Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2019) 27: 3417 – 3432

doi:10.3906/elk-1811-25

Privacy issues in post dissemination on Facebook

Burcu SAYİN[1], **Serap ŞAHİN**[1*], **Dimitrios G. KOGIAS**[2], **Charalampos Z. PATRIKAKIS**[2]
[1]Department of Computer Engineering, Faculty of Engineering, İzmir Institute of Technology, İzmir, Turkey
[2]Department of Electrical & Electronics Engineering, Faculty of Engineering, University of West Attica, Athens, Greece

Received: 03.11.2018 • **Accepted/Published Online:** 15.04.2019 • **Final Version:** 18.09.2019

Abstract: With social networks (SNs) being populated by a still increasing numbers of people who take advantage of the communication and collaboration capabilities that they offer, the probability of the exposure of people's personal moments to a wider than expected audience is also increasing. By studying the functionalities and characteristics that modern SNs offer, along with the people's habits and common behaviors in them, it is easy to understand that several privacy risks may exist, many of which people may be unaware of. In this paper, we focus on users' interactions with posts in a social network (SN), using Facebook as our research domain, and we emphasize some privacy leakages currently existing in Facebook's privacy policy. We also propose a solution to detected privacy issues, featuring a reference implementation of a tool based on a simulation, which visualizes the effect of potential privacy risks on Facebook and directs users to control their privacy. The proposed and simulated tool allows a post owner to observe the spreading area of his or her post depending on the selected privacy settings. Moreover, it provides preliminary feedback for all Facebook users that have interacted with this post, to make them aware of the possible privacy changes, aiming to give them a chance to protect the privacy of their interaction on this post by deleting it when an unwanted privacy change takes place. Finally, an online survey to increase privacy awareness in Facebook usage with over 500 volunteer participants has illuminated the need for such a tool or solution.

Key words: Social networks, social network analysis, privacy, Facebook

1. Introduction

In recent years, social networks (SNs) such as Facebook, Instagram, Twitter, and LinkedIn have become increasingly popular among people. According to the findings of Statista[1] in April 2018, Facebook is the most popular SN today with approximately 2.2 billion monthly active users. WhatsApp and Facebook Messenger follow it. Table 1 shows the number of monthly active users for each one in 2016, 2017, and 2018. Results show that there is a significant number of SN users and this number keeps increasing. This could be mainly attributed to the fact that SNs, by taking advantage on the advances in computing and telecommunications, have become a common environment in which people communicate and keep track of other people's lives at any moment. The majority of users accept their SN accounts as part of their lives, using them to share their thoughts and publish their captured moments. However, although SNs provide the means for people to stay in

*Correspondence: serapsahin@iyte.edu.tr

[1]Statista GmbH (2018). Most Famous Social Network Sites Worldwide as of April 2018, Ranked by Number of Active Users (in Millions) [online]. Website https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/[accessed 19-10-2018].

touch with their friends or acquaintances, they introduce several privacy issues for users, many of which they might be completely unaware of or misinformed about.

Table 1. Monthly active users of most popular SNs (in billions).

	2016	2017	2018
Facebook	1.6	1.97	2.2
WhatsApp	1	1.2	1.5
Facebook Messenger	0.9	1	1.3

SNs have a user-friendly interface so that anyone can create an account and join the community easily. After getting an account, a user can connect with his or her friends and share any information (posts) with them. However, if users do not adjust the privacy of their posts properly, those can be visible beyond their expectations, even outside of their social circle. This situation creates privacy risks for SN users. Hence, our objective in this paper is to highlight specific and already existing privacy risks that we detected during our analysis of Facebook and propose a solution/application to inform Facebook users about their privacy level and allow them to better control their actions in the case of any privacy change. We have selected Facebook[2] as the focus of our study because it is the most widely used SN platform today.

Facebook was launched in 2004 as a student network for Harvard University by Mark Zuckerberg and his friends. Until 2005, only students from Harvard University could join Facebook. Then it started to increase its availability through some high schools and companies [1]. In 2006, Facebook became publicly available but the privacy concerns did not change; it was still network-based. Hence, personal data were accessible throughout the whole network (i.e. public) by default.

Facebook served its users a privacy-concerning message in 2009 for the first time, by asking them to select a personal privacy setting for their personal data. However, the served default option was selected to be "Everyone", so many users accepted the default setting without being aware of the risks and without knowing where they consented. This allowed many data to become publicly available.

As a result of many debates and criticisms about possible privacy concerns, Facebook unveiled a new privacy setting page in May 2010. There were different levels of privacy setting options on the page including "Everyone, Friends of Friends (FoF), Friends Only, Other" for each data category.[3] However, it was not sufficient to prevent privacy leakages. In 2011, new changes were applied to Facebook privacy and those changes caused new problems as people could reach some users' personal data and profiles without being friends [2]. Hence, almost all data gradually became publicly available through a combination of default settings. After all that, Facebook applied a post-based privacy option that allows users to set a specific privacy preference for each post.[4] Hence, the old network-based privacy structure was removed.

In this scope, Facebook proposed four options for privacy settings: (i) Friends: only the post owner (PO)'s friends can see it; (ii) FoF: the PO's friends and friends of the PO's friends can see it; (iii) Everyone: anyone can see it, even someone who is not a Facebook member; and (iv) Custom: the PO can create a custom

[2]Official Facebook Website (2018). [online]. Website https://www.facebook.com/ [accessed 19-10-2018].

[3]McKeon M (2018). The Evolution of Privacy on Facebook [online]. Website http://mattmckeon.com/facebook-privacy/ [accessed 19-10-2018].

[4]Edens J (2018). Facebook Privacy Policy Has Become Less Transparent, Harder to Understand and Control, Experts Say [online]. Website http://www.stuff.co.nz/technology/social-networking/89645571/Facebook-privacy-policy-has-becomeless-transparent-harder-to-understand-and-control-experts-say [accessed 19-10-2018].

setting by selecting some specific friends or lists and excluding some others. Although Facebook has improved privacy control options, it was claimed that many users' information is still accessible by third parties. People were complaining about both privacy problems and the confusing structure of the novel privacy settings.

Shore et al. [3] worked on every privacy policy of Facebook from 2005 to 2015 and ranked them. Their results showed that Facebook privacy policies became increasingly incomprehensible and confusing to understand. They stated that usability, clearness, and transparency of privacy policies on Facebook are not improving and so users have fewer options to control their personal information against third parties. This study confirmed that users need some tools or applications that offer a user-friendly interface and easy privacy setting/tracking.

In this study, we focus on current privacy leakages that we detected on Facebook, and we propose a tool to increase privacy awareness. In order to collect some real data for better evaluation of our study and the proposed tool, we conducted a questionnaire via the Survey Monkey website. [5] We prepared an elaborate question list, focusing on Facebook usage and privacy awareness. More than 500 volunteers responded to our questions.[6] Five hundred of the respondents stated that they are Facebook users, with 58% within the age range of 15–24 and 33% between 25 and 44 years old. Additionally, 54.5% had a university degree, 25% are at graduate level, and 20% had a high school degree. According to the ages and educational levels of respondents, and the countries of respondents (13% Turkey, 82% Greece), the survey looks like it does not demographically represent the structure of all Facebook users. However, 46% of the respondents were not aware of Facebook's privacy policy,[7] which is an interestingly high percentage of mostly well-educated people. This result shows the need for a tool that can increase the privacy awareness of Facebook users. In the following two subsections, we explain the detected privacy issues in detail, and then we present our tool.

