Towards Inferring Communication Patterns in Online Social Networks

Ero Balsa[1], Cristina Pérez-Solà[2], and Claudia Diaz[1]

[1]KU Leuven ESAT/COSIC & iMinds, Leuven, Belgium
[2]dEIC, Universitat Autònoma de Barcelona, Catalonia &
KU Leuven ESAT/COSIC, Leuven, Belgium

October 30, 2017*

Abstract

The separation between the public and private spheres on online social networks is known to be at best blurred. On the one hand, previous studies have shown how it is possible to *infer* private attributes from publicly available data. On the other hand, no distinction exists between public and private data when we consider the ability of the OSN provider to access them. Even when OSN users go to great lengths to protect their privacy, such as by using encryption or communication obfuscation, correlations between data may render these solutions useless. In this paper, we study the relationship between private communication patterns and publicly available OSN data. Such relationship informs both privacy-invasive inferences as well as OSN communication modelling, the latter being key towards developing effective obfuscation tools. We propose an inference model based on Bayesian analysis and evaluate, using a real social network dataset, how archetypal social graph features can lead to inferences about private communication. Our results indicate that both friendship graph and public traffic data may not be informative enough to enable these inferences, with time analysis having a non-negligible impact on their precision.

1 Introduction

Privacy breaches in online social networks (OSNs) have become commonplace. Context collision Danah Boyd (2008), unexpected or regrettable disclosures Wang et al. (2011), cyberstalking Gross and Acquisti (2005), blackmailing Gross and Acquisti (2005) or law enforcement

*This is an author generated postprint of the article: Balsa, E., Pérez-Solà, C., & Diaz, C. (2017). Towards inferring communication patterns in Online Social Networks. *ACM Transactions on Internet Technology (TOIT)*, 17(3), 32. The final publication is available on `https://dl.acm.org/citation.cfm?id=3093897`.

prowling Marks (2006) are only a few of the privacy threats users may face in OSNs. *Inferences* are a particular type of privacy threat that rely on correlations between data to learn private, non available data from publicly available data Heatherly et al. (2013); Zheleva and Getoor (2009). Information publicly available on social networks enables inferences about data that users have not publicly shared, such as their sexual orientations Jernigan and Mistree (2009), their political views, or the use of illegal substances Kosinski et al. (2013), among other types of attributes.

In order to shield themselves from these privacy threats, OSN users may use a range of privacy preserving tools and strategies, e.g., they may plainly refrain from uploading certain information to the site or deliberately lie in order to blur certain details of their persona. They may use *privacy settings* that enable them to control the visibility of their data, choosing between *posting* information *publicly*, or communicating *privately* with a small subset of their friends. Privacy settings effectively allow users to limit the amount of information publicly available; yet the service provider is still able to access privately shared information. More stringent cryptographic access control tools such as Scramble Beato et al. (2011), can effectively prevent the service provider from accessing the OSN users' private information. Still, encryption does not conceal traffic data. Even when all data and communications are encrypted, the service provider is still able to monitor the users' communication patterns, namely, who the users communicate with, how much, and how often, as well as other activities performed by the users on the site.

Communication patterns potentially reveal who users are most intimate with, their affinity in terms of age, religion and kinship, or their political views, among other attributes. They may also expose the strength of the users' relationships and how they evolve with time. In particular, users may choose to favour private messaging over public communication (i.e., *posts*) to hide discreditable or sensitive information. Examples include two people having a romantic affair, community leaders secretly organising an event, hiding from the broader public who the leaders and their social circle are in the community, or users reaching out to others for help with sensitive issues (e.g., bullying, medical advise).

Yet *hiding* communication patterns in the same way that encryption hides messages is impossible, and alternative strategies must be devised, such as *obfuscation*. Obfuscation tools send *dummy traffic* on behalf of the users to conceal their communication patterns: an eavesdropper, such as the OSN provider, observes a mix of real and dummy traffic; and is as a result no longer able to retrieve an accurate representation of the users' real communication patterns.

For dummy traffic to work, it must be *indistinguishable* from real traffic. Even if encryption prevents the service provider from distinguishing between real and dummy traffic based on the content of the messages, other attributes such as the timing or size of the messages may

be exploitable. In particular, OSNs pose a particularly challenging scenario as the wealth of data available may give away information about how users communicate. Do two users communicate more when they have more friends in common? Does their number of friends affect their communication patterns? Can we tell how a user communicates by looking at other publicly available information in the OSN?

Previous research has focused on modelling the OSN structure and studying inferences of private *attributes* from publicly available data. However, little is known about the feasibility of inferring *communication patterns*.

In this paper, we take the first steps towards this goal by performing, to the best of our knowledge, a first study on the feasibility of inferring private communication patterns from publicly available friendship and communication traffic data. To this end, we propose a model for communication inference in OSNs and analyse a dataset from a Belgian social network, *Netlog*[1], to determine how both friendship graph and public traffic data can expose private communication patterns.

This paper provides several contributions. Firstly, we study the likelihood with which an adversary can infer the private communication patterns of a user even when it only has access to OSN encrypted data or data stripped from its content. Examples of such scenario include an OSN analyst that *only* obtains metadata from the service provider or an OSN provider that implements end-to-end encryption and provides traffic data to a law enforcement agency.

Secondly, we provide an analysis of the topological and communication properties of Netlog, partially replicating previous studies. This allows us to further confirm the existence of key OSN properties such as the power-law degree distribution or the fact that users only communicate with a small subset of their friends.

Lastly, our results inform design strategies of obfuscation tools to achieve indistinguishability between real and dummy traffic, e.g., preventing dummy traffic to be filtered out when it does not match expected correlations with other available OSN data. Besides, our study can also inform OSN communication models, and thus allow researchers to simulate realistic communication patterns in OSNs.

2 Related work

In this section we shortly review how our work extends, differs from and complements previous work on the field.

[1]Netlog was merged with *Twoo* in 2015. Accessing `en.netlog.com` in April 2017 automatically redirects to `www.twoo.com`.

2.1 Modelling of online social networks

The problem of inferring communication patterns is analogous to deriving a model of communication in OSNs. Most efforts on OSN modelling have however been devoted to derive a model of the *topology* of the *social graph*, namely, a model of the characteristics of the network structure Ahn et al. (2007); Kossinets and Watts (2006); Kumar et al. (2010); Mislove et al. (2008, 2007), while fewer works have attempted to model activity and communication behaviour.

Network Topology

Several distinguishing properties of social graph topology have been identified in the literature. *Power-law degree distributions* Barabási and Albert (1999), *small diameters* Watts and Strogatz (1998), *assortativity* Mislove et al. (2007), *community structure* and *network modularity* Ferrara and Fiumara (2012) are among the most representative. These properties have in turn informed social network graphs *generators* that attempt to generate synthetic yet realistic social network graphs Sala et al. (2010). Still, these models are limited in that they only describe the topology of the network and do not capture users' activity, this is, how users communicate and the actions they perform in the OSN.

Users' Activity

Benevenuto et al. have proposed a model of OSN user behaviour that, to the best of our knowledge, is the most ambitious and comprehensive so far Benevenuto et al. (2009, 2012). They provide, among other features, a characterisation of session timing, the frequency and type of activities performed in the OSN and the number of friends users interact with. Similarly, Gyarmati and Trinh Gyarmati and Trinh (2010) have proposed a model of the number of logins and session duration per user based on their analysis of four popular social networks. These models however do not study how these activity features relate to one another or to other OSN data (e.g., topology). We provide a first contribution in this direction.

Despite the fact that no general model for user interaction has been proposed so far, social network activity has received significant attention in the literature, unveiling recurrent, characteristic patterns. Users have been found to typically communicate with a small subset of their friends Chun et al. (2008); Golder et al. (2006); Wilson et al. (2009) and to reply to most messages and posts received on the OSN, i.e., OSN interactions feature high *reciprocity* Chun et al. (2008); Jiang et al. (2010); Wilson et al. (2009). User communication also exhibits marked temporal patterns, e.g., differences between workdays and weekends Golder et al. (2006), or the fact that communication between two users seldom persists over time Viswanath et al. (2009). In this paper we analyse a different OSN dataset to confirm many of these findings. Moreover, we

also take up these previously identified OSN properties to carry out our work on the feasibility of inferences.

2.2 Inferences on OSNs

In this paper we study the feasibility of inferring private communication patterns on an OSN. Previous works have focused instead on inferring sensitive attributes such as sexual orientation or political affiliation. Their methodology is however very similar to ours, namely, based on Bayesian analysis. We shortly review some of these prior contributions.

He et al. He et al. (2006) have proposed an analysis framework based on Bayesian networks to infer personal attributes of OSN users based on the attribute values their friends declare. To test the suitability of the framework, they synthetically generate attributes for a network of users in LiveJournal[2], demonstrating that their framework successfully exposes relationships between the attributes of a user and her friends. Heatherly et al. Heatherly et al. (2013) show that combining both non sensitive attribute values and friendship links leads to more accurate inferences of sensitive values. Moreover, they propose countermeasures based on the removal of links and attributes to thwart potential inferences based on them.

