Probabilistically Assessing Informational Privacy Awareness in Social Networks

Giuseppe Primiero, Franco Raimondi, Balbir Barn
Department of Computer Science
Middlesex University, United Kingdom
E-mail:G.Primiero—F.Raimondi—B.Barn@mdx.ac.uk

***Abstract*—User privacy in social networks represents one of the greatest challenges of the information age. One important aspect is to make users aware of freely shared data exposure on a network and, accordingly, to establish information openness to non-linked nodes based on user's behaviour. In this paper we formulate the above aspect as a well-defined problem and propose a first proof-of-concept model and tool for informational privacy assessment by users of a social network. The model is built on a Theory of Informational Privacy where privacy requirements are evaluated by probabilities in a Bayesian network. We propose some examples of how the framework could be used along with a number of required extensions.**

I. Introduction

Online social network sites represent a recent object of academic studies (see e.g. [BE07], [Mah13]). Among the various different aspects of interest, they are analysed as one of the largest threat to personal privacy in the digital age (for a survey on transparency tools, see [Hed09]). Freely shared information can be used to infer further (supposedly secure) data, aiming at re-identification for commercial and even criminal purposes. This has triggered a large amount of studies in the last few years, with attention to potential attacks to privacy based on voluntary information disclosure [GA05], [IWPL11], on incomplete personal information [XZL08], control on web-tracking activities [TOT$^+$14], in relation to trust [DHP07], just to name a few. Alongside with scientific research, a large amount of work is being invested in policy-making.

One aspect of this research area concerns methodologies and practices to define, monitor and rise awareness of social networks users' privacy, a topic that ranges from theoretical and socio-demographic considerations [HJ10], [HB14], [JC14] and design of access control [KK10], to requirements elicitation [OCS$^+$13]. For an overview of the literature see [Hug11]. The notion of privacy awareness and risk assessment is available in the literature as privacy concern, see [TQKH12] for an overview of the existing literature. A missing feature of these analyses, though, is a *diachronic* analysis of the privacy concern model: how does perceived privacy *change* over a given span of time in relation to (possibly non-linked) nodes of the same network? And how can evaluation of such change be monitored so as to offer a more dynamic view on the privacy concern?

Such evaluation of one's privacy can be formulated with respect to quantitative parameters, such as the amount of information content shared, the degree of distance to any given non-linked node and the likability that a given content might appeal to such a non-connected user. The problem can be formulated explicitly and more precisely as follows:

Definition 1 (The Privacy Change Assessment Problem)**.** *Given a fixed span of time t, two users U_1, U_2 separated by a degree n of intermediate users, and a measurable amount of network activity of U_1 over t, how does the probability change over t that some piece of information i shared by U_1 will eventually reach U_2 through the shared network?*

This specific formulation of information privacy for social networks does not seem to have been tackled yet in the literature. Its analysis offers the advantage of providing a measurable criterion of *privacy change*, with the aim of making users aware of how their behaviour in the network increases the possibility of information leakage. On such basis, network design and policy making issues can also be further addressed. In the present paper, we offer a proof-of-concept working model to approach this problem.

In previous work [BPB15], we have provided a model of informational privacy based on [Flo05] to a representation suitable for consumption by software engineers to evaluate information privacy concerns in the design process. The language is used to inform the development of a Bayesian network which we use to encode aspects of the informational privacy theory. In the present work, we adapt the informational privacy model to the case of social networks and use an adapted Bayesian encoding to perform on-the-fly evaluations of user's privacy relative to a non-connected link. The model is currently at design stage and a large scale testing is yet to be performed. Nonetheless, we believe it is a novel and significant approach to user privacy evaluation and its implementation may be very useful in rising users' awareness concerns in their information sharing processes.