1.1. Currently detected privacy issues on Facebook

After our analysis, we detected two important privacy leakages within the Facebook privacy policy:

- *Comment owner's (CO's) interactions (i.e. like/comment) may be visible beyond the expected area:* When the user interacts with a post that is visible to "Friends only" and if the PO changes the privacy setting of the post beyond Friends, Facebook does not inform COs of this change. Hence, COs' interactions are visible to more people than they initially expected. According to our survey, 49% respondents claim that they disagree with this privacy policy of Facebook, while 34% of them are neutral.

- *CO's interactions may be inaccessible to their owners:* Suppose that user A and user B are friends and user C is only B's friend. Then assume that user A posts something with privacy of "Friends only" while he/she also tags user B in the post. Hence, the post privacy becomes "Friends and B's Friends". Then suppose that user C interacts with this post. After a while, user A removes user B's tag. In this case, not only is user C no longer able to reach his/her comments/likes anymore, but also he/she cannot even see it in their activity log. This situation creates a privacy issue for user C, because he/she cannot see his own comments/likes while other people can still see them. According to our survey, 63% of volunteers claim that they disagree with this privacy policy of Facebook, while 24% of them are neutral.

[5]Survey Monkey (2019). Website https://www.surveymonkey.com [accessed 15-03-2019].
[6]Results of questionnaires on Github (2019). Survey results for privacy awareness on Facebook usage [online]. Website https://github.com/burcusyn/PrivacyAwarenessOnFacebook [accessed 15-03-2019].
[7]Facebook Privacy Policy (2019). Website https://www.facebook.com/about/basics/manage-your-privacy [accessed 15-03-2019].

1.2. Solution proposal for detected privacy issues on Facebook

Researchers have suggested various solutions and developed applications to protect SN users' privacy [4–18], especially on Facebook. Some of them were even adapted by the Facebook platform. However, as we stated above, privacy issues still exist because Facebook is a dynamic platform; it grows continuously with each new friend connection, new post/comment/share, etc. Furthermore, Facebook continuously proposes new updates and Facebook users may change the privacy settings of their profile/posts, etc., so new privacy leakages may arise.

Considering these points, we propose the following contributions to handle detected privacy leakages:

- Analyzing and measuring the impact of these leakages on post dissemination.
- Increasing the privacy awareness of Facebook users so that they can protect their interactions on other people's posts during the post dissemination process.
- Visualizing the problem and proposed solution with the design of a Facebook privacy tool.
- Analyzing and measuring the utility of our proposed solution under different awareness levels of Facebook users.

We propose a Facebook privacy tool to clarify the impact of privacy leakages and our solution. The solution scheme demonstrates the functionality via simulations and a mock-up design. We provide a discussion on how this tool can be implemented in Section 3.4.

The proposed tool serves two post-based user categories: (i) POs and (ii) COs. Users in the PO category of a post can see the spreading area of this post depending on the initial privacy settings. Moreover, they can simulate the effect of change to the initial privacy setting of that post to observe an updated spreading area or change in possible number of interactions on that post, considering the fact that some COs of the post may change their actions depending on this privacy change. Users in the CO category can set some rules (delete my comment/like, inform me, etc.) for their interactions to protect their privacy. Thus, the tool will apply the predefined rules by the COs in the case of any privacy change by a specific post.

The rest of this paper is organized as follows. We review the related works in Section 2. We provide our methodology in Section 3 and present the following four points: (i) our proposed solution for mentioned privacy issues, (ii) an overview of our dataset and details of the experimental process, (iii) a possible implementation of our solution, and (iv) the applicability of the proposed solution. Section 4 concludes the paper with a discussion on the novelty of the proposed solution, including a comparative analysis with regard to the existing solutions, and presents the future work.

2. Related works for privacy issues on Facebook

This section revises the privacy issues in Facebook's functionalities and/or actions, including a discussion about the solutions that were proposed and how our approach differs.

The first notable privacy breach took place in December 2007, where Facebook received feedback about privacy concerns of its users. This issue had attracted the interest of the Federal Trade Commission (FTC), which entered talks with Facebook regarding online privacy and advertising rules that should be respected. Following the talks in 2011, Facebook decided to undergo evaluation of its privacy features by an independent authority annually. Furthermore, in April 2015, Facebook announced a serious change in installation of user's data by

third party applications, which enables tracking the amount of data that have already been compromised. Finally, the most recent event recorded in March 2018 caused a severe blow to Facebook's popularity, after Facebook was proven to have knowledge of massive theft of user's data.[8] To solve this issue, Facebook initiated closer tracking of installed applications and restricted their access to user's data. Facebook has also created tools displaying the data, which can be accessed by an application, allowing users to decide whether to grant permission to the application to access their data and/or uninstall the respected application.

Many researchers focusing on the analysis of privacy concerns and behaviors of SN users have shown interest in Facebook's privacy issues [4–7]. Furthermore, Ho et al. [8] and Tuunainen et al. [9] proved the existence of privacy issues in SNs, such as lack of awareness and lack of flexibility in current privacy tools. They also emphasized that a privacy tool should increase SN users' awareness of privacy. Talukder et al. [10] claimed that privacy settings may hide sensitive information (SI) for SN users (gender, political views, etc.), but that information can still be exploited by an adversary via some prediction techniques. For this purpose, they developed a tool, called "Privometer," to measure the level of privacy leakage of a user and direct them to a self-sanitization process (by recommending some options such as "Ask your friend to hide political view information") for their profile. Similar work was presented by Dong et al. [11, 12]. At the same time, Costantino et al. [13] created an application to measure and raise the privacy awareness of users regarding the visibility and access of online pictures. The results of this search revealed that, although most users restrict access to their photos to only friends, the main problem arises from the fact that not all their friends can be considered as trusted. Johnson et al. [14] explored mitigation of privacy concerns, which could alter the online behavior of interested users. They stated that combining privacy settings with the selection of trusted friends provided a better and more efficient outcome.

Although some researchers use social links to represent interactions between SN users and focus on interaction analysis [15, 16], Wilson et al. [17, 18] proposed the idea that social links do not actually represent real interactions among users. Instead, a few linked users can create an interaction graph (IG) with a big diameter and small number of super nodes. The IG aims to define the real interactions between a source node (PO) and the users that interact with his or her post (COs). This graph is different from the graphs created via social links, because friendship or other social relations do not represent the actual interactions among users. Interaction analysis may depend on many parameters, including individuals' behavioral factors.

Wilson et al. [17, 18] proved their approach over a dataset consisting of real Facebook user traces (real interactions). They created an IG based on this dataset, which included the same nodes with the social link graph, but only taking a subset of the links (only real interactions). Some of their findings are listed below:

- "Most users have no interaction with up to 50% of their friends."
- "For the vast majority of users, 20% of their friends account for 70% of all interactions."
- "Nearly all users can attribute all of their interactions to only 60% of their friends."

Although our study is different in approach, we will use these findings as a baseline in our experiments, because they analyzed the real interactions among SN users.