Mislove et al. on the other hand use *community detection* to infer the attributes of the users in the network Mislove et al. (2010). Relying on the assumption that unknown users' attributes can be inferred from the attributes of the people in their community, Mislove et al. show that as few as 20% users with known attributes may allow very accurate inferences over the attributes of the remaining users. Zheleva and Getoor combine both community detection techniques and Bayesian analysis to infer, among other attributes, gender, location and marital status; suggesting that whereas friendship links do not necessarily enable accurate inferences, community membership does Zheleva and Getoor (2009). Chaabane et al. exploit semantic relationships between user data to infer unknown attributes Chaabane et al. (2012). They rely on a measure of similarity to assign to the unknown attributes of a user the known value of other, *similar* users.

Inferences may also allow to learn not only private information about the present state of the OSN and its users, but also about future events, e.g., the evolution of the OSN. *Link prediction* attempts to infer future interactions between OSN users taking into account the current state of the OSN. Some of the existing approaches Liben-Nowell and Kleinberg (2003); Al Hasan and J. Zaki (2011) are based on assigning a score to pairs of nodes to represent their proximity or similarity, an idea we also leverage in this paper.

[2] `http://www.livejournal.com/`

2.3 Obfuscation tools for traffic analysis resistance

Our work is further inspired by the design of obfuscation tools for traffic analysis resistance in online social networks Balsa et al. (2012). The goal of these tools is to prevent an adversary (be it the service provider or an external adversary) from profiling the users' communication patterns, namely, to accurately determine with whom and how often OSN users communicate. In order to do that, these tools generate and send out *dummy traffic*, i.e., fake, cover traffic that prevents an observer from accurately determining with whom the user *actually* communicates. To do this effectively, dummy messages must be indistinguishable from real ones. A first step to achieve this is to encrypt all communications, so that the adversary cannot distinguish real and dummy messages based on their content. Content is however not the only piece of information that may leak information about which messages are real and which ones are not. *Correlations* between the number of messages a user sends to a friend and other types of features may undermine the plausibility of the dummy traffic being generated. For example, if the number of private messages two users exchange and the number of posts that they leave on each other's wall are correlated, an obfuscation tool needs to take this into account to generate dummy traffic. Otherwise, if Alice does not communicate with Charlie and Alice's obfuscation tool sends *dummy messages* to Charlie, the adversary can potentially dismiss those dummy messages as obfuscation because she knows that if Alice had actually sent messages to Charlie she would have also posted on his wall. Our work therefore aims to inform the design of obfuscation tools for traffic analysis resistance by providing a first analysis of the relationship between OSN communication and other OSN features.

3 Communication Inference in Online Social Networks

3.1 A model of communication in online social networks

We model an online social network (OSN) as a *mixed multigraph* $G := (V, F, P, M)$. The set of *nodes* V represents the OSN users. The set of *friendships* F represents friend relations between the OSN users. The *multiset* of *posts* P, represents messages publicly posted on users' *walls*. The multiset of *messages* M represents the private messages sent on the OSN. Friendship relationships are represented with undirected edges while posts and messages are represented with arcs (directed edges).

For a specific OSN user $\mathrm{a} \in V$, say Alice, $F(\mathrm{a})$ denotes Alice's set of friend relationships. The set of posts sent and received by Alice are denoted as $\underrightarrow{P}(\mathrm{a})$ and $\underleftarrow{P}(\mathrm{a})$, respectively. The sets of sent and received messages are analogously represented as $\underrightarrow{M}(\mathrm{a})$ and $\underleftarrow{M}(\mathrm{a})$. The absence of an arrow indicates that both directions are considered thus $M(\mathrm{a}) = \{\underrightarrow{M}(\mathrm{a}) \cup \underleftarrow{M}(\mathrm{a})\}$ and

Table 1: Notation summary

Symbol	Definition
$G := (V, F, P, M)$	Mixed multigraph representing the OSN
V	Set of OSN users (nodes)
F	Set of friendships (edges)
P	Multiset of public posts (arcs)
M	Multiset of private messages (arcs)
$F(\mathrm{a})$	Alice's set of friend relationships
$\underrightarrow{P}(\mathrm{a})$	Multiset of posts sent by Alice
$\underleftarrow{M}(\mathrm{a})$	Multiset of messages received by Alice
$P(\mathrm{a})$	$\{\underrightarrow{P}(\mathrm{a}) \cup \underleftarrow{P}(\mathrm{a})\}$
$M(\mathrm{a})$	$\{\underrightarrow{M}(\mathrm{a}) \cup \underleftarrow{M}(\mathrm{a})\}$
$\underrightarrow{P}(\mathrm{a}, \mathrm{b})$	Multiset of posts Alice sent to Bob
$M(\mathrm{a}, \mathrm{b})$	Multiset of messages exchanged between Alice to Bob, i.e., $M(\mathrm{a}, \mathrm{b}) = \{\underrightarrow{M}(\mathrm{a}, \mathrm{b}) \cup \underleftarrow{M}(\mathrm{a}, \mathrm{b})\}$
$V_F(\mathrm{a})$	Set of nodes that are friends with Alice
$V_{\underrightarrow{P}}(\mathrm{a})$	Set of nodes to whom Alice sent posts
$V_M(\mathrm{a})$	Set of nodes that sent to or received messages from Alice, i.e., $V_M(\mathrm{a}) = \{V_{\underrightarrow{M}}(\mathrm{a}) \cup V_{\underleftarrow{M}}(\mathrm{a})\}$
$\bar{\bar{S}}$	Cardinality of the set S
Superscript T	Specifies time frame

$P(\mathrm{a}) = \{\underrightarrow{P}(\mathrm{a}) \cup \underleftarrow{P}(\mathrm{a})\}$.

The set $\underrightarrow{P}(\mathrm{a}, \mathrm{b})$ represents the posts Alice sent to Bob and, in the same way, $\underrightarrow{M}(\mathrm{a}, \mathrm{b})$ represents the messages Alice sent to Bob.

We denote as $V_F(\mathrm{a})$ the set of nodes in the induced subgraph formed by the set of friendships of Alice, this is, the set of nodes representing the friends of Alice. In the same way, $V_P(\mathrm{a})$ (respectively, $V_M(\mathrm{a})$) is the set of nodes in the induced subgraph formed by the multiset of posts P (analogously, messages M) sent and received by Alice, this is, the set of users that sent and received posts (correspondingly, messages) to and from Alice.

We denote as $\bar{\bar{S}}$ the cardinality of a set S, e.g., $\bar{\bar{V}}_F(\mathrm{a})$ denotes the number of friends Alice has on the OSN.

We use a superscript T to refer to communication taking place on a specific time period T, e.g., $V_{\underrightarrow{M}^T}(\mathrm{a})$ represents the set of users Alice sent a message to during time period T.

Table 1 presents a summary of all the notation described above.

3.2 Evaluating the feasibility of communication inference in online social networks

We model as random variables, R, both *unknown* variables to be inferred and *evidence* variables to perform inferences from. We denote the probability distribution of a random variable as $\mathrm{P}[\, R = x \,]$, e.g., $\mathrm{P}[\, \overrightarrow{\bar{\bar{M}}}(\mathrm{a}, \mathrm{b}) = x \,]$ represents the probability of a number x of messages sent by Alice to Bob. Similarly, $\mathrm{P}[\, R_1 \mid R_2, R_3, \ldots, R_k \,]$ denotes the conditional probability of R_1 given evidence from random variables $\{R_2, R_3, \ldots, R_k\}$, e.g., $\mathrm{P}[\, \overrightarrow{\bar{\bar{M}}}(\mathrm{a}, \mathrm{b}) = z \mid \overleftarrow{\bar{\bar{P}}}(\mathrm{b}, \mathrm{a}) = x \,]$ represents the probability that Alice sends z messages to Bob given that Bob left x posts on Alice's wall.

We use Shannon entropy Shannon (1948), denoted as $H(R)$, to measure the amount of uncertainty about the expected value of a random variable R. In other words, we use entropy to measure how easy it would be to infer the value of R. Low values of entropy represent easy inferences, namely, some values $R = r$ are far more likely than others. Conversely, high values of entropy indicate harder inference problems, as there is significant uncertainty about the actual value that R may take. The conditional entropy, denoted as $\mathrm{H}(\, R_1 \mid R_2, R_3, \ldots, R_k \,)$, measures the uncertainty about the expected value of R_1 when information about random variables $\{R_2, R_3, \ldots, R_k\}$ is available. Conditional entropy ranges from 0 to $H(R_1)$. When values $R_2 = r_2, R_3 = r_3, \ldots, R_k = r_k$ univocally determine $R_1 = r_1$, conditional entropy is 0. Conversely, when values of $R_2, R_3, \ldots, R_k$ provide no additional information about R_1, conditional entropy equals $H(R_1)$, as values $R_2, R_3, \ldots, R_k$ reveal nothing about R_1.

Both entropy and conditional entropy are related to *mutual information* through the following expression: $I(R_1; R_2) = H(R_1) - H(R_1|R_2)$. We have favoured mutual information over other measures of statistical dependence, such as correlation coefficients, for its *equitability*, i.e., its ability to detect general, not only linear or monotonic, dependence Khan et al. (2007); Kinney and Atwal (2014). We have chosen however to present results in this paper in terms of conditional entropy, which is trivial to obtain from the corresponding mutual information.

Practicalities The computation of both entropy and conditional entropy depends on the estimation of the probability distribution of the random variables involved. We *quantise* random variables to reduce the set of values they may take Balsa et al. (2012). Quantisation collapses several values on one *category* of values, effectively increasing the number of samples available per category. This reduces the error on the probability distribution estimation, albeit at the expense of coarser random variable values. Moreover, shorter lists of values allow for a speedier thus more efficient computation of the mutual information. To measure the underlying estimation error we resort to Bayesian Inference, using the methods described in Balsa et al. (2012).

