A Tool for assessing Privacy Awareness in Social Networks

Giuseppe Primiero, Lalith Athiappan, Franco Raimondi, and Balbir Barn

Department of Computer Science
Middlesex University London (UK)
{G.Primiero,L.Athiappan,F.Raimondi,B.Barn}@mdx.ac.uk

Abstract. Social network activities directly affect the exposure of users' personal data. Monitoring changes in users' behaviour is therefore crucial to privacy concerns and it facilitates balancing the trade-off with trust and security in the network. This paper proposes a proof-of-concept model for computing privacy levels based on behavioural changes. The model relies on a Bayesian Network whose nodes are elements of a Theory of Informational Privacy. We detail the conceptual frame and illustrate its implementation as a tool in the Facebook API. Preliminary evaluation is offered through three study cases, exemplifying different privacy profiles. We conclude by highlighting a number of feasible extensions.

1 Introduction

Online social networks have seen a massive growth in recent years and are playing a key role in shaping the nature of human interaction. In the first quarter of 2016, the number of active users on Facebook reached over 1.65 billion [27]. Increasingly though, interactions mediated by online social networks (OSN) are not without risk, with key concerns at both individual and societal level related to privacy, trust and security. Individual sense for the first of these notions is contingent on various relational concepts and it comes to the fore through interactions either as the right to *limit* or to *control* access to a personal domain. The complexity of such notion is evident in its being a continuum between two dimensions, where total privacy may be as undesirable as total transparency. Individual privacy is also relative with respect to temporal change. An individual on an OSN such as Facebook may choose to participate in interactions with full knowledge of a relative privacy concern. At a later time, the net awareness of combining the effects of further interactions on the total privacy concern may no longer be transparent.

In this paper, we explore the specific question of how a user's privacy state changes across time and how it is moderated by sharing activities on an OSN. We propose that monitoring such change can offer a dynamic view on the privacy concern, a parameter on the network's overall trust and a measure of the user reputation. Furthermore, a better understanding of such state change can help inform policy makers addressing the trade-offs between privacy and security.

Our contribution to this problem is an operationalisation of existing theories of informational privacy using Bayesian networks to analyse user interactions on OSN and the development of an appropriate tool.

The reminder of this paper is structured as follows. In Section 2 we formalise our research question on how to account for the change in a user's privacy measure within the context of relevant background literature. In Section 3 we review the notions underlying a theory of Informational Privacy for the context of social networks. In Section 4 we propose an implementation in Java of this theory through a Bayesian Network. In Section 5 we present a Facebook plugin implementing the model. In Section 6 we offer a preliminary evaluation through a limited number of study cases. Finally, in Section 7 we briefly reconsider our contribution in view of the current approaches to privacy in social networks and their relation to trust and reputation and suggest further work.

2 Background and Formalisation of Research Question

Threats to personal privacy in social networks [23] focus in particular on voluntary information disclosure [8, 15], incomplete personal information [31] and control on web-tracking activities [28]. A survey on transparency tools is available in [12]. The relationship between privacy and trust, issues of methodology, user reputation, various theoretical and socio-demographic considerations [13, 11, 17], design of access control [18], requirements elicitation [25], are all addressed in a survey by Hugl [14]. Hence, privacy awareness and risk assessment are largely available in the literature as privacy concerns, see [29], but these models fail to consider how user's privacy *changes* across time. Monitoring such change can offer a dynamic view on the privacy concern, representing a parameter on the network's overall trust and a measure of the user reputation. The problem can be more precisely formulated as follows:

Problem 1 (Privacy Change Assessment). Given a measurable amount of network activity by user U_1 over a fixed span of time Δt, how does the probability p that information i shared by U_1 is exposed to some non-connected user U_2 change over Δt?

Let us indicate with $NA_{\Delta t}(U_1)$ the network activity of User1 over some span of time $t' - t$, with $t < t'$ temporal units; $p(U_1 i U_2)$ indicates the probability that information i issued by User1 reaches User2; then our problem can be restated by the following formula:

$$PC(\Delta t) = (p(U_1 i U_2)_{t'}) - (p(U_1 i U_2)_t) \mid NA_{\Delta t}(U_1)$$

with $PC(\Delta t)$ indicating a measure of *privacy change* over time. The evaluation of such parameter is based on the assessment of the user's network activity (including reactions to such activity from other users), which hence needs to be decomposed in – and calculated from – measurable units.