In general, our study is different from all the above-mentioned works, since it benefits from results that point to the mitigation of concern, accepting the fact that not all of our friends are evenly trusted. Therefore,

[8]Newcomb A. (2018). A Timeline of Facebook's Privacy Issues and Its Responses [online]. Website https://www.nbcnews.com/tech/socialmedia/timeline-facebook-s-privacy-issues-its-responses-n859651 [accessed 19-10-2018].

we propose a novel privacy tool that manages to act on our behalf whenever the privacy settings of a friend's post are changed. The proposed tool also visualizes the effects of a change in a post's privacy settings, which can affect its popularity, as the number of interactions taking place is also changed. If this tool can be adopted to Facebook, POs having inconsistent behavior in privacy decisions will deliver their posts to fewer people when the COs use this tool and protect the privacy of their comments. Following this, and to the best of the authors' knowledge, for the first time a tool could prompt the owner to decide whether to alter their interaction on a Facebook post.

3. Methodology

3.1. Proposed solution to cover privacy leakage on Facebook

We propose a novel Facebook privacy tool to increase privacy awareness of Facebook users. It is expected to dynamically track users' action logs, check for any changes in privacy settings of a post, detect possible privacy issues, inform users, and suggest some solutions in order to direct them to protect their own privacy.

The proposed tool provides the following facilities:

Fa. shows the PO the spreading area of a specific post based on its privacy setting,

Fb. shows the PO a possible IG that a specific post may get,

Fc. continuously checks whether the privacy setting of a post has been changed by its owner and sends a notification to the CO, and

Fd. allows any CO to set some rules for their interactions (comments/likes) on other people's posts. Defined rules will trigger a specific action when the described conditions are met (e.g., delete my interaction from a post if its privacy settings are changed to "Public").

Before showing a possible design and implementation of the proposed tool, we will first demonstrate our experiments, which serve as a background implementation, and emphasize the effect and importance of the detected privacy risks.

3.2. Experimental work for the proposed solution

Our experimental work consists of two parts: (i) research questions regarding expected functions of our tool are analyzed, and (ii) a simulation of the proposed tool demonstrating its operations for a Facebook user is created.

In these experiments, the SNAP Facebook Dataset,[9] which was collected from real Facebook users by Stanford University, is used. Each node represents a real Facebook user, but the corresponding information is anonymized. Edges represent friendship relations. In addition, nodes and edges may have some features, such as gender, age, and political views. According to an example from the SNAP webpage, if a user supports "Democratic Party" and specifies this on his/her Facebook profile, then this dataset keeps it as a feature of "political=anonymized feature 1". Hence, it keeps anonymized values for each feature. We transformed this dataset into a NetworkX[10] graph to simplify our tests as shown in Figure 1 in Section 3.2.2.

3.2.1. Functionality analysis of the proposed tool

To satisfy the aforementioned facilities of the tool and evaluate them on our dataset, we should answer the following research questions:

[9] Leskovec J, Krevl A (2019). SNAP Datasets: Stanford Large Network Dataset Collection [online]. Website https://snap.stanford.edu/data [accessed 15-03-2019].

[10] NetworkX (2019). Software for Complex Networks [online]. Website https://networkx.github.io/ [accessed 15-03-2019].

RQ1. How does the level of a PO's popularity affect the spreading area of his/her post and the number of interactions it may get from other COs?

This question is related to facilities Fa and Fb. We consider the number of a PO's friends (directly connected neighbor edges) as his/her level of popularity. For the simulation, we classified POs according to different popularity levels. Then we study the relation between their popularity level and the size of the IG created for their posts.

RQ2. How does a change in the privacy setting of a post affect the CO's decision on whether to still interact with it or not, and how do the rules predefined by a CO affect the number of interactions that the related post may get?

This question considers facilities Fc and Fd. It is mainly caused by the privacy preferences of POs and those may be one of the following in our study: (i) Friends: Only friends of the PO can see the post. (ii) FoF: Friends of the PO and friends of the tagged friends can see the post. To choose this option, the PO should tag some people. Tagging on Facebook means that the PO selects some of his/her friends to add in his/her post. It is a behavioral action that a PO performs with a real-time decision. However, we selected this probability as 20% based on the study of Wilson et al. [17, 18]. (iii) Public: Anyone can see the post even if he/she is not a Facebook member. After specifying these limitations, we performed our tests as explained below.

We first tried to discover the popularity level of each node in the dataset. We accepted the popularity of a node as the number of their neighbors (so the degree of the node represents its popularity level). Hence, we found the degree of each node, and then sorted them according to their popularity levels in descending order in a list (L_{pop}). Then we created ten different popularity classes and classified all nodes in the dataset according to their popularities as following: (i) 1st class: first 10% of nodes in L_{pop} (so this class keeps the most popular 10% nodes in the dataset); (ii) 2nd class: after the first 10% popular nodes, take the next first 10% remaining in L_{pop}, etc. Hence, the 10th class represents the least popular nodes in the dataset. As a result, we created ten different groups of nodes (Gr_{pop}) according to their popularity, and each group $g \in Gr_{pop}$ contains approximately 400 nodes, accordingly. For each group g, we performed the steps given below:

1. For each node Nd in g:
   - **a.** Find the friendship graph (FG) of Nd.
   - **b.** Find a possible IG for friends case.
   - **c.** Determine the friends to tag for the FoF case.
   - **d.** Create FoF graph of Nd.
   - **e.** Find a possible IG for the FoF case.

In step 1.b, we measure the size of the possible IG for each Nd. Based on the study of Wilson et al. [17, 18], we accepted the probability of a node to interact with his/her friends' posts as 20%. Hence, we assigned a decision value (whether to interact with the post or not) for each node in the FG of Nd, and then we created a possible IG. The decision value is specified randomly with 20% probability for yes (interact) and 80% probability for no. We call the probability of interacting with a friend's post p_F (so $p_F = 20\%$). This step was repeated 50 times for each node in the FG of Nd, and the average number of interactions was calculated. We used this average value to specify the size of the possible IG. The IG includes the nodes that have decision value "yes"

and the edges come from the friendship relations between these nodes. The size of this graph is equal to the number of its nodes.

In step 1.e, first of all, we stochastically find friends who will remove (withdraw) their interaction when the privacy setting of the post changes from Friends to FoF. Then we observe the change in the spreading area of the post, while testing for different probability values for a node to remove his/her interaction. We assign a removal decision value to each node (representing Facebook users' awareness level), based on the probability of withdrawal (p_w). The whole process is repeated for p_w values of 10%, 20%, 30%, 40%, and 50%, respectively. In each trial, we update IG for the "friends" case and extend it with the addition of new nodes from FoF. For the extension operation, we assign a decision value to each FoF node (same with p_F, so $p_{FoF} = 20\%$), and we add the ones with decision "yes" to the IG. The final graph represents the IG for the FoF case. Hence, the size of this graph is recorded as the number of interactions (nodes) for the FoF case.

Table 2 shows the change in the number of interactions per popularity for different p_w values. Rows start with the most popular nodes and continue downwards to the least popular ones. For each row, the columns represent:

Table 2. Change in number of interactions vs. p_w.