Table 2: Description of the Netlog interaction dataset

Type	Data			Time period
Friendship	User 1 ID	User 2 ID	Day & time	Dec'02 - Oct'11
Posts	Poster ID	Recipient ID	Day & time	Dec'02 - Oct'11
Messages	Sender ID	Recipient ID	Day & time	May'11 - Oct'11

4 A Case Study: Netlog

4.1 The Netlog dataset

We have performed our study using a dataset from Netlog, a Belgian OSN.[3] Our dataset comprises communication data from the Dutch-speaking subnetwork in Netlog. Specifically, it includes three different sets of interaction data[4]:

Friendship requests and acceptances In Sect. 3.1 we have modeled friendship as an undirected edge between two users. We consider two users Alice and Bob to be friends (and thus a friendship edge is added to the social graph between the node representing Alice and the node representing Bob) when the dataset contains both a friendship request from Alice to Bob and a friendship acceptance from Bob to Alice.

Private messages Messages that are only visible to the sender and recipient of the message. The dataset contains traffic data of both sent and received private messages.

Public posts Messages that users leave on other users' personal pages and are publicly available. The dataset contains traffic data of both sent and received private messages.

Table 2 outlines the data obtained for each type of interaction and the time period for which complete data is available. Note that the dataset contains no personal attributes or the contents of any message or post, but only metadata. Moreover, the dataset was de-identified, namely, names were replaced by a random identifier.

Figure 1 sums up some statistics about the dataset. We use the acronym 'AT' (*All Time*) to tag those figures that refer to all the time for which data are available. Otherwise, figures refer to the six-month period of messages data. We use the terms *posting* and *messaging users* to refer to users that posted and sent at least one post or message, respectively. *Active* users either posted or sent at least one message and *strictly active* sent at least one of each. Note that active users are a small fraction of the total number of users in the network, as previously reported in the literature Benevenuto et al. (2009).

[3] See Footnote 1 in page 3.

[4] The dataset includes additional datafields which are not used for the study performed in this paper.

Number of users	3 834 304
Number of posts (AT)	175 731 008
Number of messages	70 170 964
Posting users (AT)	1 763 931
Posting users	180 182
Messaging users	379 611
Active users	443 398
Strictly active users	270 327
Average friend. degree	24.96
Std. dev. friend. degree	161.1

Figure 1: Dataset statistics

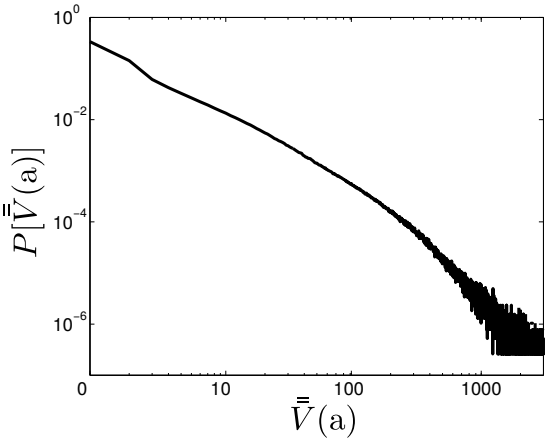


Figure 2: Degree probability distribution

Next we provide an analysis of the main features of both social graph and communication activity in the network, discussing to what extent they match or deviate from the previous studies we have referred to in Sect. 2. These features are the ones we use below in Sect. 4.2 to perform our evaluation on the feasibility of inferring private communication patterns.

4.1.1 Social graph topology

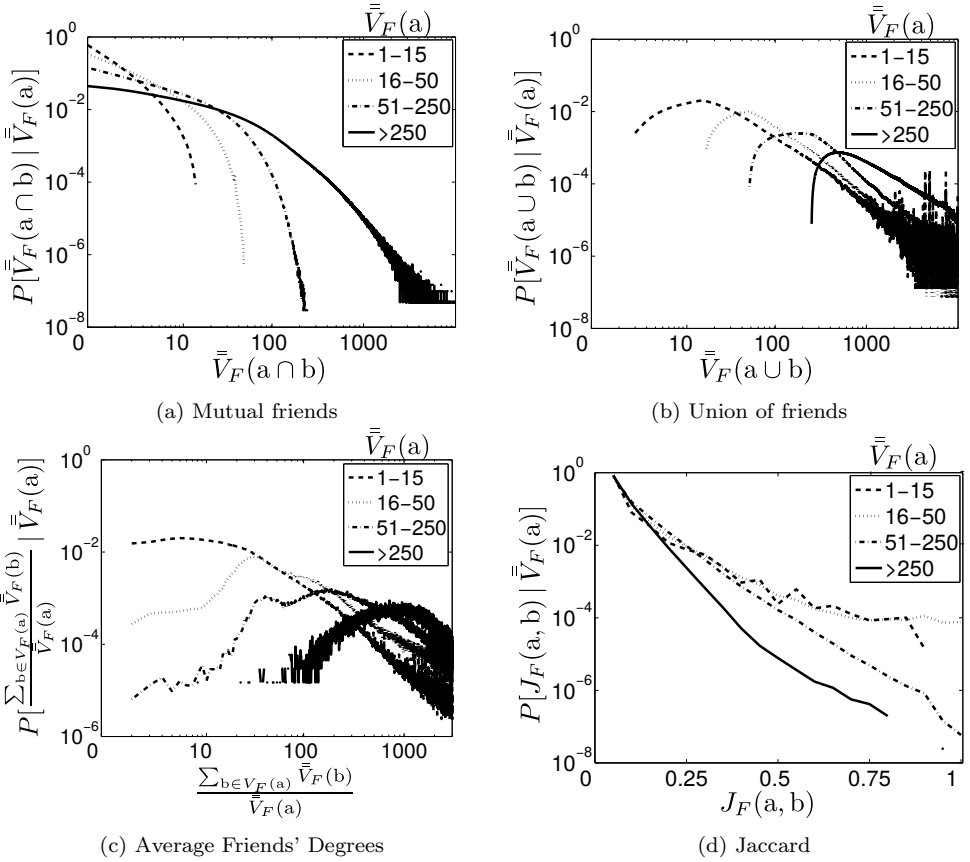
Figure 2 displays the distribution of the number of friends each user has, $\mathrm{P}[\,\bar{\bar{V}}_f(\mathrm{a})\,]$, which approximately follows a power-law with $\alpha = 2.2$.

Figure 3 represents the probability distribution of different features of Netlog's social graph topology. Figure 3a shows the probability distribution of the number of friends Alice has in common with each of her friends, namely —slightly abusing notation[5]—, $\mathrm{P}[\,\bar{\bar{V}}_F(\mathrm{a}\cap\mathrm{b}) \mid \bar{\bar{V}}_F(\mathrm{a})\,]$.

Figure 3b shows the probability of the number of different people that two friends, Alice and Bob, can jointly count among their friends, i.e., $\mathrm{P}[\,\bar{\bar{V}}_F(\mathrm{a}\cup\mathrm{b}) \mid \bar{\bar{V}}_F(\mathrm{a})\,]$. That number is strongly correlated with the degree of Alice because users tend to become friends with people that have a similar amount of friends in the OSN. This has been referred in the literature as *homophily* Mislove et al. (2010) and is shown in Fig. 3c, which shows the probability of the average degree of Alice's friends given Alice's own degree, namely, $\mathrm{P}[\,\frac{\sum_{\mathrm{b}\in F(\mathrm{a})}\bar{\bar{V}}_F(\mathrm{b})}{\bar{\bar{V}}_F(\mathrm{a})} \mid \bar{\bar{V}}_F(\mathrm{a})\,]$.

Lastly, Fig. 3d shows the probability distribution of the Jaccard coefficient between any two friends, i.e., $\mathrm{P}[\,\mathrm{J}_F(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_F(\mathrm{a})\,]$, where $\mathrm{J}_F(\mathrm{a},\mathrm{b}) = \frac{\bar{\bar{V}}_F(\mathrm{a}\cap\mathrm{b})}{\bar{\bar{V}}_F(\mathrm{a}\cup\mathrm{b})}$. Note that the greater the degree of Alice the lower the Jaccard coefficient is likely to be. The probability that Alice and Bob have the same friends decreases as the degree of any of them increases.

