy setting is not sufficient. In Fig 3a, all Alice's friends have loose privacy setting, but Alice has higher privacy level (lower information leakage probability). In Fig 3b, Alice's friends have tight privacy setting, but Alice has lower privacy level. Second, considering specific information leakage path is important. We found the critical node is Dave, whose information spreading probability (ISP) largely impact Alice's privacy level. The ISP of Dave is 0.36 and 0.53 in Fig 3a and Fig 3b, respectively.

Through this example, we can see that the privacy level of Alice depends on the choice of UD, network topology, and the ISP of nodes along the paths from Alice to the UD. All these factors have been included in the TAPE framework.

B. Comparisons between PA algorithms

In this subsection, we compare two PA algorithms by applying them in special cases. We assume there are 4 binary privacy settings for each OSN user, which is the column index of Table IV. For example, '0000' means 4 privacy settings are all set to be 'hidden'. We investigate the PA computing results in following scenarios, and compare it with the IRT model [8].

1) Most users (85.8%) have very tight privacy settings (i.e. 0000).
2) Most users (85.5%) have have tight privacy settings for some information, and loose privacy setting for the other information (e.g. 0101).
3) Most users (85.4%) have very loose privacy settings (i.e. 1111)
4) The percentage of users with different privacy settings are uniformly distributed.

The results are shown in Table IV. For example, in case 1, 85.8% of users choose setting '0000'. For those users, IRT model in [8] computes IRT PA=0, and TAPE computes Rank PA=0.54. We observe the two models have similarities.

1) Generally, the more privacy settings are set to be open, the lower the corresponding PA value is.

TABLE IV: PA calculation examples

Special cases		Privacy settings: $s_1 s_2 s_3 s_4$ (0=hidden,1=open)															
		0000	0001	0010	0100	1000	0011	0101	0110	1001	1010	1100	0111	1011	1101	1110	1111
1	Stats	**85.8%**	0.9%	0.9%	0.9%	0.9%	0.9%	1.0%	0.9%	1.0%	1.1%	0.9%	1.0%	0.8%	0.9%	1.0%	0.9%
	IRT PA	**0.00**	-1.28	-1.28	-1.28	-1.28	-1.38	-1.38	-1.38	-1.38	-1.38	-1.38	-4.57	-4.57	-4.58	-4.56	-1.67
	Rank PA	**0.54**	0.41	0.41	0.41	0.41	0.29	0.29	0.29	0.29	0.29	0.29	0.16	0.16	0.16	0.16	0.04
2	Stats	1.1%	0.9%	1.1%	0.8%	1.0%	0.8%	**85.5%**	1.2%	1.0%	1.1%	0.9%	0.9%	0.9%	1.1%	0.9%	0.9%
	IRT PA	1.88	1.06	1.06	1.06	1.06	0.00	**0.00**	0.00	0.00	0.00	0.00	-1.06	-1.06	-1.06	-1.06	-1.87
	Rank PA	0.75	0.62	0.62	0.62	0.62	0.50	**0.50**	0.50	0.50	0.50	0.50	0.37	0.37	0.37	0.37	0.25
3	Stats	0.9%	0.9%	1.1%	0.9%	0.9%	0.9%	1.0%	1.0%	1.1%	1.0%	1.1%	1.0%	0.8%	1.2%	0.9%	**85.4%**
	IRT PA	1.67	1.47	1.47	1.47	1.47	1.38	1.38	1.38	1.38	1.38	1.38	1.28	1.28	1.28	1.28	**0.00**
	Rank PA	0.96	0.84	0.84	0.84	0.84	0.71	0.71	0.71	0.71	0.71	0.71	0.59	0.59	0.59	0.59	**0.46**
4	Stats	6.3%	6.4%	6.5%	5.9%	6.6%	6.2%	6.1%	6.2%	6.5%	6.3%	6.4%	6.0%	6.1%	6.4%	6.1%	6.1%
	IRT PA	0.24	0.12	0.12	0.12	0.12	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.12	-0.12	-0.12	-0.12	-0.24
	Rank PA	0.75	0.62	0.62	0.62	0.62	0.50	0.50	0.50	0.50	0.50	0.50	0.37	0.37	0.37	0.37	0.25

2) The majority behavior always gets a PA value around 0.5 for Rank PA (or 0 for IRT PA), which is in between awareness and unawareness.
3) When most people adopt tight privacy settings (case 1), releasing information implies high privacy risk. Therefore the PAs of all other privacy settings in this case are less than 0.5 for Rank PA (or less than 0 for IRT PA), which means unwareness. And vice versa (case 3).
4) When the percentage of users with different privacy settings are uniformly distributed (case 4), the Rank PA values distributes from 0.25 to 0.75, which means there's neither extremely high PA nor extremely low PA.