L_{pop}	$len(FG)$	$I(F)$	$len(FoFG)$	$I(FoF)$					
				$p_w = 0$	$p_w = 10\%$	$p_w = 20\%$	$p_w = 30\%$	$p_w = 40\%$	$p_w = 50\%$
$1st$	164.21	32.94	458.78	117.85	88.60	85.32	81.85	77.25	75.47
$2nd$	88.88	17.85	391.88	92.50	76.67	74.79	73.18	70.60	69.54
$3rd$	58.92	11.79	332.85	75.72	65.42	64.22	63.17	61.26	60.66
$4th$	41.50	8.34	288.87	64.07	57.02	56.13	55.31	54.15	53.58
$5th$	30.59	6.11	262.64	57.02	51.85	51.33	50.73	49.77	49.53
$6th$	22.66	4.54	224.45	48.10	44.41	43.99	43.51	42.85	42.63
$7th$	16.91	3.39	182.26	38.79	36.13	35.81	35.48	34.91	34.78
$8th$	12.23	2.46	173.96	36.35	34.53	34.31	34.12	33.78	33.50
$9th$	8.05	1.61	134.34	27.76	26.66	26.57	26.38	26.12	26.06
$10th$	3.91	0.78	106.97	21.68	21.33	21.21	21.20	21.04	20.98

- *len(FG):* average number of nodes in friendship graph for the corresponding popularity class (Step 1.a).
- *I(F):* average number of interactions for "friends" case, based on the popularity class (Step 1.b).
- *len(FoFG):* average number of nodes in FoF for the case of corresponding popularity class (Step 1.d).
- *I(FoF):* average number of interactions for FoF case (Step 1.e). We split the *I(FoF)* column into six subcolumns, one for the number of interactions without any removal ($p_w = 0$) and others with removals to observe the change in different p_w values.

Results show the following inferences:

- The popularity level of a node directly affects the number of interactions its post may get. As shown in Table 2, a node in the 5th class gets approximately 6 interactions, while a node in the 1st class gets 33 interactions.

- Number of friends seeing the post increases when we expand privacy settings (see columns *len(FG)* and *len(FoFG)*). For example, the visibility of a post created by a node in the 1st class increases from 164 to 459.
- Effect of the proposed tool can be observed in the case of removals ($p_w = \{10\%, 20\%, 30\%, 40\%, 50\%\}$). Spreading area of the post gets bigger when we change the privacy from "Friends" to "FoF" and the number of interactions it may get also increases dramatically. However, if COs can set rules for their comments, the rate of increase in the number of interactions may be smaller than the PO's expectation. By using this property, anyone can take precautions to protect his/her privacy by setting rules for his/her interactions on others' posts. For instance, the 1st row in Table 2 shows that although the PO's expected number of interactions for the FoF case is approximately 118 ($p_w = 0$), if users set some rules to protect their privacy with a probability of 50% ($p_w = 50\%$), this number decreases to approximately 75.

3.2.2. Simulation of the proposed tool for an individual node on Facebook

We provide a visual representation for our experiments in Section 3.2.1. We created a basic simulation on Spyder (Python 3.6), using the PyQt5 module. It works on the SNAP dataset and includes eleven steps. We explain the details of each step below, and also demonstrate a complete run for an average popular node as a PO. In this simulation, we used the following probability values: $p_F = 20\%$, $p_{FoF} = 20\%$, and $p_w = 20\%$.

Step 1: Visualization of dataset: Figure 1 shows the visualization of the SNAP Facebook dataset, which includes 4039 nodes. This figure also shows the graphical user interface (GUI) of our tool. The user can click the buttons at the bottom of the screen and continue the actions through the following steps.

Step 2: Selecting a PO: Here, to create an understandable example, we have selected an averagely popular node as the PO (with ID 766 from the 5th popularity class).

Step 3: Creating a post (Shared with "Friends" as default): A default post is created.

Step 4: Spreading the post and visualizing the spreading area ("Friends" case): Post is spread through the PO's friends, and then the spreading area is visualized in Figure 2. Size of the friendship graph is $len(FG) = 36$.

Step 5: Creating and visualizing a possible IG ("Friends" case): The tool creates a possible IG for the post based on the method we presented in Section 3.2.1. Then it visualizes the created graph as shown in Figure 3. The number of interactions for the "Friends" privacy setting is $I(F) = 5$.

Step 6: Changing privacy setting of the post from "Friends" to "FoF": The privacy setting of the post then changes to FoF. Now the PO expects to have a larger number of interactions on his/her post.

Step 7: Spreading the post and visualizing the spreading area (FoF case): The post is spread again, this time through PO's friends and friends of the tagged ones with $p_{FoF} = 20\%$. The tool visualizes the updated spreading area, which can be seen in Figure 4. As expected by the PO, the accessible number of nodes for this post increases as shown by the difference between Figure 2 and Figure 4. The size of the spreading area for FoF case is $len(FoF) = 48$.

Step 8: Creating and visualizing a possible IG (FoF case): The tool creates a new IG for the FoF case and visualizes it in Figure 5. Although the spreading area for the FoF case is bigger than the "Friends" case as mentioned in Step 7, it can be seen from Figure 5 that there is a decrease in the number of interactions compared to Figure 3 when the privacy changes from "Friends" to "FoF". Even if we increase the spreading area of a post by changing its privacy settings from friends to FoF, there is a possibility that the number of interactions it gets may decrease ($I(F) > I(FoF)$) according to the privacy awareness ($p_w = 20\%$) of the

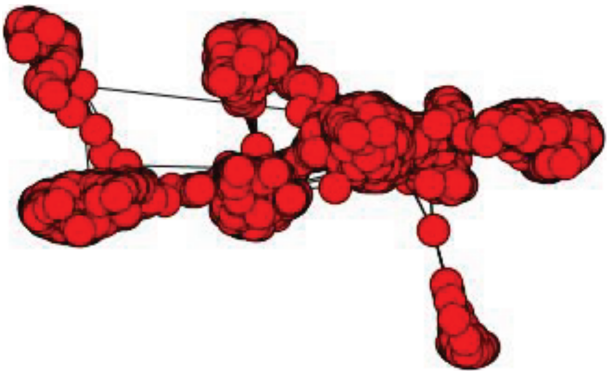

Figure 1. SNAP Facebook graph (shows all nodes in dataset).

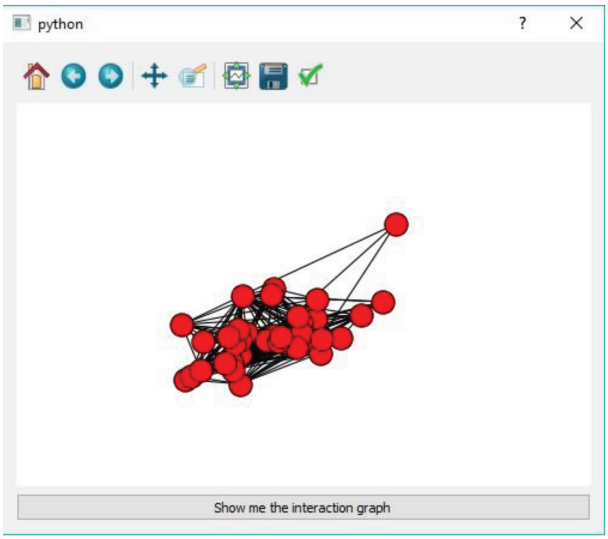

Figure 2. Spreading area (privacy: friends).