The role of Bayesian models for privacy is currently explored in several directions. In [30], a probabilistic model is developed for information leakage over networks: the main difference in our approach is that we do not assume or require an adversary, and are interested only in assessing the sender's behaviour. In this light, our tool is designed to instruct a potential modification of such behaviour when it leads to undesired exposure. In [10], the focus is on privacy breaches, rather than unaware personal practices which might lead to undesired information leakage. Moreover, some approaches differ also depending on which of the stakeholders are involved. For example, in [1] the focus is on the difficulties that *developers* encounter in defining and implementing privacy decisions and in [19] the interest is in soliciting more informed decisions on app-selection from mobile users. In our case, we are interested in monitoring a series of different metrics on the user side only, providing an intuitive qualification of what it means for information to become more or less likely to reach outside of the immediate range of connected users, when that is not the user's intention.

When considering information leakage without adversaries in online social networks, much of the literature focuses on the privacy settings configuration of the service provider: [21] focuses on the exploits due to conflicts between privacy configurations and network functionalities; [20] identifies minimal information sets and sharing practices required for successful interactions, but it does not provide any specific monitoring tool; [22] offers a sophisticated calculus to evaluate the trade-offs between privacy and risks and delivers a decision method to predict the user's intention to share information online, a contribution which is quite close to ours in spirit, but it does not necessarily aims at improving user's behaviour. In [9], a holistic approach to privacy in social networks encompassing surveillance, institutional considerations and social privacy is supported: we consider such an approach essential, and suggest that improving on user's behaviour is also crucial to preserve values and security over social networks. Previous studies (e.g. [29]) have shown how the privacy concerns, albeit significant, do not directly affect users' acceptance of social networking web sites [3], while they do moderate the effects of perceived usefulness, ease of use, and intention to continue to use the service, i.e. trust in it, as shown by recent empirical studies [26, 11].

The literature reviewed suggests that the Privacy Change Assessment Problem formulated above has not yet been addressed. Computing a value of privacy change can help making users aware of behaviour which might increase risks of information leakage. On such basis, network design and policy making issues can be further addressed: a user with high privacy levels can be given higher reputation and a network with low information leak can be deemed more trustworthy. In the present paper, we approach this problem with a proof-of-concept model based on a Bayesian network. The model requires definitions of the elements used as prior and posterior probabilities to build a measurable notion of user's network activity. To this aim, we start from the basic elements of the informational theory of privacy developed in [6]. Such theory of informational privacy is a discursive proposal evaluated through the use of argumentation. Some effort has been already made to make the theory suitable for consumption by software engineers through the development of a conceptual model formalised using

UML [2], but there is no mechanism for its operationalisation. In this context, our contribution in this paper is along two dimensions. Firstly, we contribute a theory for informational privacy for social networks reified through the use of measurable units of actions. Secondly, the theory is operationalised through an implementation using Bayesian networks and the development of a tool. The provision of a tool is a necessary first step towards quantifiable evaluations of privacy awareness in social networks.

3 A theory of Informational Privacy for Social Networks

Privacy in the real world has been defined in relation to other people's behaviour or public attention. Its meaning is fluid in different contexts, leading to several distinct definitions. In [4], privacy is conceptualized positively by a user's personal space (physical dimension), personal integrity (psychological), interaction with others (social), and personal data (informational); and negatively as the absence of invasion of the individual sphere by the government, businesses, or other actors. In considering privacy for Internet-based applications and on-line social networks in particular, we focus on the *informational dimension.* A theory of informational privacy is offered in [6, 7], construed around four basic notions:

1. the ontological features of agents determine the *information accessibility*;
2. accessibility defines the *informational gap* between agents;
3. the interaction environment establishes a degree of *ontological friction*;
4. friction regulates the *information flow* within the system.

By the first item one understands the agents' behaviour and attitude towards information sharing; by the second, the larger the gap, the less agents know about each others; the third item qualifies structural properties of the world (analogical or digital) in which agents share information; by the latter term, one refers to the structural properties of information channels. Informational privacy then becomes a function of both the information gap between agents and the information flow in the infosphere. In previous work [2], we offered a UML presentation and an axiomatic translation for these concepts, usable for software engineering purposes.

For Online Social Networks, privacy has been mainly referred to as data protection, while the aspects related to awareness of information leakage and behaviour's monitoring practices for information sharing processes have been less considered. To offer such a treatment, we start by appropriate informal and formal counterparts of the above notions, specifically formulated for the social networking environments. Let us start defining a notion of Information Access:

Definition 1 (Information Access). *A measure of an agent's activity on the network.*

The value for information access is calculated in Section 5 by extracting data related to user's actions on the network. These will depend on the type of social network, but will typically include: amount and types of items posted, whether or not these allow geo-location visibility, and the increase in outreach. These reflect the attitude of the agent towards sharing practices on the ONS.