We claim that the Rank PA is better than IRT PA. Rank PA value is normalized to (0,1), which has an intuitive meaning and is comparable. It can also be easily extended to a probability method. IRT PA value has a range of $(-\infty, +\infty)$. We need further normalization to make IRT PA meaningful and fit probability based methods. For example, in case 2, the Rank PA is distributed from 0.25 to 0.75, and IRT PA is distributed from -1.87 to 1.88. In case 4, the Rank PA is distributed from 0.25 to 0.75, but the IRT PA is distributed from -0.24 to 0.24. In such situation, we cannot directly compare IRT PAs between different cases.

C. Datasets

Two datasets were constructed to perform larger scale experiments. **Dateset I** contains 514 Facebook users, including one graduate student at URI, his friends, and his friends of friends. Note that the privacy setting data is usually not provided by OSN providers. In this dataset, very detailed privacy settings were obtained through a survey. **Dataset II** contains 957 thousand Facebook users, sampled by Metropolis-Hasting random walk (MHRW) by the authors of [18]. Four privacy settings were crawled for each user, including 'add as friend', 'photo', 'view friends' and 'send message'. For each, the privacy setting is binary, either open or hidden, from the crawler points of view. Some features of the two datasets are listed in Table V.

D. Privacy Leakage Probability

It is well known that the reliability of data transmission can drop significantly as the distance (i.e. the number of hops) between the source node and the destination node increases. In the context of privacy protection, does the information leakage probability heavily depend on the distance between the PIO and the UD?

TABLE V: Datasets summary

	Dataset I	**Dataset II**
# of unique users	514	957K
Avg real degree	215.8	95.2
Avg sampled degree	2.1	3.8
Max sampled degree	18	124
# of privacy settings	16	4
Privacy setting type	5 levels	binary

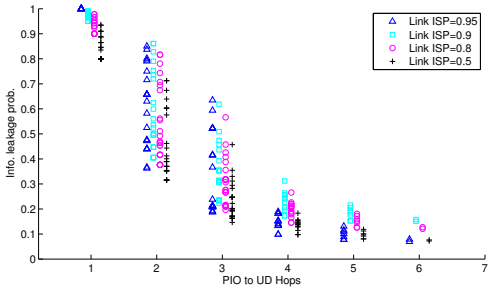


Fig. 4: Information leakage probabilities vs. PIO to UD hops.

In this subsection, we study the relationship between the privacy leakage probability (the y-axis in Fig 4) and the hop distance from the PIO to the UD (the x-axis in Fig 4).

We first randomly pick 100 nodes and put them in set S. In each round of simulation, we pick one node from S as PIO, and pick one other node from the k-hop neighbors ($k = 1, 2, 3, 4, 5, 6$) of the PIO as UD. For each pair of PIO and UD, we measure the distance between them (hops), compute the information leakage probability using TAPE, and plot Fig 4. Each point represents one result. The x-axis indicates the hop-count distance between PIO and UD, and y-axis is the ISP. The color indicates link ISP, which is chosen as 0.5, 0.8, 0.9, and 0.95 respectively. We have several observations

- The information leakage probability to 1-hop neighbors (i.e. friends) can be greater than the link ISP. This is because Alice's friend not only gets the information from Alice directly, but also through other paths. For example, Alice's friend Bob may not heard what Alice said, but he could get the message from Charlie who is a friend of Alice.
- As expected, when the hop distance increases, the information leakage probability has a decreasing trend.
- When the distance is small, the information leakage probability varies in a large range. The hop count is not a dominating factor. The PA, PT and network topology

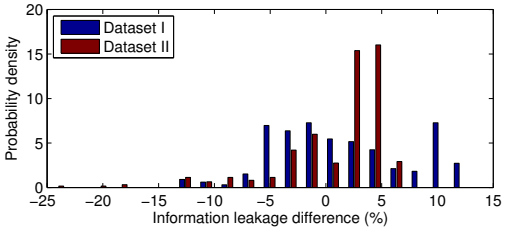


Fig. 5: Histogram of information leakage prob. changes

jointly determine users' privacy level. A user who is 3 hops away can be more likely to obtain Alice's person information than a user who is 2 hops away.

- As the link ISP decreases, information leakage probability decreases. In the future work, incorporating the estimation of link ISP will yield even a larger variation in the ISP values.