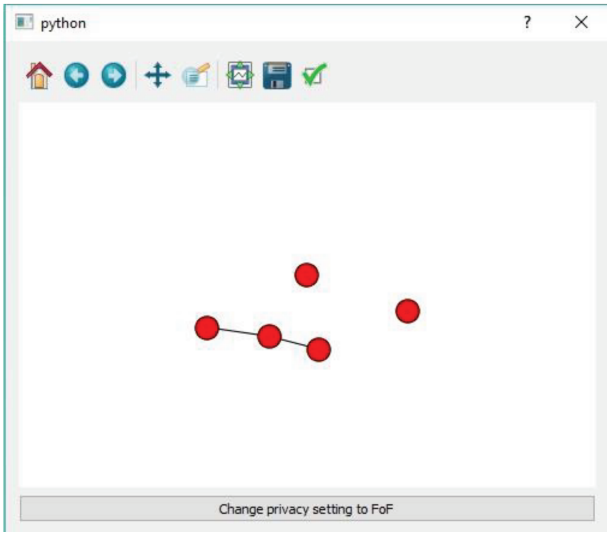

Figure 3. Interaction graph (privacy: friends).

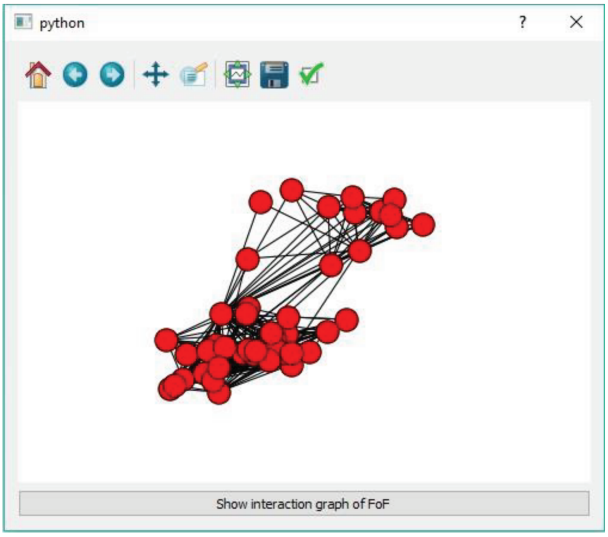

Figure 4. Spreading area (privacy: FoF).

COs. In this case, the PO's expectation about the increasing number of interactions on his/her post cannot be satisfied. This is not valid for all simulation scenarios, but there is a possibility that a decrease may happen if the COs' awareness increases. The number of interactions decreased to $I(FoF) = 4$ for this example.

Step 9: Changing privacy setting of the post to public: Privacy setting of the post is changed to public.

Step 10: Spreading the post and visualizing the spreading area (public case): This time, the post spreads through the whole network. The tool visualizes the updated spreading area for the public case in Figure 6. The size of the spreading area is 4039 (equal to the size of the dataset).

Step 11: Creating and visualizing a possible IG (public case): The tool creates a possible IG for the public case and visualizes the created graph as shown in Figure 7. Number of interactions in public case is 803.

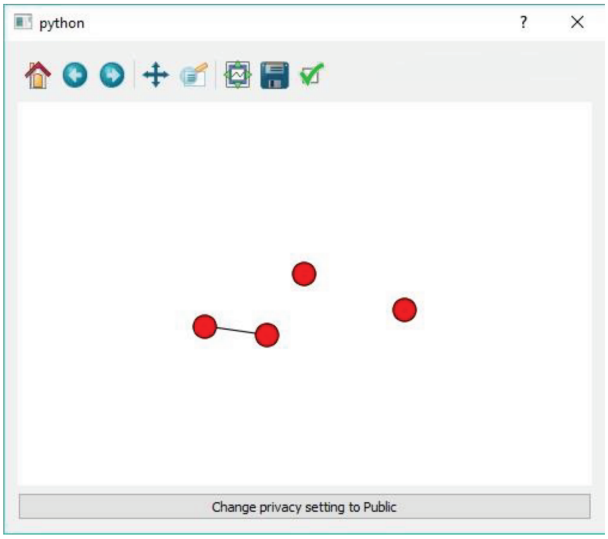

Figure 5. Interaction graph (privacy: FoF).

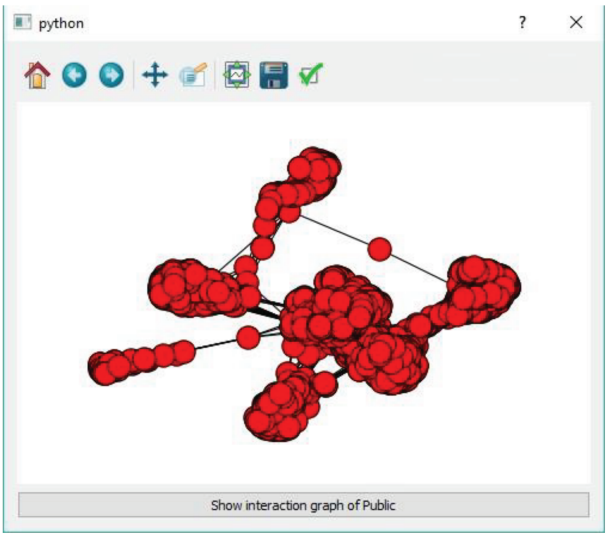

Figure 6. Spreading area (privacy: public).

3.3. Possible implementation of the proposed solution

We designed a mock-up to show how our tool will work. The mock-up was created via the proto.io website,[11] and a link to the implementation video of the mock-up[12] is given.

Our mock-up design includes eight main screens. It only shows dummy values such as the spreading area or number of interactions; it does not work with real data. The work flow is summarized in Figure 8 and further explanation can be found below.

Figure 7. Interaction graph (privacy: public).

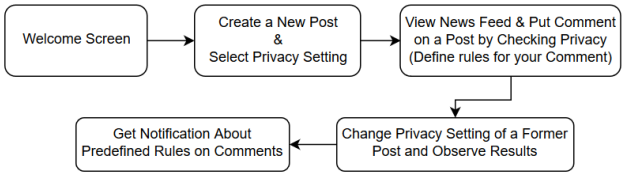

Figure 8. Flowchart of the mock-up.

[11] Proto.io (2019). Website https://proto.io/ [accessed 15-03-2019].
[12] Mockup Link (2019). Website https://www.dropbox.com/s/tmnjxf46r225rzq/mockup.mp4?dl=0 [accessed 15-03-2019].

1. **Welcome (Figure 9):** This screen welcomes users and informs them about the scenario embedded in the mock-up, the addressed privacy issue, and the content of images. Then it directs users to follow the mock-up.

2. **Create a post (Figure 10):** Post creation is simulated by using a dummy Facebook screen. In this screen, the post is ready as a default, and the PO is expected to proceed with this default post.

3. **Select a privacy setting (Figure 11):** This screen directs the PO to select the privacy setting of his/her post as "Friends". According to this setting, the spreading area of this post is visualized via a subscreen, as shown in Figure 12. As seen in the figure, dummy numbers for the size of the spreading area for each case are given at the top, with red, orange, and green colors. Furthermore, the corresponding area is represented with circles. In real situations, those numbers will show the personal data for each PO.