The second basic notion is that of Network Friction:

Definition 2 (Network Friction). *A measure of an agent's network fluidity.*

This value measures some quantifiable notion of *density* of the user's network: in particular, it accounts for the flow of information to nodes in the network that have a degree of separation $n > 1$, in terms of shared items that are visible beyond the limits controlled by the users. By this parameter, we aim at reflecting the 'ontology' of the real world to a precise interpretation in its digital counterpart.

The two concepts above are used to infer values of two further notions:

Definition 3 (Information Gap). *A measure of the degree of accessibility to personal data, as a function of informational access.*

This notion quantifies the exposure of user's data on the network, and it is computed by weighting user's activities according to her visibility settings. In other words, the reachability of data is inversely proportional to the information gap: a piece of information posted with a *Public* setting has more effect on reducing the gap to other users than a post with a *Private* setting.

Definition 4 (Information Flow). *A measure of the degree of fluidity of personal data as a function of network friction.*

This value expresses the proportion between the fraction of user's activity which is publicly accessible (according to the properties of the network used, e.g. Likes and Events in Facebook) and the fraction of network activity which is instead shared with some restrictive properties (e.g. only to some pre-defined group of friends or family-members circle). This notion gives us a general idea of how much of the overall activity is exposed, hence rendering a partial image of the user's behaviour. While Definition 3 interprets a quantifiable exposure based on the privacy settings for the information produced by the user, Definition 4 gives the proportion of the user activity which is more exposed.

Figure 1 offers an intuition for the relationship between the values of Information Gap, Information Flow and Privacy. It presents privacy on the y axis on a $0 - 2$ scale, Information Flow on the x and Information gap on the z axes on a $0 - 1$ scale. The axes have decimal points. The intuition at the basis of the Bayesian Model presented in Section 4, is that for the maximum of Flow and the Maximum of Gap one gets the minimum of Privacy; the less information is shared and the more close the network is, the more privacy increases. The model will assess how this relation *changes* in virtue of the user's activity on the Network, also taking into account how the Network itself grows.

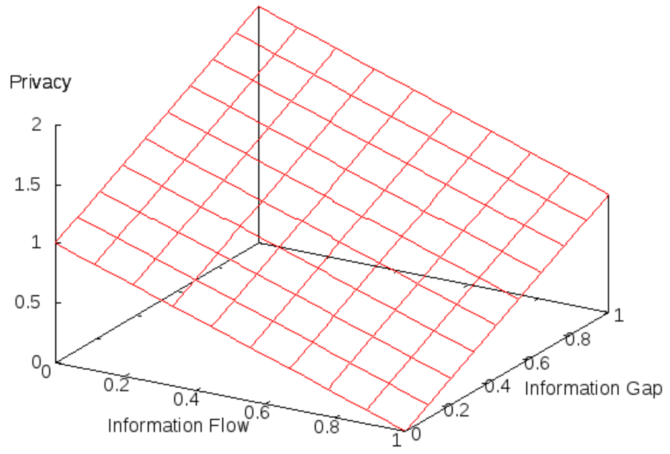


Fig. 1. Privacy as a function of Information Gap and Information Flow

4 Bayesian Network implementation

A Bayesian network (BN) is a compact representation of probabilistic relationships between properties in a domain. A BN is constructed as a directed acyclic graph where nodes represent properties with values depending upon a probabilistic correlation between two nodes [16]. A BN also describes a probability distribution over all its variables using conditional probability tables (CPTs). Each node may have a CPT that describes the probability of each possible state of the node given each possible outcome of the parent nodes. Nodes without parent nodes are given prior probabilities.

In the following, we implement the theory of Informational Privacy for Social Networks presented in section 3 in the Java API for the Netica Software [24]. The testing has been performed on a Ubuntu 14.04 LTS system, with 4Gb of memory, Intel Core i7, 2M CPU @ 1.70ghZ and Java OpenJDK 1.7. For easy reference to the lines mentioned in the explanation below, the code is presented in Appendix I and is also available for download at `https://github.com/gprimiero/InfoPriv`.

The implementation starts from creating nodes according to the Definitions from Section 3 (lines 14-31). Nodes for Information Access and Network Friction are parent nodes with prior probabilities; this means they determine respectively Information Gap (line 33) and Information Flow (line 35) as dependent nodes; the latter two are used to compute the conditional probability value of Information Privacy (lines 37-41). Each node is given a probabilistic value on a scale $0-100\%$. In this formulation, all prior values are balanced at $50-50\%$, indicating a standard or average evaluation for network access (line 47) and network fluidity (line 58). Informally, this can be understood as saying that we start from an expectation of moderate use of the network from the user in a standardly dense network. Both values can mean different things in distinct social, geographical and cultural contexts, hence an extended experimental approach will be needed in the future to characterise these more precisely. We assume in the present work low proportions of public activity (lines 52-53) and exposure to external nodes (lines 64-65), setting them at only 10%. Computation of the value of information privacy (IP, lines 67-71) is given by the following equation:

$$IP = (max(InfoGap) + min(InfoFlow))$$

Its dual, information openness (IO), is computed as follows:

$$IO = (min(InfoGap) + max(InfoFlow))$$

As for the graph in Figure 1, these are values on a scale $0.00 - 2.00$, obtained by the sum of the two $0 - 100\%$ scales for Information Gap and Information Flow respectively. We assign variable names to the evaluation range of our conditional probabilities (lines 73-95):