[5]According to our notation rules the correct form should be $\mathrm{P}[\,\overline{\overline{V_F(\mathrm{a})\cap V_F(\mathrm{b})}} \mid \bar{\bar{V}}_F(\mathrm{a})\,]$



(a) Mutual friends

(b) Union of friends

(c) Average Friends' Degrees

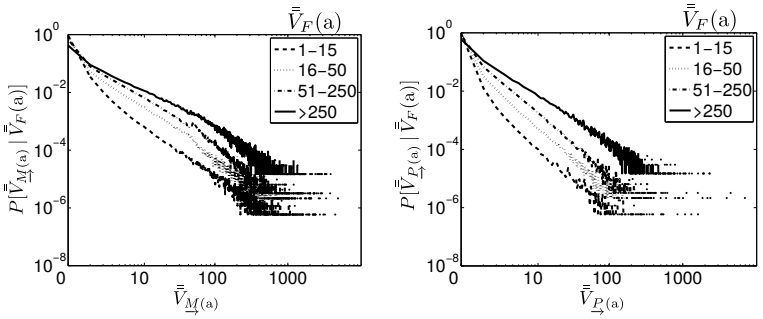
(d) Jaccard

Figure 3: Graph Features

4.1.2 Users' Communication Activity

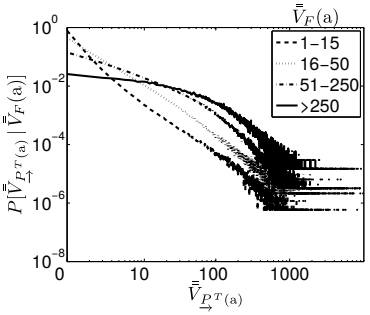
Figure 4 shows the distribution of the number of friends each user Alice communicates with, depending on her number of friends $\bar{\bar{V}}_F(a)$, showing that the more friends a user has, the greater the number of people $\bar{\bar{V}}_{\underrightarrow{M}}(a)$ and $\bar{\bar{V}}_{\underrightarrow{P}}(a)$ she sends messages or posts to, respectively. Not surprisingly, the number of people a user communicates with increases over time, as shown in Fig. 4c. Yet all three figures show that OSN users only communicate with a small subset of their *friends*, confirming previous results Golder et al. (2006); Huberman et al. (2008); Viswanath et al. (2009).

Figure 5 shows the probability distribution of the number of messages and posts a user Alice sends to each of her friends. Figure 5a shows that the number of messages does not depend on the number of friends Alice has, whereas Fig. 5b shows there is a slight dependency between the number of posts Alice sends to Bob and her number of friends, i.e., the fewer friends Alice has, the less posts she will send to each of them. Interestingly, this trend tends to disappear



(a) Number of people Alice messages.

(b) Number of people Alice posts to.



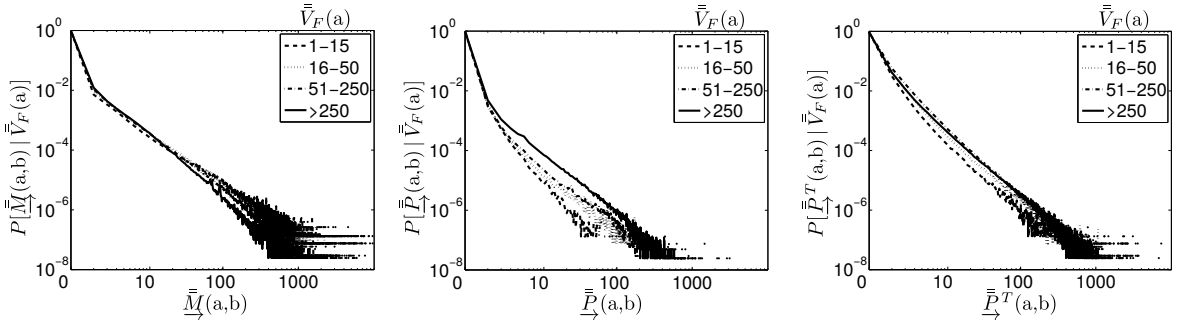
(c) Number of people Alice posts to (AT).

Figure 4: Distributions of the number of people a user sends messages and posts to.

over time, as shown in Fig. 5c.

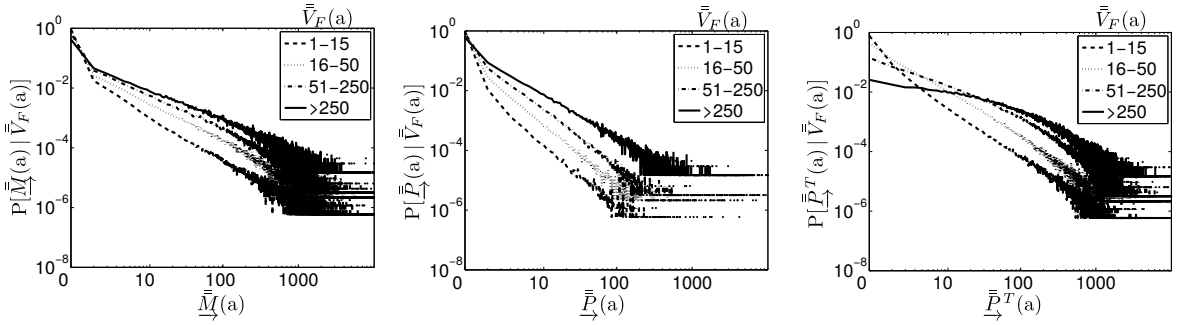
Figure 6 describes the probability distribution of the total number of messages (Fig. 6a) and posts (Figs. 6b and 6c) sent by Alice, supporting Fig. 4 in that the total number of messages and posts a user sends depends on the number of friends she has.

Lastly, Fig. 7 represents the degree of reciprocity for both messages (7a) and posts (7b and 7c). The figures represent the histogram of the pairs of counts of messages (or posts, according to the figure) that Alice sent to Bob and Bob sent to Alice. Darker areas represent a higher incidence of those pairs of values in the network. For example, the number of times that Alice sends 3 messages to Bob and Bob sends 3 messages to Alice is much higher (in the dark area of the figures) than the number of times that Alice sends 3 messages to Bob and Bob sends 100 messages to Alice (in the light area of the figures). As shown in all three figures and confirming what has been previously reported Chun et al. (2008); Jiang et al. (2010); Wilson et al. (2009), users have a strong tendency to reciprocate the messages and posts they receive.



(a) Messages Alice sends to Bob (b) Posts Alice sends to Bob (c) Posts Alice sends to Bob (AT)

Figure 5: Distributions of the number of messages and posts Alice sends to Bob.



(a) Messages (b) Posts (c) Posts (AT)

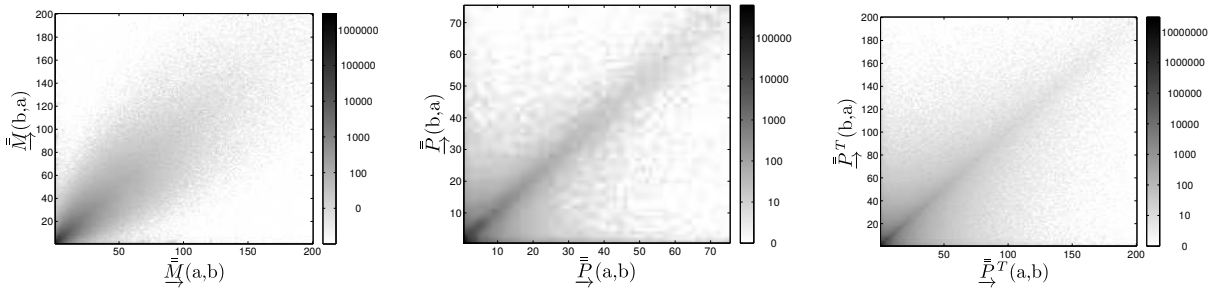
Figure 6: Probability distributions of the total number of messages and posts Alice sends.

4.2 Inferring private communication in Netlog

In this section we present the results of our evaluation on the feasibility of inferring private communication patterns. Unless otherwise stated, all figures included in this section follow the same representation formula. They display conditional probability distributions $\mathrm{P}[\, Z \mid X \,]$ where Z represents the variable to be inferred (e.g., number of messages sent by Alice to Bob) and X represents the evidence variable (e.g., the number of friends Alice and Bob have in common). In the figures, the x-axis represents values of the independent variable $X = x$, and the y-axis the probability $\mathrm{P}[\, Z = z \mid X = x \,]$. The figures may also feature error bars, which represent the standard error on a 99% confidence interval. Table 3 summarises the features that we have chosen to analyse in our study.

4.2.1 Messaging behaviour based on network topology features

We have analysed the relationship between the network topology and the number of private messages users send.



(a) Messages (b) Posts (c) Posts (AT)

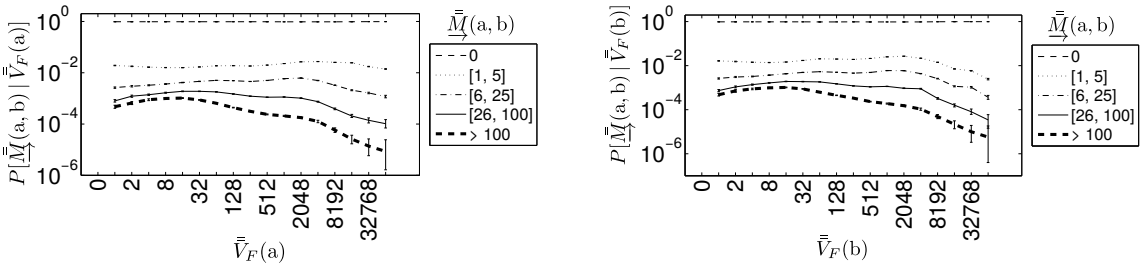
Figure 7: Communication Reciprocity

Table 3: List of features involved in our inference analysis. The *posts graph* is the communication graph *induced* by the multiset of posts P