4. **View news feed (Figure 13):** This screen simulates some dummy posts created by POs or COs' friends. This property is referred to as "News Feed" on Facebook.

5. **Put a comment (Figure 14):** This screen directs the user to interact with a specific post, by leaving a comment, and so the user becomes a CO. A dummy comment is already placed in this screen, so the CO just needs to proceed to the next screen.

6. **Check privacy (Figure 15):** When the CO wants to leave his/her comment, he/she is asked to check the privacy of the comment before doing so. Hence, this screen shows a message that asks whether the CO prefers to set a rule for this comment to protect its privacy. If so, the mock-up proposes some rule options for him/her (i.e. "Send me a notification!", "Delete my comment without notifying me!", and "Delete my comment and notify me!"), and he/she can select one of these rules.

7. **Change privacy setting of former posts (Figure 16):** This screen represents the PO's previously shared posts. The PO is directed to select a specific post created by him/her before and then try to change the privacy setting of it to public. Before applying the corresponding change in the privacy settings, the mock-up demonstrates the newly created spreading area, as shown in Figure 17. The PO can see the results of this change in the number of interactions on this post as percentage values at the bottom of the screen. These results are affected by the rules created by COs of the post. Numbers in Figure 17 are dummy, used for only this simulation, and they should reflect the real data in the original application.

8. **Get notification from tool (Figure 18):** This screen demonstrates a notification that the tool sends to COs in the case of a change in any of a post they interacted with. Thanks to this notification, COs can control the privacy level of their interactions affected by the changes in privacy settings.

3.4. Applicability of the proposed tool on Facebook

The proposed tool demonstrates some crucial privacy risks for Facebook users, while sharing posts as a PO and interacting with shared posts as a CO. Hence, it gives a valuable warning for Facebook users, who may have either a PO or a CO role at any time. Facebook should choose one of the following three options to handle those problems:

1. Facebook may prevent the POs from changing the privacy setting of their former posts.

2. If Facebook consciously allows POs to change those privacy settings, then the proposed tool may served centrally as a service to Facebook users. This requires Facebook to meet requirements in RQ1 and RQ2, proposed in Section 3.2.1. For RQ1, Facebook should send COs a notification in the case of a privacy change on a post that they have already interacted with. For RQ2, Facebook should allow COs to access all of their activity logs, including their whole interactions, although the visibility of related posts has changed. This solution sounds handy, but actually it creates an extra complexity for Facebook. In fact, Facebook faces a serious processing cost, because there is a huge amount of data processed over Facebook every minute. According to Facebook statistics in December 2017,[13]: (i) "Facebook generates 4 new petabytes of data per day", (ii) "Users generate 4 million likes every minute", and (iii) "350 million photos per day are uploaded by users".

   Statistics show that the data processed over Facebook are huge and dynamic. That is why implementing our tool to track each specific post creates an enormous cost for Facebook. In addition, each Facebook user may not be concerned with his/her privacy to the same extent. Under these circumstances, Facebook may consider giving this service for payment. Hence, users may pay for the service to protect their privacy.

3. This tool may be implemented and provided to Facebook users by third parties. However, this may cause other privacy risks because the third party should demand the daily Facebook data via a contract. Moreover, the application served by third parties should be installed by all Facebook users; otherwise, they cannot handle the whole privacy risks. This is unrealistic and inapplicable.

According to our survey, 61% of respondents think that this service should be given by Facebook, while 29% claim that this tool can be given as a service by any trusted third party and 10% think that this solution is inapplicable. Finally, 85% certainly agree that this problem should be solved by Facebook through policy regulations and service refinements. After evaluating these options, let us assume that Facebook did not disclose the privacy leakages, did not offer the proposed service, and even did not permit third parties to adopt our solution. Our study still plays an important role by giving notice to Facebook users that they face serious privacy issues; 95% of respondents verified that our mock-up will increase the privacy awareness of users, confirming the real impact of our study.

4. Discussion and conclusions

This study was triggered by the fact that SN users continuously face privacy issues, although they might not even be aware of them, and there is a rising need for the development of a tool to allow them to easily control their privacy settings. We focused on the issue of the privacy scope of personal posts on Facebook. Therefore, we analyzed current privacy risks on Facebook and specified two detected ones: (i) personal interactions may be visible beyond the expected area, and (ii) personal interactions may be inaccessible to owners. After that, we proposed a tool that provides a solution to both problems, but the experimental work mainly focused on the first privacy issue. A simulation focused on post creation and spreading, observing the change in spreading area according to the PO's popularity level and the preferred privacy setting for the post.

Table 3 shows a comparison of the existing models with the proposed one in terms of challenges and solutions in detail. The most critical issue in our study is that similar studies related to privacy threats mostly

[13]Smith K (2019). 47 Incredible Facebook Statistics and Facts [online]. Website https://www.brandwatch.com/blog/47-facebook-statistics/ [accessed 15-03-2019].

result from technological factors (e.g., copy/paste capability), privacy settings of the user's profiles or posts, or lack of information of members. However, the reason for the detected privacy problem in this study is unexpected human behaviors. The Facebook privacy policy and Facebook members assume that all friends are credible and the predefined privacy settings will remain the same. According to general security rules, we cannot trust any of those entities to ensure security. Therefore, an awareness of this risk should be created to evaluate all kinds of technological products, software, and services while defining a privacy policy.

Many researchers focused on privacy issues in SNs [4–14], but we detected two important privacy leakages that still exist on Facebook. We can compare our study with "Privometer" [10], which investigates a different privacy issue, as explained in Section 2. Both works use real Facebook member profile data and present a simulation to clearly show and verify the problem and its possible effects. In this study, problem verification and its impacts are analyzed for different awareness levels of COs and for different popularity levels of POs in Section 3.2.1 (Table 2). This simulation is similar to the simulation part of Privometer, which visualizes different awareness levels of members on a friendship graph by their effects on visibility of private information. Both of them propose a tool to explain the privacy risk, present a preventative course of action to members, and visualize its usability. For the validation of the proposed solutions and the suggested tool, our study also includes survey results that were collected from more than 500 people. This survey presents a mock-up design to explain the impact of the problem and the proposed tool, asking them about the usability and the impact on improving awareness. Finally, our study also presents a discussion about the applicability of the proposed solution, verifying it with the questionnaire.

Table 3. Comparison of related works in terms of challenges and solutions.