- *open* is the value associated with the presence of information gap;
- *close* is the value associated with its absence;
- *easy* is the value associated with the presence of information flow;
- *difficult* is the value associated with its absence.

We finally calculate the value of IP as the sum of *open* and *difficult* and the value of IO as the sum of *close* and *easy* (lines 97-111).

5 Tool

We have implemented the Bayesian model of informational privacy as a plugin for Facebook, available at `http://ta.mdx.ac.uk/sompri/`. The main aim of the tool is to offer the user a simple and accessible set of meters and a detailed metric of the data used in order to compute the change in exposure from a selectable previous measurement. After the user concedes authorization, the plug-in retrieves posts, details, events etc. from a given Facebook profile, computes values based on Definitions 1 to 4 and uses them as input for the Bayesian Network described in the previous section. Concretely, these values are made available through the Facebook Graph API [5], which provides access to the Facebook social graph. The results are displayed in graphical form as shown in the user examples in Figures 3 and 4. The intended effect is to make the user aware of the way her use of the platform exposes her to unintended information leakage, how such exposure changes over time and possibly to influence the user's behaviour.

Given a time interval Δt, ranging from 1 day to 1 year in our implementation, we define the value of Information Access (Definition 1) in the interval $(-t, 0)$ as:

$$\begin{aligned} IA(-t,0) = {} & \text{N.of likes}(-t,0) + \text{N.of locations}(-t,0) + \\ & + \text{N.of events}(-t,0) + \\ & + \text{N.of posts}(-t,0) + \text{N.of new friends(-t,0)} \end{aligned} \tag{1}$$

where 0 is the current time, "N. of new friends", "N. of likes" and "N. of posts" have their obvious meaning in the time interval, "N. of locations" is the number of places that the user has tagged in pictures, posts, etc., "N. of events" is the number of the events that the user has *accepted* to attend. Intuitively, all the parameters in this formula capture how much a user is active in a network; for instance, if the total number of friends in a time interval decreases, the variation is negative and this, in turn, contributes negatively to IA.

We define the value of Network Friction (Definition 2) in a time interval t as:

$$\begin{aligned} NF(-t,0) = {} & \text{N.of likes}(-t,0) + \text{N.of locations}(-t,0) \\ & + \text{N.of events}(-t,0) + \text{N.of public posts}(-t,0) \end{aligned} \tag{2}$$

where "number of public posts" includes the number of posts whose accessibility is set to everyone and the posts by *other* users on the user's timeline. Intuitively, the parameters in this formula capture the activities that are relevant to the fluidity of information from the user through the network. A low value means that the user is more likely to make information "move" around the network in the interval $(-t,0)$ than in the next interval.

We assign a reference value of 0.5 to $IA(-2t,-t)$ and to $NF(-2t,-t)$ and we scale the new values of of $IA(-t,0)$ and $NF(-t,0)$ as $0.5\frac{IA(-t,0)}{IA(-2t,t)}$ and $0.5\frac{NF(-2t,t)}{NF(-t,0)}$. IA and NF are the priors of our Bayesian network which also means that the posteriors defined below inherit the interval time.

The value for Information Gap (Definition 3) is:

$$IG = \frac{0.3NP_{\mathrm{C}} + 0.5NP_{\mathrm{AF}} + 0.8NP_{\mathrm{FofF}} + NP_{\mathrm{No}} + NP_{\mathrm{Ev}}}{\text{N.of posts}(-t,0)} \tag{3}$$

obtained as the weighted sum of the number of posts in time interval $(-t,0)$ with the following privacy settings: custom (C), all friends (AF), friends of friends (FoF), no privacy settings (No), and public posts open to everyone (Ev). In the present definition we have assigned a heavy weight of 0.8 to posts shared to friends of friends; a medium weight of 0.5 to posts shared to one's friend only; and a low weight of 0.3 to posts with the custom setting; the posts that have highest exposure get a full weight. The weights can be adjusted to reflect different intuitions concerning the exposure parameter. This sum is then divided by the total number of posts (included in IA) to obtain IG.