The reminder of this paper is structured as follows. In Section II we revise the basics of Floridi's Ontological Theory of Informational Privacy and adapt its notions to the context of social networks with digital information. In Section III we propose an implementation of the theory through a Bayesian Network in Java. In Section IV, we consider a hypothetical

study case, where the probabilistic value of information leakage is assessed by a user at different stages, so as to reflect changes in her behaviour over the network. In Section V we briefly consider the required improvement and extensions to both theory and implementation expected in the future, and in Section VI we reconsider our contribution in view of the current approaches to privacy in social networks.

II. A THEORY OF INFORMATIONAL PRIVACY FOR SOCIAL NETWORKS

User's privacy can be defined on a set of primitive notions. Such primitives can differ based on the environment of reference: for example, while privacy in the real world has been standardly defined in relation to other people's behaviour or public attention, privacy for the digital world has been mainly referred to data protection in the use of digital media. In recent years, the methodological approach of the Philosophy of Information has suggested that a common framework can be designed for users in the digital age. Accordingly, a theory of informational privacy for digital environments can be designed taking inspiration from a theory of privacy of information packets in the infosphere at large. Such is Floridi's theory of informational privacy, presented in [Flo05] and then briefly re-presented in [Flo06]. According to Floridi, a theory of informational privacy can be construed around four basic notions: ontological features such as agents and the environment in which they interact determine the accessibility of information between agents; this determines the informational gap among agents in the infosphere, such that the larger the information gap between agents the less they know about each other; the world establishes a degree of ontological friction, and the latter regulates the information flow within the system. A minimal axiomatic presentation of the theory can be given as follows, with pairs of agents $(A, B), (C, D)$ embedded in a portion W of the infosphere and relational symbols $>, <$ standing for the obvious reflexive (total) order:

Axiom 1. $InfoGap(A, B) > InfoGap(C, D) \rightarrow InfoPrivacy(A, B) > InfoPrivacy(C, D)$

Axiom 2. $InfoAccess(A, B) > InfoAccess(C, D) \rightarrow InfoGap(A, B) < InfoGap(C, D)$

Axiom 3. $OntoFriction(InfoFlow(W[A, B])) > OntoFriction(InfoFlow(W[C, D])) \rightarrow InfoAccess(A, B) < InfoAccess(C, D)$

Informational privacy then becomes a function of both the information gap between agents and the information flow in the infosphere (see Figure 1 for a UML-style presentation).

In the present context, our aim is to redefine the basic concepts of Floridi's theory for the specific context of information sharing processes in the social networking environment. To this aim, we propose the following definitions:

Definition 2 (Information Access)**.** *A measure of an agent's openness to social networking activities.*

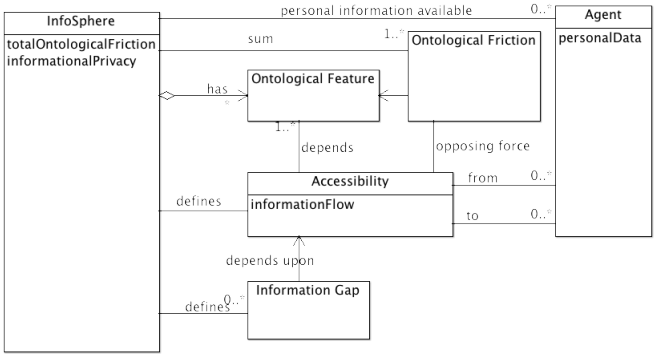

Fig. 1. Floridi's Theory of Ontological Informational Privacy

This value can be calculated on a time scale, or on the number of accesses in a given period of time. The use of such a notion in terms of time actively spent on a social networking site can be useful in countering the findings in [Cha10] that see privacy concerns to be negatively correlated with users frequency, but not with usage time.