Data source	Features	Source visibility
Friendship	$\bar{\bar{V}}_F(\mathrm{a})$; $\bar{\bar{V}}_F(\mathrm{a} \cap \mathrm{b})$; $\bar{\bar{V}}_F(\mathrm{a} \cup \mathrm{b})$; $\mathrm{J}_F(\mathrm{a},\mathrm{b})$	Public
Posts	$\underrightarrow{\bar{\bar{P}}}(\mathrm{a},\mathrm{b})$; $\underrightarrow{\bar{\bar{P}}}(\mathrm{b},\mathrm{a})$	
	$\underrightarrow{\bar{\bar{P}}}^T(\mathrm{a},\mathrm{b})$; $\underrightarrow{\bar{\bar{P}}}^T(\mathrm{b},\mathrm{a})$	
Posts graph	$\bar{\bar{V}}_P(\mathrm{a} \cap \mathrm{b})$; $\underrightarrow{\bar{\bar{V}}}_P(\mathrm{a} \cap \mathrm{b})$; $\underleftarrow{\bar{\bar{V}}}_P(\mathrm{a} \cap \mathrm{b})$	
	$\bar{\bar{V}}_{P^T}(\mathrm{a} \cap \mathrm{b})$; $\underrightarrow{\bar{\bar{V}}}_{P^T}(\mathrm{a} \cap \mathrm{b})$; $\underleftarrow{\bar{\bar{V}}}_{P^T}(\mathrm{a} \cap \mathrm{b})$	
	$\bar{\bar{V}}_{P^T}(\mathrm{a} \cup \mathrm{b})$; $\underrightarrow{\bar{\bar{V}}}_{P^T}(\mathrm{a} \cup \mathrm{b})$; $\underleftarrow{\bar{\bar{V}}}_{P^T}(\mathrm{a} \cup \mathrm{b})$	
Private messages	$\bar{\bar{M}}(\mathrm{a},\mathrm{b})$; $\underrightarrow{\bar{\bar{M}}}(\mathrm{a},\mathrm{b})$; $\underrightarrow{\bar{\bar{M}}}(\mathrm{b},\mathrm{a})$	Private

Messages sent given topological degree (number of friends) Figures 4 and 6 have shown that with an increasing number of friends Alice is slightly more likely to send messages to more of her friends as well as to send slightly more messages overall. However, those increases are not linear with the number of friends, which means that Alice only communicates with a small subset of her friends and distributes a limited *"budget"* of messages among them. We analyse whether the number of Alice's friends has an impact on this distribution, namely, Alice may distribute her *"budget"* equally among her friends —sending less messages to each friend—, or not —ignoring certain friends, whether old or new. Besides, if Alice has more friends, more of those friends may send her messages and, considering strong reciprocity (cf. Fig. 7), that may have an impact on how Alice chooses to communicate.

Figure 8a shows the conditional probability distribution of the number of messages Alice sends to Bob given her number of friends, namely, $\mathrm{P}[\,\underrightarrow{\bar{\bar{M}}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_F(\mathrm{a})\,]$. Our results suggest that the number of friends users have is not a good indicator of the number of messages they send to any of their friends. Alice seems to choose the number of messages she sends to any of her friends



(a) Probability distribution of the number of messages Alice sends to Bob given Alice's number of friends.

(b) Probability distribution of the number of messages Alice sends to Bob given Bob's number of friends.

Figure 8: Inferring messages given topological degree

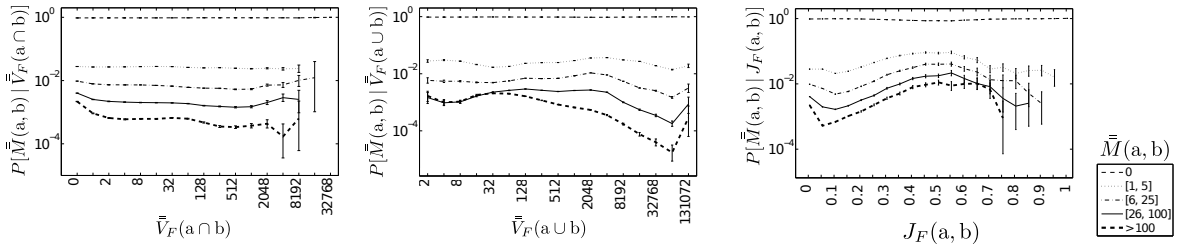
Table 4: Entropy of number of messages given topological degree

	Bits
Ref.: $H(\vec{\bar{\bar{M}}}(\mathrm{a},\mathrm{b}))$	0.2044
$H(\vec{\bar{\bar{M}}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_F(\mathrm{a}))$	0.2033
$H(\vec{\bar{\bar{M}}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_F(\mathrm{b}))$	0.2037

regardless of her own number of friends. This is in line with previous results Golder et al. (2006); Huberman et al. (2008); Viswanath et al. (2009) showing that not only Alice communicates with just a small subset of friends regardless of the total number of friends amassed on the social network, but also that the amount of messages she sends to any of her friends is neither affected. We posit that most of those additional friends are mere acquaintances that Alice is not interested in communicating with on a regular basis, i.e., additional friends do not disrupt Alice's stable communication patterns with a small circle of friends.

Besides, Alice may choose to message more popular friends (i.e., Bob has many friends and contacts on the network) or on the contrary favour less well-connected people (e.g., Bob has fewer friends and can therefore devote more of his attention to her). From our analysis, the number of messages Alice sends to Bob is independent from the number of friends Bob has, as shown in Fig 8b. The analysis of the entropies, as shown in Table 4, further confirms this. Knowing the number of friends either Alice or Bob have barely reduces the uncertainty on the number of messages Alice sends to Bob. This suggests that Alice's motivation to message Bob is neither determined by Bob's popularity nor lack thereof.

Messages exchanged given subnetwork graph We have analysed the relationship between the number of messages two friends *exchange* with respect to various features of their local



(a) Probability distribution of the number of messages Alice and Bob exchange given their number of common friends.

(b) Probability distribution of the number of messages Alice and Bob exchange given the size of the union of their friends sets.

(c) Probability distribution of the number of messages Alice and Bob exchange given the Jaccard index of their friends sets.

Figure 9: Messages exchanged given subnetwork graph.

subnetwork graph. This allows us to evaluate whether the network of friends surrounding two users may reveal any information about the volume of their private communication. We posit that if two users have several friends in common, they may belong to a tight social circle (e.g., their family) and may therefore be likely to be in touch and often communicate. On the other hand, a very large number of common friends may also mean that both users belong to a large and loose community of acquaintances (like a company's employees or university alumni) and may therefore seldom communicate.

The total number of friends two users have may also reveal information about how much they communicate. Users that together total a small number of friends may message each other more often, as they have few other friends to communicate with. Conversely, when either Alice or Bob (or both) have a large number of friends, the probability that they message each other might be smaller, as they have many other friends they can communicate with. Lastly, it may be that users only communicate more with each other when the amount of common friends they share make up a certain percentage of their friends. If Alice and Bob share many common friends but those are all the friends they have, they may simply be using the social network to keep in touch with a loose community of acquiantances. Conversely, if their common friends represent just a fraction of their friends, they may be part of a tighter circle of friends and communicate more often.

Figure 9 shows the probability of the number of messages two users exchange given their mutual friends (Fig. 9a), the *union* of their friends, (Fig. 9b), and their Jaccard coefficient (Fig. 9c). None of the three features seems to provide information about the number of private messages two users exchange. The probability of any number of messages stays relatively constant for numbers of mutual friends below 1024. Beyond that number the error increases significantly —as few users have more than 1024 mutual friends—, yet nothing indicates a potential change in trend.

Table 5: Entropy of messages exchanged given subnetwork graph

	Bits
Ref.: $H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}))$	0.2751
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_F(\mathrm{a}\cap\mathrm{b}))$	0.2745
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_F(\mathrm{a}\cup\mathrm{b}))$	0.2734
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \mathrm{J}_F(\mathrm{a},\mathrm{b}))$	0.2738

Table 5 displays the results of the entropies analysis. This confirms that all three features provide little information, with the union of friends being only slightly more informative. Note that the Jaccard index depends on both the number of mutual friends (non-informative) and the union of friends (slightly more informative), thus the effect of the former may diminish the amount of information provided by the latter.

We posit that because two users can often belong to several loose communities where they share many common friends (e.g., colleagues, schoolmates, university alumni and neighbours, to name a few) this is not a good indicator of the volume of communication between two people. Similarly, the union of the set of friends and the Jaccard coefficient between two users do not provide significant information about the number of messages they exchange. The former supports previous results in that it seems that communication between users does not depend on how many friends they have, i.e., users communication patterns with their close group of friends is unaffected by the number of friends each of them has on the network. The latter is on the other hand a combination of both the intersection and the union sets of friends. The combination of both features does not seem to provide significant information about the amount of private communication. Again, this may be due to the fact that users belong to different loose communities that may or may not make up the great majority of their friends. High variability in size, number and overlap of these communities may explain why the Jaccard coefficient provides no additional information.

4.2.2 Messaging behaviour based on posting behaviour

We have analysed the relationship between private communication patterns and public communication patterns.

Messages sent given sent or received posts Users may choose to send private messages to people they are not willing to publicly reveal they communicate with. On the other hand, users may also largely use both messages and posts interchangeably to communicate with their

friends. We have analysed whether any relationship exists between the number of posts and messages two users exchange. Figure 10a represents the probability of the number of messages Alice sends to Bob given the number of posts she writes to him in the same period of time (i.e., 6 months). The probability of having sent at least one or more messages significantly rises when Alice leaves a post on Bob's wall, steadily increasing for even larger numbers of posts. The same is true when we consider the number of posts Alice receives from Bob, shown in Fig. 10b, suggesting that the number of posts Alice sends to Bob and the ones she receives from him are equally informative to infer the number of messages they have exchanged. This is not surprising given the high communication reciprocity observed in the network (cfr. Fig. 7). Knowing that Alice sent a specific number of posts to Bob does not however precisely determine the number of messages she has sent to him, as the probability to send *any* number of messages increases with the number of posts at a rather similar and marginally incremental rate for any number of messages.