The value for Information Flow (Definition 4) is:

$$IF = \frac{Li + E + Sh + NP_{\mathrm{No}} + NP_{\mathrm{Ev}}}{NP_{\mathrm{Pr}} + NP_{\mathrm{C}} + NP_{\mathrm{AF}} + 0.5NP_{\mathrm{FofF}}} \tag{4}$$

defined as the ratio between the total number of likes (Li), Events (E), *shared* posts (Sh) and public posts (i.e. a portion of NF) over the total sum of posts with restricted privacy settings as defined above. IG and IF are the posterior probabilities of our Bayesian network.

6 Results and Evaluation

As concrete case studies, we have considered three Facebook users with different profiles over two 1-year intervals (i.e. it compares the difference in value between the activity over 1 year with the activity over the next year).

Case 1. A software developer using Facebook to interact with friends and family. The overall activity of the user has decreased due to an increased workload and to a house move. The total number of posts decreased from 31 to 16 and the number of likes decreased from 15 to 6. As expected, the value of privacy has increased from 1.00 to 1.32.

Case 2. An academic using Facebook mainly to communicate research, interact with students and with some family members. Also in this case, the number of posts decreased from 124 in the period 2013-14 to 94 in 2014-15. However, the number of friends posting on this person's wall increased from 12 to 30 and, moreover, the number of shares increased from 10 to 28. As a result, the privacy of the user decreased from 1 to 0.69 (see Figure 3). This case shows that, in our model for Facebook, social network privacy is affected not only by direct user's choices, but also by the behaviour of someone's friends.

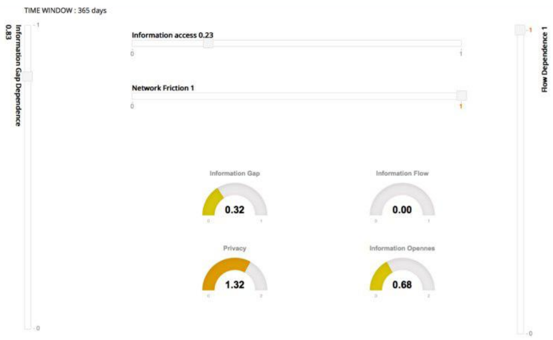
Case 3. A PhD student, using Facebook mainly to keep in touch with friends and family during her study periods abroad. For this user, in the given period, the total number of tagged places has increased from 36 to 78, and the total number of events from 50 to 57. As a result, the privacy of the user has decreased from 1 to 0.75 (see Figure 4). This case shows that, in our model, social network privacy takes into account also how much the user exposes information on her non-digital ontology, through the mention of places and events.

The data for the three users is summarized in Table 1.

	IA	IG	NF	IF	IP
User1	0.23	0.32	1.00	0.00	1.32 (+0.32)
User2	0.47	0.49	0.20	0.80	0.69 (-0.31)
User3	1.00	0.61	0.14	0.86	0.75 (-0.25)

Table 1. Comparison of the Profile Cases.

The current testing is only meant to give an intuitive idea of results for three different types of agents. They offer an indication of how the tool matches different behaviours, but it can be improved in several ways. The next stage


TIME WINDOW : 365 days

Information Gap Dependence 0.83

Information access 0.23

Network Friction 1

Flow Dependence 1

Information Gap 0.32

Information Flow 0.00

Privacy 1.32

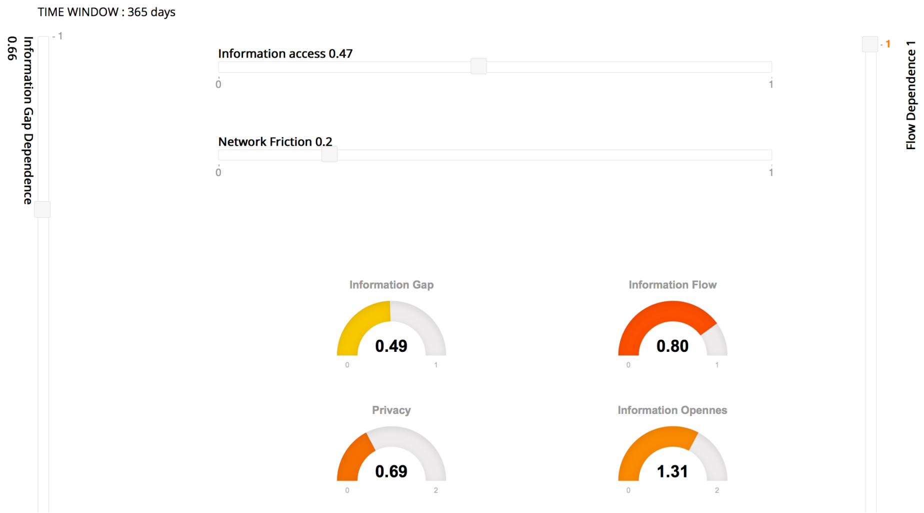
Information Opennes 0.68

Fig. 2. Facebook plug-in: results from Profile Case 1.

of this research requires a large-scale evaluation; the design needs to account for distinct a priori probabilities for groups or typologies of users and distinct dependencies between prior and posterior probabilities; the evaluation should consider average values over groups of users; and the test should include a more culturally diversified set of users.