Definition 3 (Network Friction)**.** *A measure of an agent's vicinity to another agent, in terms of number of intermediate links or proportion of common nodes.*

This value is used to evaluate the proximity of two users not directly connected to each other. The rationale behind this definition can be specified in different ways: one way is to look at the minimal degree of separation between two users and attribute higher friction to a longer chain, more lubrication to shorter ones; another way is to consider the proportion of connections of a user that are shared with another user, and attribute more friction to a lower proportion, and more lubrication to a higher one. Notice that this value can be further qualified: one can consider only nodes of first degree that are common between the two agents, or for larger and more complex networks, one might be interested in connections of higher degree. This definition can thus be further refined in view of different contexts or uses. We consider the two concepts above as fundamentals; in the modelling offered in the next Section III, these are defined as prior probabilities and used to infer the probabilistic values of two further notions:

Definition 4 (Information Gap)**.** *A function of the degree of accessibility of personal data (depending on informational access).*

This value is used to compute the exposure that personal information of the user receives on the network, and it is directly dependent from the value of access. It can be understood as the probability that a certain gap towards a piece of data be bridged, given the behaviour on the network of the user owning that data.

Definition 5 (Information Flow)**.** *A function of the degree of fluidity of personal data (depending on network friction).*

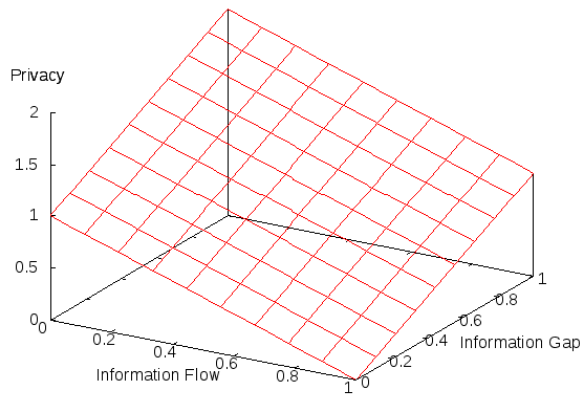

Fig. 2. Privacy as a function of Information Gap and Information Flow

This value is used to compute how easy is for data to transfer from User 1 to User 2 over the network, and it is directly dependent on the lubrication value of the network. In our model, an agent that tends to have high (resp. low) availability of technology and activity on the network has *high information access* (resp. low information access). This value determines an *absence of information gap*, to express the fact that personal data of such a user is reachable (resp. presence of information gap to express that data is not reachable). An agent with a low (resp. high) relative distance from another agent (or proportion of nodes in common with another agent, depending on the preferred interpretation) will display a lubricated (resp. frictional) network, inducing presence (resp. absence) of Information Flow. The two values of Information Gap and Information Flow are finally used to assess Information Privacy.

These four item can be plotted on a graph in order to visualize their relation. Figure 2 provides an intuition for the relationship between the three values Information Gap, Information Flow, and Privacy