Hence, these results suggest that private messaging is more likely to take place when public posting has taken place. *However,* there is no correlation between the volume of public and private communication, i.e., it is not possible to assume that if Alice leaves a large volume of posts on Bob's wall she will also send her a large number of messages. Therefore, users do not interchangeably choose to send messages or posts to their friends, as if that was the case there should be a correlation between both types of communication. Still, these results have implications for communication modelling and obfuscation, namely, a realistic and plausible model of user communication must consider overlapping subsets of friends for both private and public communication.

We have tested whether having access to a longer history of posting behaviour enables better inferences, the rationale being that long term observation of public behaviour enables to more acurately determine who are a users' *"true"* friends. Figure 10c represents the probability that Alice sends z messages to Bob on a 6-month period given that Alice wrote to him x posts in the previous 9 years. The probability that Alice sends messages to Bob still increases with the number of posts she or he left on his or her wall. However, the relationship between posts and messages seems to be weaker, suggesting that communication profiles are not stable and therefore previous posting history may not be as reliable a predictor of recent messaging behaviour as the evidence of recent posts. Another hypothesis suggests that users choose to message close, long-term friends through alternative communication channels to social media, favouring the latter for casual, sporadic conversation with a clique of friends.

Table 6 shows the entropies of the distributions represented above. Note that the entropy of the messages probability distribution barely changes conditioned on the number of posts. This highlights that, in spite of the trends shown in the pictures, information about the number of

Table 6: Entropy of sent messages given sent/received posts

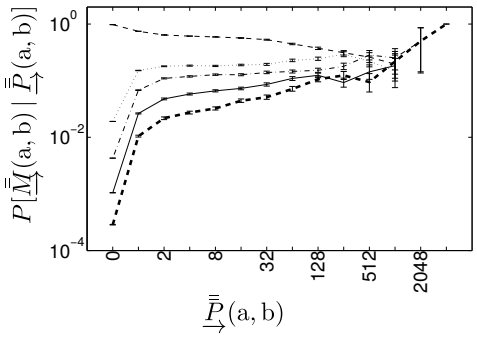
	Bits
Ref.: $H(\bar{\bar{\underrightarrow{M}}}(\mathrm{a},\mathrm{b}))$	0.2044
$H(\bar{\bar{\underrightarrow{M}}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{\underrightarrow{P}}}(\mathrm{a},\mathrm{b}))$	0.1989
$H(\bar{\bar{\underrightarrow{M}}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{\underrightarrow{P}}}(\mathrm{b},\mathrm{a}))$	0.1996
$H(\bar{\bar{\underrightarrow{M}}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{\underrightarrow{P}}}^T(\mathrm{a},\mathrm{b}))$ (AT)	0.2031
$H(\bar{\bar{\underrightarrow{M}}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{\underrightarrow{P}}}^T(\mathrm{b},\mathrm{a}))$ (AT)	0.2033

publicly exchanged posts does not provide significant information about the private messages. Still, this may predominantly be due to the fact that, in this OSN, users exchange no messages regardless of the number of posts they have left for each other.

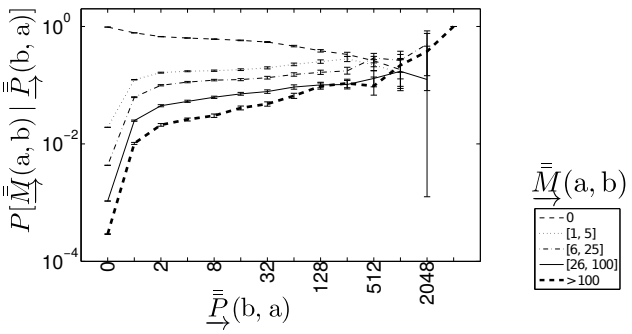
Lastly, we have analysed whether considering *proportions* or *percentages* instead of absolute numbers may lead to better inferences. Figure 10e represents the probability that Alice sends to Bob a certain percentage of all the messages she sends when she has written to Bob a certain percentage of all the posts she wrote (when Alice has written at least 5 posts) and Fig. 10f when she has received from Bob a certain percentage of all the posts people have posted to her (when Alice has received at least 5 posts). Both figures show that considering proportions of posts instead of absolute numbers do not enable better inferences either. This result further debunks the idea that users interchangeably use posts and messages to communicate with their friends, as otherwise both percentages would be correlated.

Exchanged messages given posting friends We have analysed the relationship between the number of messages two friends *exchange* with respect to their shares of *posting friends*, namely, those friends they send to or receive posts from. Specifically, we have considered the number of *mutual* posting friends (Figs. 11a and 11b) and the *union* of posting friends (Fig. 11c). Note that this is a *"hybrid"* analysis that combines evidence of public communication with the graph structure or network it induces. The first hypothesis is that if Alice and Bob leave posts to the same set of people, they may be more likely to also communicate with each other. The same reasoning we have proposed in section 4.2.1 applies to the analysis below of more complex features such as intersection, union and Jaccard coefficient over sets of posting friends.

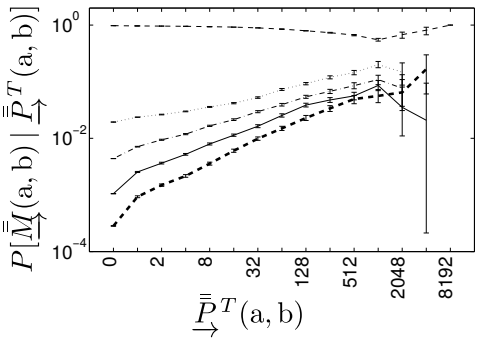
Our analysis of the data shows however that none of these features enable inferences of private message communication. The number of friends that both Alice and Bob have sent to or received messages from provides little information about the number of messages Alice and Bob exchange, regardless of whether we consider the posts on the same period of time



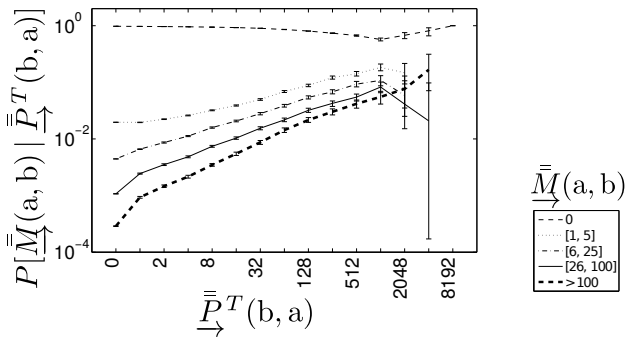
(a) Probability distribution of the number of outgoing messages given number of outgoing posts.



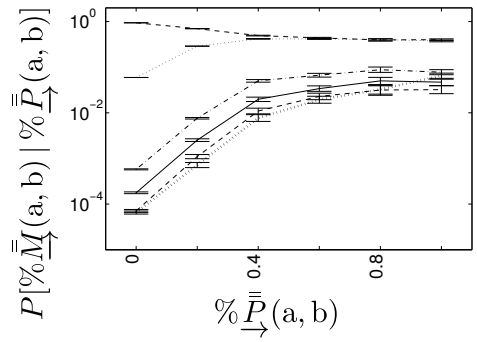
(b) Probability distribution of the number of outgoing messages given number of incoming posts.



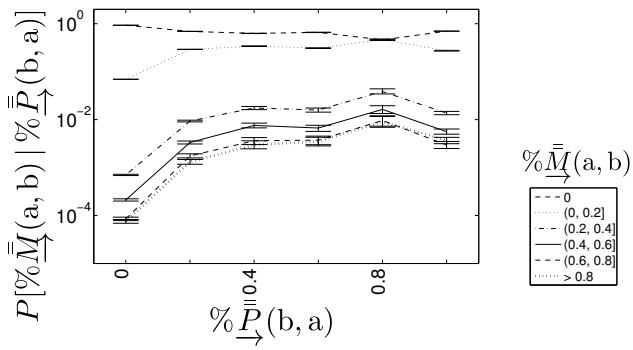
(c) Prob. distribution of the number of outgoing messages given number of outgoing posts (AT).



(d) Prob. distribution of the number of outgoing messages given number of incoming posts (AT).



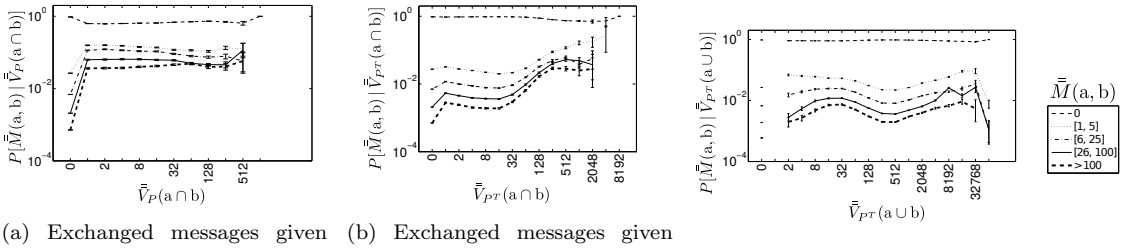
(e) Prob. distribution of the percentage of outgoing messages given percentage of outgoing posts.



(f) Prob. distribution of the percentage of outgoing messages given percentage of received posts.

Figure 10: Probability distribution of messages sent given sent/received posts.