7 Conclusions

The task of this paper has been threefold. First, to propose a modular, extensible and conceptually grounded theory of informational privacy for social networks, based on the approach of the philosophy of information. Second, to provide a



TIME WINDOW : 365 days

Information Gap Dependence 0.66

Information access 0.47

Network Friction 0.2

Flow Dependence 1

Information Gap 0.49

Information Flow 0.80

Privacy 0.69

Information Opennes 1.31

Fig. 3. Facebook plug-in: results for User 2.

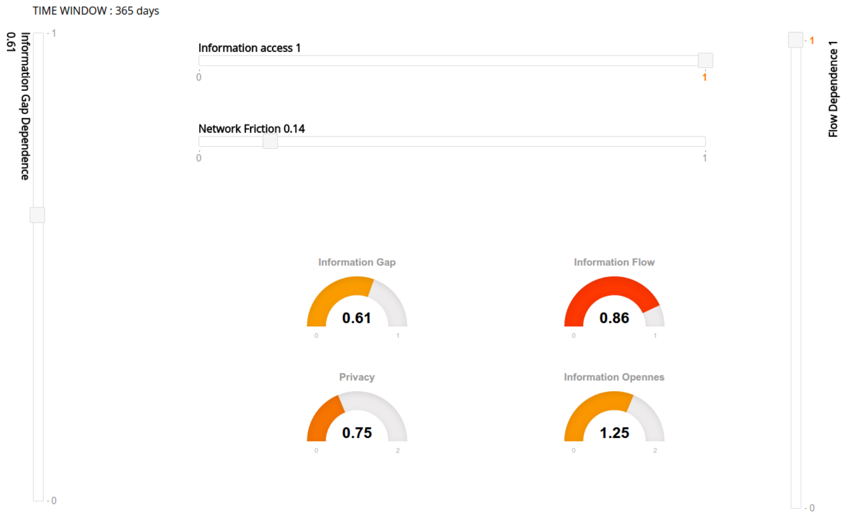
working Bayesian model of such theory in a Java algorithm. Third, to propose its implementation as a valid approach to the need of a diachronic modelling of privacy concerns, and in particular as a way to approach the Privacy Change Assessment Problem in a social network such as Facebook.

We forsee several obvious improvements to the tool. The diachronic approach to privacy assessment can be parametrised also in view of the frequent changes in privacy policies of social network sites: assuming one can quantify some prior probability value for the standard privacy setting of the network and given the possibility to change such value in view of the frequent corporate changes in policy, users can then calculate required changes in their own settings, reducing the risk of not acting in view of changes in policy. Such a possibility from service providers can strengthen trust in the network also in terms of ethical considerations. The tool itself can be developed with privacy-enhancing features, so as to make the user able to select relevant and appropriate metrics. The number of technical elements that can be codified in the Bayesian Network could be increased: currently an extension is being tested that includes an additional node for third party applications installed on the user's profile.

Finally, we aim at using this tool for a direct evaluation of user reputation in social context, where the ability to prevent (or, in some contexts, to facilitate) information leakage can be used to qualify the (in)suitability of users to informational privacy (or openness).

References

1. Rebecca Balebako and Lorrie Faith Cranor. Improving App Privacy: Nudging App Developers to Protect User Privacy. *IEEE Security & Privacy*, 12(4):55–58, 2014.



TIME WINDOW : 365 days

Information Gap Dependence 0.61

Information access 1

Network Friction 0.14

Flow Dependence 1

Fig. 4. Facebook plug-in: results for User 3.