III. Bayesian Network implementation

A Bayesian network (BN) is a compact representation of probabilistic relationships between properties in a domain. The BN is constructed as a directed acyclic graph where nodes represent properties that can take values depending upon a probabilistic correlation between properties (the arc between two nodes) [Jen96]. A BN also describes a probability distribution over all variables in the BN 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. We offer in the following an implementation of the theory of Informational Privacy for Social Networks presented in section II in the Java API for the Netica Software, https://www.norsys.com/netica-j.html. 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. The code is also available for download at https://github.com/gprimiero/InfoPriv

```
1  import norsys.netica.*;
2  import norsys.neticaEx.aliases.Node;
3
4  public class BuildNet {
5  //main method with constructors
6   public static void main (String[] args){
7     try {
8   Node.setConstructorClass
9       ("norsys.neticaEx.aliases.Node");
10  Environ env = new Environ (null);
11  Net net = new Net();
12  net.setName("Informational_Privacy");
13
14 //Nodes of the Informational Privacy Network
15  Node InfoAccess = new Node
16  ("InfoAccess", "lowAccess,highAccess",net);
17  Node InfoGap = new Node
18  ("InfoGap", "Gap_Present,Gap_Absent",net);
19  Node NetworkFriction = new Node
20  ("NetworkFric", "Friction,Lubricated",net);
21  Node InfoFlow = new Node
22  ("InformationFlow", "absent,present",net);
23  Node InfoPriv = new Node
24  ("InfoPriv", "absent,present",net);
25
26  InfoAccess.setTitle ("Information_Access");
27  InfoGap.setTitle ("Information_Gap");
28  NetworkFriction.setTitle ("Network_Friction");
29  InfoFlow.setTitle ("Information_Flow");
30  InfoPriv.setTitle ("Informational_Privacy");
31
32  //Dependencies between the nodes
33  InfoGap.addLink (InfoAccess);
34  // link from Information_Access to
        Information_Gap
35  InfoFlow.addLink (NetworkFriction);
36  // link from Network_Friction to
        Information_Flow
37  InfoPriv.addLink (InfoGap);
38  // link from Information_Gap to
39  //Informational_Privacy
40  InfoPriv.addLink (InfoFlow);
41  // link from Information_Flow to
42  //Informational_Privacy
43
44  // Define the first prior probability:
45  //Access is calculated in terms of
46  // availability of the technology
47  InfoAccess.setCPTable(0.50,0.50);
48
49  //Define the first dependent probability:
50  //InfoGap is dependent from Access
51           //InfoAccess low high
52  InfoGap.setCPTable ("lowAccess", 0.90, 0.10);
53  InfoGap.setCPTable ("highAccess", 0.10, 0.90);
54
55  // Define the second prior probability:
56  //Network Friction is calculated in terms of
57  //proportion of common nodes
58  NetworkFriction.setCPTable (0.50, 0.50);
59
60  //Define the second dependent probability:
61  //Information Flow
62  //dependent from Network Friction
63        //Friction Frictioned Lubricated
64  InfoFlow.setCPTable ("Friction", 0.90, 0.10);
65  InfoFlow.setCPTable ("Lubricated", 0.10, 0.90);
66
67  InfoPriv.setEquation
```

```
68    ("InfoPriv
69    (InfoGap,InformationFlow) = absent||present");
70    InfoPriv.equationToTable (1, false, false);
71    net.compile();
72
73    //set the value of the first prior positive
74    float open = InfoGap.getBelief("Gap_Present");
75    System.out.println
76    ("\nThe positive value of the
77    information gap is " + open );
78
79    //set the value of the first prior negative
80    float close = InfoGap.getBelief ("Gap_Absent");
81    System.out.println
82    ("\nThe negative value of the
83    information gap is " + close);
84
85    //set the value of the second prior positive
86    float easy = InfoFlow.getBelief ("present");
87    System.out.println
88    ("\nThe positive value of the
89    information flow is " + easy);
90
91    //set the value of the second prior negative
92    float difficult = InfoFlow.getBelief
          ("absent");
93    System.out.println
94    ("\nThe negative value of the
95    information flow is " + difficult);
96
97    //get the value of the result
98    //with gap and no flow
99    System.out.println
100   ("\nGiven the max of infogap
101   and the min of flow,\n"+
102   "the probability of informational privacy is "
103   + (open + difficult));
104
105   //get the value of the result
106   //with no gap and flow
107   System.out.println
108   ("\nGiven the min of infogap
109   and the max of flow,\n"+
110   "the probability of informational openness is "
111   + (close + easy));
112
113   net.finalize();
114     }
115     catch (Exception e) {
116   e.printStackTrace();
117   }
118 }
```

The implementation of the above model of informational privacy as BN contains nodes according to the Definitions from Section II (lines 14-31): Information Access and Network Friction are parent nodes with prior probabilities; they determine respectively Information Gap (line 33) and Information Flow (line 35); 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 the present initial formulation, all values are balanced: a $50 - 50\%$ value on low/high Information Access reflects a moderate use of technology and network access (line 47); a $50 - 50\%$ value on friction/lubrication of the Network reflects a medium chain length (or range of common nodes, depending from the preferred interpretation) between User 1 and User 2 (line 58). The Information Gap is the first conditional probability, based on the Information Access value (lines 52-53), and Information Flow the second one, based on Friction (lines 64-65). Their values are associated almost one-to-one with their prior probabilities. We now wish to calculate the value of information privacy (IP, lines 67-71) and on its basis that of its dual, information openness (IO). The former is computed as