(Fig. 11a) or a longer history (Fig. 11b). The probability that Alice and Bob exchange at least one message substantially increases when the number of mutual posting friends is greater than



(a) Exchanged messages given number of mutual posting friends.

(b) Exchanged messages given number of mutual posting friends (AT).

(c) Exchanged messages given union of posting friends.

Figure 11: Probability distributions of exchanged messages given posting friends.

one. As for the exact number of messages exchanged, this evidence variable does not provide enough information. The same occurs when considering the union of posting friends, namely, those friends that at least Alice or Bob have sent to or received a post from. The analysis of the conditional entropy confirms these results, as shown in Table 7 where we include the results of further analyses which we have not represented in the figure mainly due to the high similarity with the ones above.

There may be a number of reasons that explain this. The simpler is that users in this particular network favour public posting over private messaging, therefore exhibiting a generalised lack of private communication that barely correlates with any other examined feature. A more complex explanation is that the wide variety of types of users, the communities they belong to and the contexts and situations in which they choose to communicate contribute to multiple forms of communication that fail to emerge as consistent patterns to enable inferences. In the next section we further discuss the results presented up to this point and their practical implications.

5 Discussion

We mentioned in the introduction that our study has two main implications.

Firstly, regarding the feasibility of inferences, we set out to determine to what extent an adversary could infer the private communication patterns of OSN users from OSN metadata. Our analysis shows that it may not be possible to infer private communication patterns in OSNs from publicly available data such as graph topology or public communication. In the particular case of Netlog, an adversary would not be able to easily infer private communication patterns from publicly available traffic patterns or the social graph topology. Were our results to be confirmed for other social network data or features thereof, this would represent a *"natural" privacy protection* against inferences. However, it is important to understand that it is not

Table 7: Conditional entropies given posting friends sets

	Bits
Ref.: $H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}))$	0.2751
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_{P^T}(\mathrm{a}\cap\mathrm{b}))$ (AT)	0.2744
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_{P^T}(\mathrm{a}\cap\mathrm{b}))$ (AT)	0.2744
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_{\overrightarrow{P^T}}(\mathrm{a}\cap\mathrm{b}))$ (AT)	0.2730
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_{\overleftarrow{P^T}}(\mathrm{a}\cap\mathrm{b}))$	0.2712
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_{P}(\mathrm{a}\cap\mathrm{b}))$	0.2726
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_{\overrightarrow{P}}(\mathrm{a}\cap\mathrm{b}))$	0.2721
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_{\overleftarrow{P}}(\mathrm{a}\cap\mathrm{b}))$	0.2736
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_{P^T}(\mathrm{a}\cup\mathrm{b}))$ (AT)	0.2737
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_{\overrightarrow{P^T}}(\mathrm{a}\cup\mathrm{b}))$ (AT)	0.2731
$H(\bar{\bar{M}}(\mathrm{a},\mathrm{b}) \mid \bar{\bar{V}}_{\overleftarrow{P^T}}(\mathrm{a}\cup\mathrm{b}))$ (AT)	0.2731

possible to generalise the results of our study to all OSNs.

Whereas many of our results are consistent with previous findings (e.g., the fact that users communicate with a small subset of their friends or the fact that their communication exhibits a high degree of reciprocity Chun et al. (2008); Golder et al. (2006); Wilson et al. (2009)), that should not lead us to assume that the absence of correlations between the OSN data and features we have considered is a universal property of all OSNs. These results relate to one particular social network (Netlog) and a subset of all posible features that we have chosen to examine. Future studies may uncover relationships between OSN features that enable inferences on private communication, be it in other social networks (e.g., Facebook, Twitter, Google+ or Renren, to name a few), from other available data such as *Likes*, *comments* or tagged photos or from more complex graph or communication features (e.g., eccentricity, clustering coefficients or centrality). In this sense, and similarly to the initial models of OSN topological structure that only captured one or two (sometimes even conflicting) features Newman et al. (2002); Backstrom et al. (2006); Kumar et al. (2010), our study is a first step and contribution in this direction.

Having said that, our analysis of the entropy of the conditional probability distributions between the features we have examined leaves little room for doubt: in this particular social network publicly available information about the graph and communication does not improve our ability to infer the number of private messages users exchange. Considering the low volume of communication our best guess would be to assume that, with high probability, two users do not communicate at all. As indicated above, OSN users tend to befriend a large number of people but only communicate with a small subset of them.

We may further advance some hypotheses to explain the absence of correlations. OSNs are known to feature *tight-knit* social structures Mislove et al. (2010), namely, users belong to different closed communities where members are akin to each other. Considering that many of these members share the same OSN features but users still communicate with only a small subset of them, the features would not help us infer who among those members users communicate with. In this sense, identifying the specific communities (e.g., close friends, relatives, co-workers) and the links across them may allow for better inferences than the features of the underlying graph structure.

More generally, communication traffic data and topological features stripped off all content or *semantics* may simply not be informative enough to perform inferences on who the users privately communicate with. One may need to know whether the friends two users have in common are co-workers or relatives, or whether the posts that two users exchange are meant to be read by a wide audience or a reduced subset of their friends. Besides, the effemeral nature of the users' communication behaviour, i.e., the fact that users' communication profiles rapidly change, may further prevent patterns and correlations between features from emerging.

Secondly, our results have also implications for the modelling of OSN communication and the design of obfuscation tools for traffic analysis resistance in OSNs.

With respect to OSN modeling, not only have we confirmed, as shown in previous studies, certain social graph topological properties such as the power-law nature of the friendship degree distribution or the fact that users communicate with a small subset of their friends. Most importantly, we have provided what to the best of our knowledge is the first analysis of the relationship between private communication patterns and other OSN data such as public communication traffic data and the social graph topology features, showing that, at least in the case of Netlog, no direct relationship exists between these. The lack of correlations greatly simplifies the modelling of OSN communication in OSNs, as each of the features can be generated and simulated independently from each other.

We recall that our inference analysis was motivated by a particular *threat model*, namely, a social network where users encrypt their communications and no content is available to either the OSN service provider or an external adversary, *"only"* communication traffic and social graph data. In this context, we wondered, would it be possible for these adversaries to infer anything about the users' private communications? And were this the case, what would it take for users or privacy technology designers to prevent that from happening?

We must recall that Netlog itself provided to its users neither encryption tools nor communication traffic or social graph data obfuscation tools. Yet none of these are limitations for our study. On the one hand, by omitting the users' communication content in our analyses we have effectively *"simulated"* encrypted (and padded) communications. On the other hand, if

Netlog provided communication traffic or social graph data obfuscation, it would have prevented us from reliably determining whether correlations between these data existed (as we would be measuring correlations between obfuscated data).

Analogously to the case of OSN communication modelling, the absence of correlations allows a designer to treat these features independently Balsa et al. (2012). Effective obfuscation requires *plausibility*. If OSN features are correlated, an obfuscation strategy must take this into account as otherwise the adversary can exploit the correlation to filter the obfuscation out. The absence of correlations thus not only prevents inferences but also simplifies the design of obfuscation tools against more powerful adversaries.

Limitations Because our dataset was stripped off all content, we could not easily prune bots and spammers off the dataset, thus their impact in our results cannot be accurately determined. Still, our results are consistent across the whole range of topological degrees and number of posts and messages sent and received, i.e., we have not identified a subset of users that exhibits a different behaviour. We therefore assume that this kind of users would have had a limited impact in our results.

6 Conclusion and future work

Users of online social networks are often provided with privacy settings that allow them to control what is publicly visible and what is private on the site. Dependence between different types of OSN data may however enable an adversary to perform inferences about the private data based on other OSN available data.

Previous work has focused on inferences about private or non disclosed attributes of OSN users. In this paper we have performed a first analysis on the feasibility of inferring private communication patterns, i.e., with whom and how often a user communicates. We have focused on traffic data because while users may use their privacy settings or use encryption to hide their messages and sensitive attributes, traffic data cannot be easily hidden away from the service provider or other external adversaries.

We have used both the friendship graph and public communication traffic data from Netlog, a Belgian OSN, to evalute to what extent these publicly available data leak information about the amount of private messages users exchange. We have found that, in Netlog, such leakage of information is minimal as the number of messages users exchange is not related to the OSN features we have examined. Still, our results cannot be generalised to all OSNs and further work is needed to confirm whether or not this generally applies to user communication in other or all OSNs.

Future work could therefore go in three different directions. First, try to replicate our analysis in other social networks. Are the results we have obtained observable in other platforms or are our results particular of Netlog? Second, analyse the relationship between other features and data. In this work we have focused on social graph topology and public communication traffic data. Our inference analysis was motivated by a specific setting, namely, one where users encrypt their communications and attributes such that the content of these are not available to an adversary. However, this may not always be the case. Whenever content is available, it may be exploitable by an adversary. The content of messages or personal attributes such as age, marital status or gender may allow better inferences about users' private communication. Also, other social network data may be available depending on the OSN site itself, such as *Likes*, comments, photo tags, *shares* and so on. These may also provide information about users' private communication. Third, future work should also examine alternative or more complex analysis methodologies. We have relied on a Bayesian framework to determine to what extent a given feature may leak information about another. Other methods should be explored. Modeling the inference problem as a binary classification problem where the goal is to predict whether there is any private communication between each pair of users may be a promising avenue. Existing classification algorithms such as random forests, support vector machines, or ensembles of classifiers could be applied to train the classifier. Besides, in terms of features, the ones used in this paper represent just a starting point. Other more complex features (e.g., eccentricity, centrality or clustering coefficient, among many others) may enable better inferences.