2. Balbir S. Barn, Giuseppe Primiero, and Ravinder Barn. An Approach to Early Evaluation of Informational Privacy Requirements. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, SAC '15, pages 1370–1375, New York, NY, USA, 2015. ACM.
3. Jiyoung Cha. Factors affecting the frequency and amount of social networking site use: Motivations, perceptions, and privacy concerns. *First Monday*, 15(12), 2010.
4. Bernhard Debatin. Ethics, Privacy, and Self-Restraint in Social Networking. In Sabine Trepte and Leonard Reinecke, editors, *Privacy Online*, pages 47–60. Springer, 2011.
5. Facebook. Graph API https://developers.facebook.com/docs/graph-api.
6. Luciano Floridi. The ontological interpretation of informational privacy. *Ethics and Information Technology*, 7(4):185–200, 2005.
7. Luciano Floridi. Four challenges for a theory of informational privacy. *Ethics and Information technology*, 8(3):109–119, 2006.
8. Ralph Gross and Alessandro Acquisti. Information Revelation and Privacy in Online Social Networks. In *Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*, WPES '05, pages 71–80, New York, NY, USA, 2005. ACM.
9. S. Gurses and C. Diaz. Two tales of privacy in online social networks. *Security Privacy, IEEE*, 11(3):29–37, May 2013.
10. Seda F. Gürses, Ramzi Rizk, and Oliver Günther. Privacy Design in Online Social Networks: Learning from Privacy Breaches and Community Feedback. In *Proceedings of the International Conference on Information Systems, ICIS 2008, Paris, France, December 14-17, 2008*, page 90. Association for Information Systems, 2008.
11. S. Hazari and C. Brown. An Empirical Investigation of Privacy Awareness and Concerns on Social Networking Sites. *Journal of Information Privacy and Security*, 9(4):31–51, 2014.

12. H. Hedbom. A Survey on Transparency Tools for Enhancing Privacy. In Vashek Matyáš, Simone Fischer-Hübner, Daniel Cvrček, and Petr Švenda, editors, *The Future of Identity in the Information Society*, volume 298 of *IFIP Advances in Information and Communication Technology*, pages 67–82. Springer Berlin Heidelberg, 2009.
13. D. J. Houghton and A. N. Joinson. Privacy, Social Network Sites, and Social Relations. *Journal of Technology in Human Services*, 28(1-2):74–94, 2010.
14. Ulrike Hugl. Reviewing person's value of privacy of online social networking. *Internet Research*, 21(4):384–407, 2011.
15. D. Irani, S. Webb, C. Pu, and Kang Li. Modeling Unintended Personal-Information Leakage from Multiple Online Social Networks. *Internet Computing, IEEE*, 15(3):13–19, May 2011.
16. Finn V Jensen. *An introduction to Bayesian networks*, volume 210. UCL press London, 1996.
17. Y Jeonga and E. Coylea. What Are You Worrying About on Facebook and Twitter? An Empirical Investigation of Young Social Network Site Users' Privacy Perceptions and Behaviors. *Journal of Interactive Advertising*, 14(2):51–59, 2014.
18. Ted Kang and Lalana Kagal. Enabling Privacy-Awareness in Social Networks. In *Intelligent Information Privacy Management, Papers from the 2010 AAAI Spring Symposium, Technical Report SS-10-05, Stanford, California, USA, March 22-24, 2010*. AAAI, 2010.
19. Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh. Privacy As Part of the App Decision-making Process. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 3393–3402, New York, NY, USA, 2013. ACM.
20. Balachander Krishnamurthy and Craig E. Wills. Characterizing Privacy in Online Social Networks. In *Proceedings of the First Workshop on Online Social Networks*, WOSN '08, pages 37–42, New York, NY, USA, 2008. ACM.
21. Yan Li, Yingjiu Li, Qiang Yan, and Robert H. Deng. Privacy Leakage Analysis in Online Social Networks. *Comput. Secur.*, 49(C):239–254, March 2015.
22. Yuan Li. Theories in Online Information Privacy Research: A Critical Review and an Integrated Framework. *Decis. Support Syst.*, 54(1):471–481, December 2012.
23. S. Mahmood. Online Social Networks: Privacy Threats and Defenses. In Richard Chbeir and Bechara Al Bouna, editors, *Security and Privacy Preserving in Social Networks*, Lecture Notes in Social Networks, pages 47–71. Springer Vienna, 2013.
24. Software Corporation Norsys. Netica Java API https://www.norsys.com/netica-j.
25. Inah Omoronyia, Luca Cavallaro, Mazeiar Salehie, Liliana Pasquale, and Bashar Nuseibeh. Engineering Adaptive Privacy: On the Role of Privacy Awareness Requirements. In *Proceedings of the 2013 International Conference on Software Engineering*, ICSE '13, pages 632–641, Piscataway, NJ, USA, 2013. IEEE Press.
26. Dong-Hee Shin. The effects of trust, security and privacy in social networking: A security-based approach to understand the pattern of adoption. *Interacting with Computers*, 22(5):428–438, 2010. Modelling user experience - An agenda for research and practice.
27. Statista. Number of monthly active facebook users worldwide as of 1st quarter 2016 (in millions). http://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/, 2016.
28. Y. Takano, S. Ohta, T. Takahashi, R. Ando, and T. Inoue. MindYourPrivacy: Design and implementation of a visualization system for third-party Web tracking. In *Privacy, Security and Trust (PST), 2014 Twelfth Annual International Conference on*, pages 48–56, July 2014.