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

the latter as

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

These will thus be evaluated on a scale $0.00 - 2.00$ as the sum of the two $0 - 100\%$ scales for Information Gap and Information Flow respectively (see again Figure 2). We assign variable names to the evaluation range of our conditional probabilities (lines 73-95): we call *open* the value associated with the presence of information gap, and *close* the value associated with its absence; we call *easy* the value associated with the presence of information flow, and *difficult* 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).

IV. Examples

Our tool is meant for privacy assessment, hence we envisage its use at regular, possibly automated intervals, to reflect changes in network use. The aim is to provide a monitoring of how network's use influences the privacy of the user, intended as the probability that personal data will reach a second user not directly connected over the same network.

Let us assume we start monitoring a user at time t_1 with a moderate use of the network (e.g. in terms of activity), and a relatively long distance from a non-connected user, or a small number of common links. This means that her Information Access value is geared towards "low access" and her Network Friction towards a low level of lubrication:

```
1              //InfoAccess low high
2  InfoAccess.setCPTable(0.80,0.20);
3                //Lubrication low high
4  NetworkFriction.setCPTable (0.80, 0.20);
```

We currently maintain the conditional probability relation of Information gap and Information Flow fixed at the (almost) 1-1 proportion with the corresponding prior.

```
1            //InfoAccess low high
2    InfoGap.setCPTable ("lowAccess", 0.90, 0.10);
3    InfoGap.setCPTable ("highAccess", 0.10, 0.90);
4
5          //Friction Frictioned Lubricated
6    InfoFlow.setCPTable ("Friction", 0.90, 0.10);
7    InfoFlow.setCPTable ("Lubricated", 0.10, 0.90);
```

The result of compiling the code above with these probabilities returns the following:

```
1  The positive value of the information gap is
       0.73999995
2  The negative value of the information gap is
       0.26000002
```

```
The positive value of the information flow is
    0.26000002
The negative value of the information flow is
    0.73999995

Given the max of infogap and the min of
    interaction,
the probability of informational privacy is
    1.4799999

Given the min of infogap and the max of
    interaction,
the probability of informational openness is
    0.52000004
```

With "positive value of Information Gap" we refer to the presence of Gap, and to its absence with "negative value of Information Gap"; hence, in this instance, the user is little exposed to access of her personal data from the outside, as a low negative value reflects a high level of gap. With "positive value of Information Flow" we refer to the presence of Flow (hence lubrication), and to its absence with "negative value of Information Flow"; hence, in this instance, the user is acting in a network that shows high friction. The result of the computation reflects a higher value of privacy than openness, as predicted by the cautious values in access and the high level of friction (low lubrication) in the Network.

Let us now assume a second monitoring instance at time $t_{m>1}$. In this interval, our user has started changing her behaviour on the network, with a more active and constant use and a heavy increase in the number of linked users, which in turn means a greater vicinity to a non-linked user. Accordingly, the values of the prior probabilities could be changed as follows (we keep the conditional dependency of the other nodes fixed):

```
          //Access       low  high
InfoAccess.setCPTable(0.30,0.70);
           //Lubrication      low  high
NetworkFriction.setCPTable (0.30, 0.70);
```

The compilation results are as follows:

```

The positive value of the information gap is
    0.34
The negative value of the information gap is
    0.66
The positive value of the information flow is
    0.65999997
The negative value of the information flow is
    0.34

Given the max of infogap and the min of
    interaction,
the probability of informational privacy is 0.68

Given the min of infogap and the max of
    interaction,
the probability of informational openness is
    1.3199999
```