Study	Type	Privacy challenge	Proposed solution	Novelty of study
[13]	Empirical survey and solution proposal for the highest privacy concern	General copy/paste capability on photo links creates privacy risk	Automatic deletion of links after 30 days	The reason for leakage is coming from a utility that supports computer usage
[14]	Empirical survey to analyze privacy concerns on post.	Analyzing Facebook users' privacy concerns on post-sharing through specific subgroups of their friendship networks	Identification of related privacy-preserving strategies for identified concerns of Facebook users	This study contributes to human-centric design efforts to improve privacy settings of SNs
[9]	Empirical survey to analyze privacy awareness on profile setup	Access to all types of profiles and collection of real data.	Improving privacy awareness of Facebook users by survey results	Survey of privacy concerns of Facebook users to analyze their preferences on profile visibility
[10]	Simulation of the privacy problem via a tool called "Privometer"	Detecting the privacy risks stemming from privacy preferences and profile setups of friends is a hard problem for Facebook members	"Privometer" allows analyzing privacy risks of member profiles in a friendship graph, visualizing hazardous friends, and proposing a sanitation mechanism	No one can manually monitor their own extending friendship graph in real time and dynamically identify privacy risk; Privometer tries to give this service offline or online at limited levels
Our study	Empirical survey, simulation of the problem and its impact, proposing a tool and evaluating possible solutions.	Detection of privacy leakages, which mostly stem from trusted relationships among friends.	Two detected privacy leakages in post/comment dissemination are simulated and validated; the authors propose a solution tool; all alternative solutions and the proposed tool are also validated and verified by a survey.	Presented privacy leakages, which are rooted in the Facebook privacy policy, are detected for the first time

We can extend our study with a more elaborated survey analysis; for instance, it can be applied to larger groups and these groups may have homogeneous demographic structures and different geographical regions. As a result of this extension, the impact of this study can be investigated more efficiently with different privacy preferences. Our approach to analyzing and handling privacy issues can be applied to different social networks, and many more privacy risks can be explored in the future. However, the realization of each solution may be computationally costly since this requires huge amounts of online data processing. This also affects the usability of SNs, and response times may be uncomfortable for users. Regulations of SNs should consider the system privacy and security to create more secure digital domains and services. Finally, as a future work of this study,

we will focus on simulating the actual interaction graph for a post by developing a prediction model to represent possible interactions depending on SN users' behaviors.

Acknowledgment

This work was supported by the Council of Higher Education in Turkey, within the scope of the Project Based Mevlana Exchange Programme (project code: MEV-2016-057). All authors assisted significantly in the research work and the manuscript drafting stage. All authors read and approved the final manuscript. The authors would like to extend their gratitude and appreciation to Professor Sorin Adam Matei from Purdue University, USA, for his valuable comments and insight, which have contributed to the improvement of this paper.

References

[1] Boyd D, Hargittai E. Facebook privacy settings: Who cares? First Monday Peer Reviewed Internet Journal 2010; 15: 8. doi: 10.5210/fm.v15i8.3086

[2] Gangopadhyay S, Dhar D. Social networking sites and privacy issues concerning youths. Global Media Journal 2014: 5 (1): 1-7.

[3] Shore J, Steinman J. Did you really agree to that? The evolution of Facebook's privacy policy. Technology Science Internet Journal 2015. doi: 10.7910/DVN/JROUKG

[4] Siddula M, Li L, Li Y. An empirical study on the privacy preservation of online social networks. IEEE Access 2018; 6: 19912-19922. doi: 10.1109/ACCESS.2018.2822693

[5] Al-Shamaileh O. I have issues with Facebook: But I will keep using it. IEEE Technology and Society Magazine 2018; 37 (2): 40-45. doi: 10.1109/MTS.2018.2826078

[6] Hossain AA, Zhang W. Privacy and security concern of online social networks from user perspective. In: International Conference on Information Systems Security and Privacy; Loire Valley, France; 2015. pp. 246-253.

[7] Al-Shamaileh O, Aloudat A, Barikzai S. User concerns about Facebook: are they important? In: 8th International Conference on Information Technology; Amman, Jordan; 2017. pp. 291-296. doi: 10.1109/ICITECH.2017.8080015

[8] Ho A, Maiga A, Aimeur E. Privacy protection issues in social networking sites. In: 7th IEEE/ACS International Conference on Computer Systems and Applications; Rabat, Morocco; 2009. pp. 271-278. doi: 10.1109/AICCSA.2009.5069336

[9] Tuunainen V, Pitkanen O, Hovi M. Users' awareness of privacy on online social networking sites case Facebook. In: The 22nd Bled eConference; Bled, Slovenia; 2009. p. 42.

[10] Talukder N, Ouzzani M, Elmagarmid AK, Elmeleegy H, Yakout M. Privometer: Privacy protection in social networks. In: IEEE 26th International Conference on Data Engineering Workshops; Long Beach, CA, USA; 2010. pp. 266-269.

[11] Dong C, Jin H, Knijnenburg BP. Predicting privacy behavior on online social networks. In: Proceedings of the Ninth International AAAI Conference on Web and Social Media; Oxford, UK; 2015. pp. 91-100.

[12] Dong C, Jin H, Knijnenburg BP. PPM: A privacy prediction model for online social networks. In: Ninth International AAAI Conference Web and Social Media; Oxford, UK; 2015. arXiv: 1606.07463v1

[13] Costantino G, Sgandurra D. Design and development of a Facebook application to raise privacy awareness. In: 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing; Turku, Finland; 2015. pp. 583-586. doi: 10.1109/PDP.2015.23

[14] Johnson ML, Egelman S, Bellovin SM. Facebook and privacy: it's complicated. In: Proceedings of the Eighth Symposium on Usable Privacy and Security; Washington, DC, USA; 2012. pp. 1-15. doi: 10.1145/2335356.2335369

[15] Laleh N, Carminati B, Ferrari B. Risk assessment in social networks based on user anomalous behaviors. IEEE Transactions on Dependable and Secure Computing 2018; 15 (2): 295-308. doi: 10.1109/TDSC.2016.2540637

[16] Vapen A, Carlsson N, Mahanti A, Shahmehri N. A look at the third-party identity management landscape. IEEE Internet Computing 2016; 20 (2): 18-25. doi: 10.1109/MIC.2016.38

[17] Wilson C, Boe B, Sala A, Puttaswamy KPN, Zhao BY. User interactions in social networks and their implications. In: Proceedings of the 4th ACM European Conference on Computer Systems; Nuremberg, Germany; 2009. pp. 205-218.

[18] Wilson C, Sala A, Puttaswamy KPN, Zhao BY. Beyond social graphs: user interactions in online social networks and their implications. ACM Transactions on the Web (TWEB) 2012; 6 (4): 17. doi: 10.1145/2382616.2382620

A. Main screens of the mock-up.

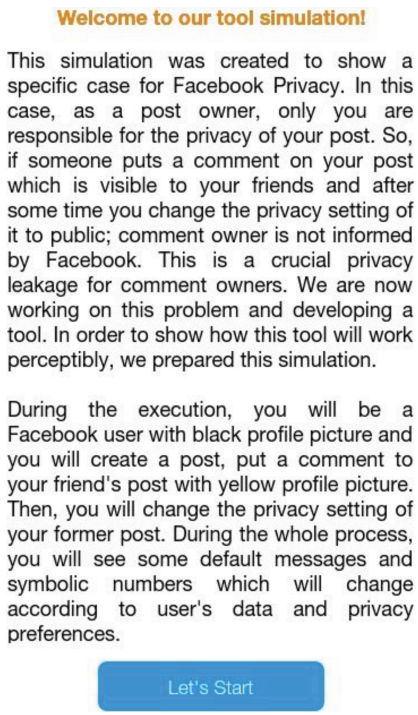

Figure 9. Main Screen 1.