Acknowledgements

This work has been partially supported by the Research Council KU Leuven through C16/15/058 and OT/13/070; the Flemish Government through FWO G.0360.11N Location Privacy and FWO G.068611N Data mining; the European Commission through H2020-DS-2014-653497 PANORAMIX and H2020-ICT-2014-644371 WITDOM; the Spanish Government through the project TIN2014-55243-P and grant FPU-AP2010-0078; and the Catalan Government through AGAUR 2014SGR-691.

References

Yong-Yeol Ahn, Seungyeop Han, Haewoon Kwak, Sue Moon, and Hawoong Jeong. 2007. Analysis of topological characteristics of huge online social networking services. In *Proceedings of the 16th international conference on World Wide Web*. ACM, 835–844.

Mohammad Al Hasan and Mohammed J. Zaki. 2011. Link prediction in social networks. In *Social Network Data Analytics*, Charu C. Aggarwal (Ed.). Springer, IBM Thomas J. Watson Research Center.

Lars Backstrom, Dan Huttenlocher, Jon Kleinberg, and Xiangyang Lan. 2006. Group formation in large social networks: membership, growth, and evolution. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 44–54.

Ero Balsa, Carmela Troncoso, and Claudia Diaz. 2012. A Metric to Evaluate Interaction Obfuscation in Online Social Networks. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 20, 6 (2012), 877–892.

Albert-László Barabási and Réka Albert. 1999. Emergence of scaling in random networks. *science* 286, 5439 (1999), 509–512.

Filipe Beato, Markulf Kohlweiss, and Karel Wouters. 2011. Scramble! Your Social Network Data. In *PETS (Lecture Notes in Computer Science)*, Simone Fischer-Hübner and Nicholas Hopper (Eds.), Vol. 6794. Springer, 211–225.

Fabrício Benevenuto, Tiago Rodrigues, Meeyoung Cha, and Virgílio Almeida. 2009. Characterizing user behavior in online social networks. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*. ACM, 49–62.

Fabrício Benevenuto, Tiago Rodrigues, Meeyoung Cha, and Virgílio A. F. Almeida. 2012. Characterizing user navigation and interactions in online social networks. *Inf. Sci.* 195 (2012), 1–24.

Abdelberi Chaabane, Gergely Acs, Mohamed Ali Kaafar, and others. 2012. You are what you like! information leakage through users' interests. In *Proceedings of the 19th Annual Network & Distributed System Security Symposium (NDSS)*.

Hyunwoo Chun, Haewoon Kwak, Young-Ho Eom, Yong-Yeol Ahn, Sue B. Moon, and Hawoong Jeong. 2008. Comparison of online social relations in volume vs interaction: a case study of cyworld. In *Internet Measurement Comference*, Konstantina Papagiannaki and Zhi-Li Zhang (Eds.). ACM, 57–70.

Danah Boyd. 2008. *Taken Out of Context: American Teen Sociality in Networked Publics*. Ph.D. Dissertation. University of California-Berkeley, School of Information. `http://www.danah.org/papers/TakenOutOfContext.pdf`

Emilio Ferrara and Giacomo Fiumara. 2012. Topological Features of Online Social Networks. *CoRR* abs/1202.0331 (2012).

Scott A. Golder, Dennis M. Wilkinson, and Bernardo A. Huberman. 2006. Rhythms of social interaction: messaging within a massive online network. *CoRR* abs/cs/0611137 (2006).

Ralph Gross and Alessandro Acquisti. 2005. Information revelation and privacy in online social networks. In *WPES*, Vijay Atluri, Sabrina De Capitani di Vimercati, and Roger Dingledine (Eds.). ACM, 71–80.

László Gyarmati and Tuan Anh Trinh. 2010. Measuring user behavior in online social networks. *IEEE Network* 24, 5 (2010), 26–31.

Jianming He, Wesley W Chu, and Zhenyu Victor Liu. 2006. Inferring privacy information from social networks. In *Intelligence and Security Informatics*. Springer, 154–165.

Raymond Heatherly, Murat Kantarcioglu, and Bhavani Thuraisingham. 2013. Preventing private information inference attacks on social networks. *Knowledge and Data Engineering, IEEE Transactions on* 25, 8 (2013), 1849–1862.

Bernardo A Huberman, Daniel M Romero, and Fang Wu. 2008. Social networks that matter: Twitter under the microscope. *arXiv preprint arXiv:0812.1045* (2008).

Carter Jernigan and Behram F. T. Mistree. 2009. Gaydar: Facebook Friendships Expose Sexual Orientation. *First Monday* 14, 10 (2009). `http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2611`

Jing Jiang, Christo Wilson, Xiao Wang, Peng Huang, Wenpeng Sha, Yafei Dai, and Ben Y. Zhao. 2010. Understanding latent interactions in online social networks. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement (IMC '10)*. ACM, New York, NY, USA, 369–382. `http://doi.acm.org/10.1145/1879141.1879190`

Shiraj Khan, Sharba Bandyopadhyay, Auroop R Ganguly, Sunil Saigal, David J Erickson III, Vladimir Protopopescu, and George Ostrouchov. 2007. Relative performance of mutual information estimation methods for quantifying the dependence among short and noisy data. *Physical Review E* 76, 2 (2007), 026209.

Justin B Kinney and Gurinder S Atwal. 2014. Equitability, mutual information, and the maximal information coefficient. *Proceedings of the National Academy of Sciences* 111, 9 (2014), 3354–3359.

Michal Kosinski, David Stillwell, and Thore Graepel. 2013. Private traits and attributes are predictable from digital records of human behavior. (2013). http://www.pnas.org/cgi/doi/10.1073/pnas.1218772110

G Kossinets and D Watts. 2006. Empirical Analysis of an Evolving Social Network. *Science* 311, 5757 (2006), 88–90.

Ravi Kumar, Jasmine Novak, and Andrew Tomkins. 2010. Structure and evolution of online social networks. In *Link Mining: Models, Algorithms, and Applications*. Springer, 337–357.

David Liben-Nowell and Jon Kleinberg. 2003. The Link Prediction Problem for Social Networks. In *Proceedings of the Twelfth International Conference on Information and Knowledge Management (CIKM '03)*. ACM, New York, NY, USA, 556–559. DOI:http://dx.doi.org/10.1145/956863.956972

P. Marks. 2006. Pentagon sets its sights on social networking websites. *New Scientist.com* (June 2006). http://www.newscientist.com/article/mg19025556.200?DCMP=NLC

Alan Mislove, Hema Swetha Koppula, Krishna P Gummadi, Peter Druschel, and Bobby Bhattacharjee. 2008. Growth of the flickr social network. In *Proceedings of the first workshop on Online social networks*. ACM, 25–30.

Alan Mislove, Massimiliano Marcon, P. Krishna Gummadi, Peter Druschel, and Bobby Bhattacharjee. 2007. Measurement and analysis of online social networks. In *Internet Measurement Comference*, Constantine Dovrolis and Matthew Roughan (Eds.). ACM, 29–42.

Alan Mislove, Bimal Viswanath, Krishna P Gummadi, and Peter Druschel. 2010. You are who you know: inferring user profiles in online social networks. In *Proceedings of the third ACM international conference on Web search and data mining*. ACM, 251–260.

Mark EJ Newman, Duncan J Watts, and Steven H Strogatz. 2002. Random graph models of social networks. *Proceedings of the National Academy of Sciences* 99, suppl 1 (2002), 2566–2572.

Alessandra Sala, Lili Cao, Christo Wilson, Robert Zablit, Haitao Zheng, and Ben Y Zhao. 2010. Measurement-calibrated graph models for social network experiments. In *Proceedings of the 19th international conference on World wide web*. ACM, 861–870.

C.E. Shannon. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal, The* 27, 3 (July 1948), 379–423. DOI:http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x

Bimal Viswanath, Alan Mislove, Meeyoung Cha, and Krishna P. Gummadi. 2009. On the evolution of user interaction in Facebook. In *Proceedings of the 2nd ACM workshop on Online social networks (WOSN '09)*. ACM, New York, NY, USA, 37–42. http://doi.acm.org/10.1145/1592665.1592675

Yang Wang, Gregory Norcie, Saranga Komanduri, Alessandro Acquisti, Pedro Giovanni Leon, and Lorrie Faith Cranor. 2011. "I regretted the minute I pressed share": a qualitative study of regrets on Facebook. In *Proceedings of the Seventh Symposium on Usable Privacy and Security (SOUPS '11)*. ACM, New York, NY, USA, Article 10, 16 pages. http://doi.acm.org/10.1145/2078827.2078841

Duncan J Watts and Steven H Strogatz. 1998. Collective dynamics of "small-world" networks. *nature* 393, 6684 (1998), 440–442.

Christo Wilson, Bryce Boe, Alessandra Sala, Krishna P.N. Puttaswamy, and Ben Y. Zhao. 2009. User interactions in social networks and their implications. In *Proceedings of the 4th ACM European conference on Computer systems (EuroSys '09)*. ACM, New York, NY, USA, 205–218. http://doi.acm.org/10.1145/1519065.1519089

Elena Zheleva and Lise Getoor. 2009. To join or not to join: the illusion of privacy in social networks with mixed public and private user profiles. In *Proceedings of the 18th international conference on World wide web*. ACM, 531–540.