As it is to be expected, the value of IP decreases, while that of OP rises. A further evolution of the model is obtained by modifying the dependent probabilities values, i.e. how Information Gap and Information Flow actually depend from respectively Access and Network Friction. In our next example we are going to make some more realistic assumptions for a third assessment at time $t_{n>m>1}$:

1) the value of Information Gap, i.e. the level of accessibility of personal data, is only partially dependent from the access of the network by the user who owns the data, as it might for example also be determined by factors such as metadata availability that make them reachable and their overlap in general search interest (e.g. by the use of common hash tags); accordingly we modify values of CPTs for Information Gap as follows:

```
InfoGap.setCPTable("lowAccess", 0.60,0.40);
InfoGap.setCPTable("highAccess",
    0.40,0.60);
```

2) the value of Information Flow, i.e. the level of lubrication of the network between two users, is only partially dependent from the length of the connection chain or the number of their common links, as it might also be determined by other factors, such as activity of the intermediate links on those data and shared interests of a sufficient number of users over that data; accordingly we modify values of CPTs for Information Flow as follows:

```
InfoFlow.setCPTable("Friction", 0.70,0.30);
InfoGap.setCPTable("Lubricated",
    0.30,0.70);
```

The compilation results are as follows:

```
The positive value of the information gap is
    0.45999998
The negative value of the information gap is
    0.54
The positive value of the information flow is
    0.58
The negative value of the information flow is
    0.42000002

Given the max of infogap and the min of
    interaction,
the probability of informational privacy is 0.88

Given the min of infogap and the max of
    interaction,
the probability of informational openness is
    1.12
```

The changes appear both in terms of evaluation of the conditional probabilities of Information Gap and Information Flow, and the resulting values on Information Privacy and Information Openness: as expected, the value of privacy has increased from the previous evaluation, as we are now taking into account the possibility that other factors be considered relevant to the combined result.

In this small experiment, we are simulating a change in the behaviour of our user at different points in time from a very cautious to a more open attitude towards information sharing. Such change is reflected in a decrease of Information Privacy value from 1.4799999 to 0.68; and an increase in the probability that a piece of data will reach a non-connected user from 0.52000004 to 1.3199999. As soon as the CPTs values are modularised in view of different parameters, the values change accordingly as in the third example above.

V. Future Work

The probabilistic framework presented in Section III still presents many areas of further development. In particular, we consider the following ones essential to a next stage of analysis:

1) *Network Design*: we have construed our BN around four nodes, each expressing one of the main notions of Floridi's theory of Informational Privacy. Although these represent relevant concepts in the design of a theory of privacy for social networks, we have highlighted how the conditional probability of the notions of Information Gap and Information Flow can be calculated in a more refined way. For example, the value of Information Gap, which expresses data accessibility, should take into account the specifics of the networks, such as metadata and link constructions in terms of weak and strong ties;
2) *GUI*: we have been working only on the underlying probabilistic evaluation of privacy awareness and ignored entirely the user interface aspect of the framework. This is certainly a crucial aspect for the successful deployment of the framework. The Netica software has a graphical user interface available, but a working integration with a social network interface should be sought also in view of user design and ease of probabilistic values interpretation;
3) *Testing*: the current presentation relies entirely on a proof-of-concept and is exemplified through a simulated experiment. More structured and systematic testing should be performed, with a significant number of users and real information sharing scenarios.