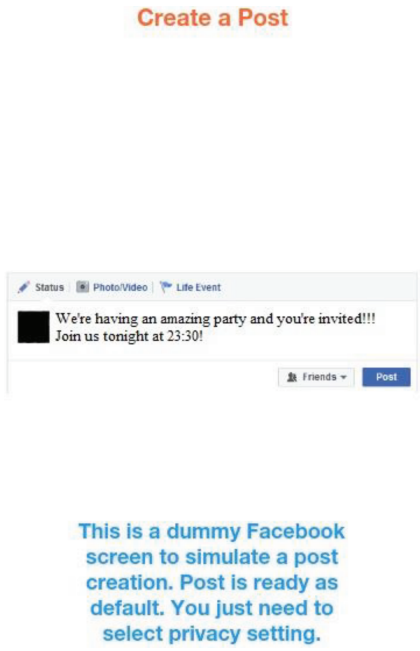

Figure 10. Main Screen 2.

Figure 11. Main Screen 3.

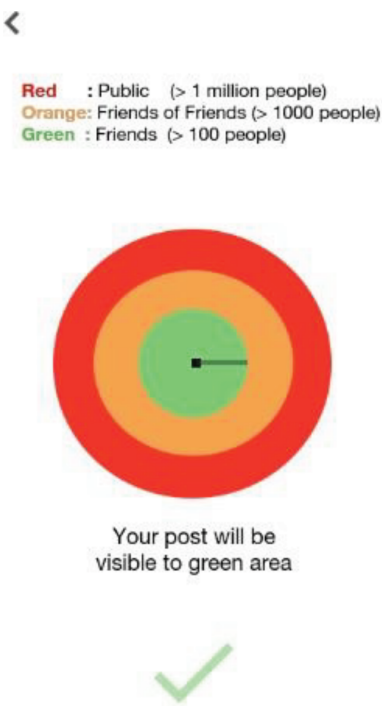

Figure 12. Mock-up screens for "Select Privacy Setting (Friends)".

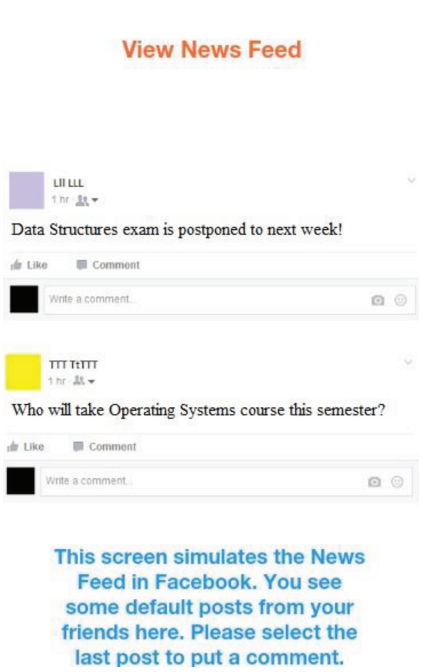

Figure 13. Main Screen 4.

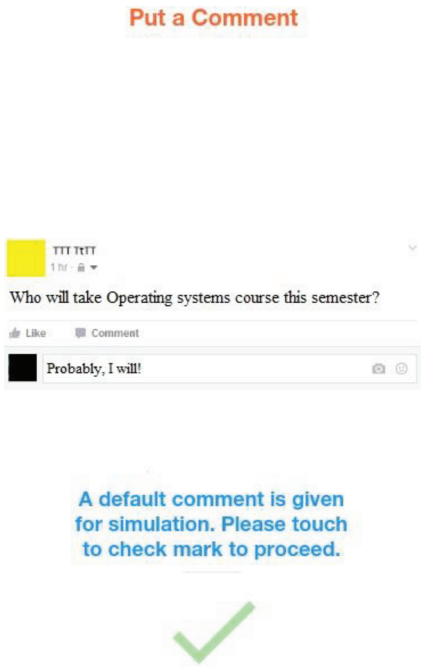

Figure 14. Main Screen 5.

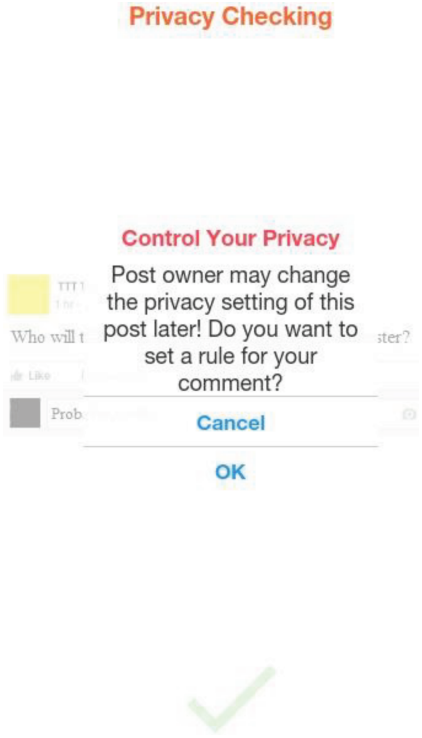

Figure 15. Main Screen 6.

Figure 16. Main Screen 7.

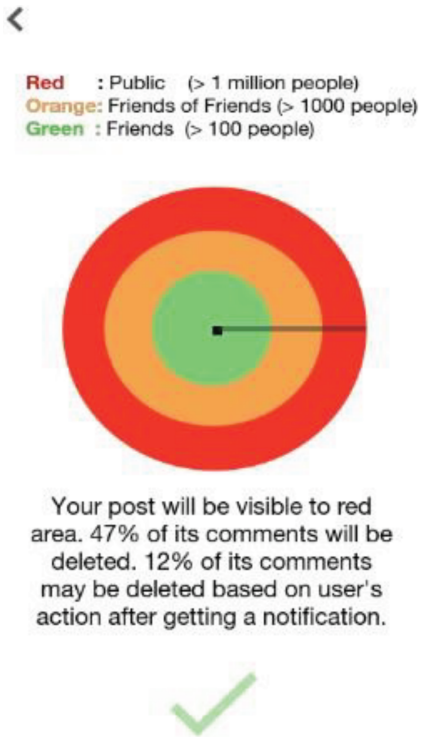

Figure 17. Mock-up screen for privacy change to public.

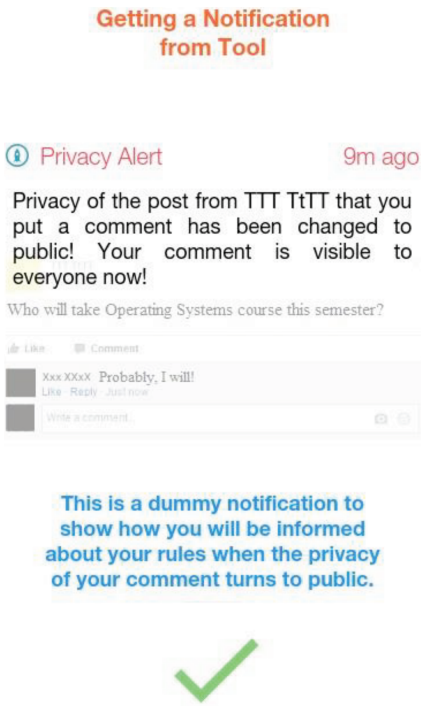

Figure 18. Main Screen 8.