E. The impact of PA and PT

Since the lack of "ground truth" about the real privacy level 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 OSNs only focuses on network topology. We construct a comparison method, referred to as the *topology-based method*, which uses the BDD to compute the information leakage probability with the fixed link ISP and fixed node ISP. 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 of the topology based method, we set link ISP to be 0.7, and set node ISP to be the average of node ISP in the proposed approach.

The experiment setup is similar to that in Section VI-D. We compute $(L_{Topology-based} - L_{TAPE})/L_{TAPE}$, which is the percentage of information leakage probability change without considering PA and PT. in Fig 5, we show the histogram of such percentage of change, for results using two datasets. It is seen that the change range is from -25% to 15%. Considering PA and PT does provide additional and useful information beyond the topology. In addition, it is seen that dataset II shows more concentrated distribution around 0. It is known that, dataset II has 4 binary privacy settings, while dataset I has 16 5-level privacy settings. We argue that the comprehensiveness of privacy setting used in TAPE can impact the performance of TAPE.

VII. Conclusion and Future Work

In this paper, we presented a new Trust-Aware Privacy Evaluation framework, as shown in Fig 1, for quantitative evaluation of users' privacy risk in OSNs. The concepts of privacy awareness and privacy trust were introduced. Simulations were performed to illustrate the computation of information leakage probability, as well as to demonstrate that TAPE captures useful information that was not captured by the topology-based methods. More importantly, TAPE sets up the stage for utilizing reliability analysis, which is a well-developed field, to solve privacy risk analysis problems. Future work includes developing better PA and PT algorithms and performing sensitivity analysis to discover how to improve user privacy in OSNs. Additionally, framework evaluation and comparisons will be performed in the future if real OSN users can be involved. This work is partially supported by NSF award #1112935 and #1112947.

References

[1] *http://news.sky.com/story/753966/mod-secrets-leaked-onto-the-internet*.
[2] *http://news.bbc.co.uk/2/hi/middle_east/8549099.stm*.
[3] *http://www.euractiv.com/Social-networks-careers-risk*.
[4] 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*, ser. KDD '11. New York, NY, USA: ACM, 2011, pp. 511–519. [Online]. Available: http://doi.acm.org/10.1145/2020408.2020489
[5] L. Fang and K. LeFevre, "Privacy wizards for social networking sites," in *Proceedings of the 19th international conference on World wide web*, ser. WWW '10. New York, NY, USA: ACM, 2010, pp. 351–360. [Online]. Available: http://doi.acm.org/10.1145/1772690.1772727
[6] 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*, ser. SIGCOMM '09. New York, NY, USA: ACM, 2009, pp. 135–146. [Online]. Available: http://doi.acm.org/10.1145/1592568.1592585
[7] 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*, ser. WWW '09. New York, NY, USA: ACM, 2009, pp. 721–730. [Online]. Available: http://doi.acm.org/10.1145/1526709.1526806
[8] S. Guo and K. Chen, "Mining privacy settings to find optimal privacy-utility tradeoffs for social network services," in *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*, 2012, pp. 656–665.
[9] S. Alim, D. Neagu, and M. Ridley, "A vulnerability evaluation framework for online social network profiles: axioms and propositions," *Int. J. Internet Technol. Secur. Syst.*, vol. 4, no. 2/3, pp. 198–206, Jul. 2012. [Online]. Available: http://dx.doi.org/10.1504/IJITST.2012.047961
[10] ——, "Axioms for vulnerability measurement of online social network profiles," in *Information Society (i-Society), 2011 International Conference on*, 2011, pp. 241–247.
[11] Y. Sun, Z. Han, W. Yu, and K. Liu, "A trust evaluation framework in distributed networks: Vulnerability analysis and defense against attacks," in *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, 2006, pp. 1–13.
[12] M. Z. Shafiq and A. X. Liu, "Modeling morphology of social network cascades," *arXiv preprint arXiv:1302.2376*, 2013.
[13] 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*, ser. WOSN '09. New York, NY, USA: ACM, 2009, pp. 7–12. [Online]. Available: http://doi.acm.org/10.1145/1592665.1592668
[14] X. Zang, H. Sun, and K. S. Trivedi, "A bdd-based algorithm for reliability graph analysis," *Department of Electrical Engineering, Duke University, Tech. Rep*, 2000.
[15] 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*, 2012.
[16] E. Adar and L. Adamic, "Tracking information epidemics in blogspace," in *Web Intelligence, 2005. Proceedings. The 2005 IEEE/WIC/ACM International Conference on*, 2005, pp. 207–214.
[17] 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. [Online]. Available: http://dx.doi.org/10.1007/s10115-011-0445-x
[18] 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 '11*, San Jose, CA, June 2011.