VI. Tackling the Privacy Problem

Our framework constitutes only one of the many different approaches that are being explored in academia, industry and among policy makers to resolve the privacy problem in the increasingly influential context of social networks. We have already mentioned, among others, the access control approach to privacy in social networks which seeks to block outwards viewers that do not have explicit access to content, e.g. through encryption [BBS+09]. The framework proposed in [KK10] complements access control methods by ensuring that people know exactly what they can and cannot do with personal information. None of these approaches covers the specific formulation offered by the Privacy Change Assessment Problem from Section I. This formulation reflects a different aspect of the issue of privacy for large networks, with a number of realistic assumptions:

- The privacy problem for social networks needs to be coordinated with an awareness problem: constraints on access or sharing practices are going to be effective only in so far as the user is made aware of how her daily practices and use of the network influence her exposure to data leakage and data reconstruction.
- The privacy problem is a diachronic one: privacy levels change across time in view of changing sharing practices and network conditions; hence a single, rigid policy is most likely to result ineffective, while a continuous monitoring is more likely to be successful.
- The privacy problem should take into account not only the nodes directly linked to the user, but also her reach to indirectly linked ones. It is in fact more relevant to data leakage measurement to forecast exposure to non-directly linked users.

Our tool for probabilistic assessment of informational privacy combines all of the above properties. In particular, while large part of current literature focuses on how privacy concerns are related to breach of trust by third parties or trusted users (see e.g. [HJ10]), the present approach strictly considers privacy changes in view of the user own activities, combined with the structure of the network. Previous studies (e.g. [TQKH12]) have shown how the privacy concerns, albeit significant, do not directly affect users' acceptance of social networking web sites, while they do moderate the effects of perceived usefulness, and perceived ease of use, on users' intention to continue to use the service. We believe that such restricted influence of privacy concerns in the acceptance of services and the modification of behaviour in using them should be seriously considered and can be alleviated by directly making users aware of privacy threats, i.e. allowing them to understand consequences of such usage. That privacy and security perception influence users' trust, attitude and usage intention is also confirmed by other recent empirical studies [Shi10], [HB14].

An additional advantage of the diachronic approach is that it could 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.

VII. 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 working modelling of such theory in the format of a Java implementation. Third, to propose a proof-of-concept application 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. We have corroborated this proposal with a simulated test and highlighted its current limitations, both in terms of developing and testing. Future work shall focus on addressing such shortcomings and deliver a full-range empirical study.

References

[BBS+09] Randy Baden, Adam Bender, Neil Spring, Bobby Bhattacharjee, and Daniel Starin. Persona: An online social network with user-defined privacy. *SIGCOMM Comput. Commun. Rev.*, 39(4):135–146, August 2009.

[BE07] Danah M. Boyd and Nicole B. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, 2007.

[BPB15] B. Barn, G. Primiero, and R. Barn. An approach to early evaluation of informational privacy requirements. In *ACM SAC15*, 2015.

[Cha10] Jiyoung Cha. Factors affecting the frequency and amount of social networking site use: Motivations, perceptions, and privacy concerns. *First Monday*, 15(12), 2010.

[DHP07] Catherine Dwyer, Starr Roxanne Hiltz, and Katia Passerini. Trust and privacy concern within social networking sites: A comparison of facebook and myspace. In *AMCIS 2007 Proceedings. Paper 339.*, 2007.

[Flo05] Luciano Floridi. The ontological interpretation of informational privacy. *Ethics and Information Technology*, 7(4):185–200, 2005.

[Flo06] Luciano Floridi. Four challenges for a theory of informational privacy. *Ethics and Information technology*, 8(3):109–119, 2006.

[GA05] 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.

[HB14] 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.

[Hed09] H. Hedbom. A survey on transparency tools for enhancing privacy. In Vashek Maty, Simone Fischer-Hbner, Daniel Cvrek, and Petr venda, 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.

[HJ10] 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.

[Hug11] Ulrike Hugl. Reviewing person's value of privacy of online social networking. *Internet Research*, 21(4):384–407, 2011.

[IWPL11] 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.

[JC14] 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.

[Jen96] Finn V Jensen. *An introduction to Bayesian networks*, volume 210. UCL press London, 1996.

[KK10] 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.

[Mah13] 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.

[OCS+13] 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.

[Shi10] 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.

[TOT+14] 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.

[TQKH12] 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.

[XZL08] 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.