29. Xin Tan, Li Qin, Yongbeom Kim, and Jeffrey Hsu. Impact of privacy concern in social networking web sites. *Internet Research*, 22(2):211–233, 2012.
30. Carmela Troncoso. Bayesian inference to evaluate information leakage in complex scenarios. In William Puech, Marc Chaumont, Jana Dittmann, and Patrizio Campisi, editors, *ACM Information Hiding and Multimedia Security Workshop, IH&MMSec '13, Montpellier, France, June 17-19, 2013*, pages 1–2. ACM, 2013.
31. Wanhong Xu, Xi Zhou, and Lei Li. Inferring privacy information via social relations. In *Data Engineering Workshop, 2008. ICDEW 2008. IEEE 24th International Conference on*, pages 525–530, April 2008.

APPENDIX 1

```java
import norsys.netica.*;
import norsys.neticaEx.aliases.Node;

public class BuildNet {
//main method with constructors
  public static void main (String[] args){
    try {
  Node.setConstructorClass
      ("norsys.neticaEx.aliases.Node");
  Environ env = new Environ (null);
  Net net = new Net();
  net.setName("Informational_Privacy");

//Nodes of the Informational Privacy Network
  Node InfoAccess = new Node
  ("InfoAccess", "lowAccess,highAccess",net);
  Node InfoGap = new Node
  ("InfoGap", "Gap_Present,Gap_Absent",net);
  Node NetworkFriction = new Node
  ("NetworkFric", "Friction,Lubricated",net);
  Node InfoFlow = new Node
  ("InformationFlow", "absent,present",net);
  Node InfoPriv = new Node
  ("InfoPriv", "absent,present",net);

  InfoAccess.setTitle ("Information_Access");
  InfoGap.setTitle ("Information_Gap");
  NetworkFriction.setTitle ("Network_Friction");
  InfoFlow.setTitle ("Information_Flow");
  InfoPriv.setTitle ("Informational_Privacy");

  //Dependencies between the nodes
  InfoGap.addLink (InfoAccess);
  // link from Information_Access to Information_Gap
  InfoFlow.addLink (NetworkFriction);
  // link from Network_Friction to Information_Flow
  InfoPriv.addLink (InfoGap);
  // link from Information_Gap to
  //Informational_Privacy
  InfoPriv.addLink (InfoFlow);
  // link from Information_Flow to
  //Informational_Privacy

  // Define the first prior probability:
  //Access is calculated in terms of
  // availability of the technology
  InfoAccess.setCPTable(0.50,0.50);
```

```

//Define the first dependent probability:
//InfoGap is dependent from Access
           //InfoAccess  low  high
InfoGap.setCPTable ("lowAccess", 0.90, 0.10);
InfoGap.setCPTable ("highAccess", 0.10, 0.90);

// Define the second prior probability:
//Network Friction is calculated in terms of
//proportion of common nodes
NetworkFriction.setCPTable (0.50, 0.50);

//Define the second dependent probability:
//Information Flow
//dependent from Network Friction
       //Friction Frictioned Lubricated
InfoFlow.setCPTable ("Friction", 0.90, 0.10);
InfoFlow.setCPTable ("Lubricated", 0.10, 0.90);

InfoPriv.setEquation
("InfoPriv
(InfoGap,InformationFlow) = absent||present");
InfoPriv.equationToTable (1, false, false);
net.compile();

//set the value of the first prior positive
float open = InfoGap.getBelief("Gap_Present");
System.out.println
("\nThe positive value of the
information gap is " + open );

//set the value of the first prior negative
float close = InfoGap.getBelief ("Gap_Absent");
System.out.println
("\nThe negative value of the
information gap is " + close);

//set the value of the second prior positive
float easy = InfoFlow.getBelief ("present");
System.out.println
("\nThe positive value of the
information flow is " + easy);

//set the value of the second prior negative
float difficult = InfoFlow.getBelief ("absent");
System.out.println
("\nThe negative value of the
information flow is " + difficult);

//get the value of the result
```

```
  //with gap and no flow
  System.out.println
  ("\nGiven the max of infogap
  and the min of flow,\n"+
  "the probability of informational privacy is "
  + (open + difficult));

  //get the value of the result
  //with no gap and flow
  System.out.println
  ("\nGiven the min of infogap
  and the max of flow,\n"+
  "the probability of informational openness is "
  + (close + easy));

  net.finalize();
    }
    catch (Exception e) {
  e.printStackTrace();
  }
}
```