Constrained Social-Energy Minimization for Multi-Party Sharing in Online Social Networks

Sarah Rajtmajer
Department of Mathematics
Pennsylvania State University
University Park, PA 16802
smr48@psu.edu

Anna Squicciarini
College of Information
Sciences and Technology
Pennsylvania State University
University Park, PA 16802
asquicciarini@ist.psu.edu

Christopher Griffin
Department of Mathematics
United States Naval Academy
Annapolis, MD 21402
griffinch@ieee.org

Sushama Karumanchi
College of Information
Sciences and Technology
Pennsylvania State University
University Park, PA 16802
sik5273@psu.edu

Alpana Tyagi
College of Information
Sciences and Technology
Pennsylvania State University
University Park, PA 16802
aft5121@psu.edu

ABSTRACT

The development of fair and practical policies for shared content online is a primary goal of the access control community. Multi-party access control, in which access control policies are determined by multiple users each with vested interest in a piece of shared content, remains an outstanding challenge. Purposeful or accidental disclosures by one user in an online social network (OSN) may have negative consequences for others, highlighting the importance of appropriate sharing mechanisms. In this work, we develop a game-theoretic framework for modeling multi-party privacy decisions for shared content. We assume that the content owner (*uploader*) selects an initial privacy policy that constrains the privacy settings of other users. We prove the convergence of users' access control policies assuming a multi-round consensus-building game in which all players are fully rational and investigate a variation of rational play that better describes user behavior and also leads to the rational equilibrium. Additionally, in an effort to better approximate human behavior, we study a bounded rationality model and simulate real user choices in this context. Finally, we validate model assumptions and conclusions using experimental data obtained through a study of 95 individuals in a mock-social network.

CCS Concepts

•**Security and privacy** → **Social network security and privacy;** *Economics of security and privacy; Social aspects of security and privacy;* •**Human-centered computing** → **Social content sharing; Social networks;** *Social media; Social tagging systems; Synchronous editors;* •**Theory of computation** → **Network games;** *Algorithmic game theory;*

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

Keywords

Multi-party access control, Collective sharing, Game theory, Bounded Rationality

1. INTRODUCTION

The ubiquity of online social networks (OSNs) and their expanding role in the day to day lives of user-members has led to a substantial amount of shared personal content, and along with it, increased exposure and privacy-related concerns. Although OSNs provide privacy control mechanisms, they typically adopt a discretionary approach, wherein content "uploaders" are also in charge of managing the content entirely and other stakeholders have no control over their access control decisions. For instance, when one user uploads a photo and tags friends who appear in the photo, tagged friends cannot restrict who can see this photo. Since multiple associated users may have different privacy concerns about a piece of shared content, lack of infrastructure for collaborative privacy control increases the potential risk of purposeful or accidental information leakage.

In addition, as noted by [20], government sectors are adopting social networks to exchange information and establish specialized groups/communities/task forces. These environments need to protect and control shared data due to its potential sensitivity and criticality.

Users face several potential dangers when over-sharing. Social media usage may lead to strained interpersonal relationships [31], decreased productivity [24] and susceptibility to negative feedback in disinhibited online communication [46], in addition to breeches of privacy. Users' perception of these and other risks inform their general comfort level with sharing personal information and content online. Individual behavior online, like individual behavior offline, is also subject to social norms and peer influence [18,21,33]. Notions of what is appropriate in content sharing online is defined comparatively, so that subtle shifts in local behavior may have much farther-reaching consequences for the network as a whole. In sum, unlike the SN site which is ultimately a business operating with a business model, users are individuals with more complex incentives, concerns and

considerations operating voluntarily within the constraints of the SN.

In this paper, we propose a social-energy minimizing game framework for modeling the privacy-related decisions of users in an OSN. We consider a "multi-party" access control decision making process, considering the case of a user/owner of a piece of shared content, and another set of users who have some relationship to that content and accordingly a personal stake in its access control settings. We establish a utility function representation of users' interests based on the core notions of inherent personal comfort and peer pressure.

The emphasis on multi-party access control represents a shift from the traditional approach taken in the access control community for two main reasons. First, the access control community has long investigated models and techniques to facilitate single subjects' access to resources according to well-defined, secure policies. Little, if any, attention has been given to group-driven access control decisions. The primary exception to this being secret sharing (see e.g. [40]). Second, the underlying goal has been to maintain confidentiality rather than facilitate controlled sharing. As such, the decisions offered by these mechanisms are single-user driven and often binary and based on inflexible policies. As others have noted [20], the inflexibility of binary decisions typically offered by current access control systems is a major inhibitor to information sharing when dealing with events in social networking sites. The lack of collective access control for resource sharing can threaten the protection of user data, including violating privacy expectations of content owners (and stakeholders) [12], due to the inability to determine whether a given disclosure meets the privacy expectations of the group.

Because ensuring "collective" access control requires taking into account multiple users' input, we tackle the problem of access control by formally modeling how users' access control decisions are made, and use those models to develop actionable and practical access control models that can be applied in a variety of real-world contexts. For example we consider which goals, reasons and influential dimensions affect users' privacy decisions.

The remainder of this paper is organized as follows: In Section 2 we provide a brief survey of related work to seat our results in the literature. We discuss our problem statement in Section 3. The model is formalized in Section 4, where results under rationality assumptions are given. In Section 5 we discuss a bounded rationality variation of our model and provide basic simulation results. In Section 6 we discuss experimental results from human trials. We conclude in Section 7.

2. RELATED WORK

There is a growing body of work on game-theoretic approaches to security [2,8,47] including leader-follower Stackelberg game models [34–36, 41, 49]. Stackelberg games have recently been used to model various security issues, with emphasis on classic attacker-defenders problems [36], although some limitations to the applicability of these games have been noted [51]. In particular, as noted by Pita et al., it is important to determine whether optimality assumptions hold when humans are involved in the decision making processes [34, 35]. Yet, recent contributions show that with proper modifications, Stackelberg games are suitable even with bounded rationality [35]. While this work does not use explicit Stackelberg formalisms, we do assume an initial decision constrains all players; this decision is not necessarily made with that fact in mind.

In this paper, we explore a social-energy minimization game for multi-party access control problems. With the exception of [17, 37] which explore the single-owner scenario, we are not aware of any work using game theory in this way to deal with access control problems.

This work is related in general to the body of work on game theory in social networks, both offline and online. Fundamental research efforts exploring cooperation in structured human populations include [32, 39, 48]. In the realm of online social networks, game theoretic models have been implemented for the study of the evolution of various social dilemmas, influence, bargaining, voting and deception [4, 14, 22, 25, 26, 38]. Most closely related to our work is the subset of this research concerning agent-based decision-making related to privacy and security in online social networks. Chen et al. model users' disclosure of personal attributes as a weighted evolutionary game and discuss the relationship between network topology and revelation in environments with varying level of risk [10]. Hu et al. tackle the problem of multi-party access control in [19], proposing a logic-based approach for identification and resolution of privacy conflicts. In [20] these authors extend this work, this time proposing adopting a game-theoretic framework, specifically a multi-party control game to model the behavior of users in collaborative data sharing in OSNs. The primary difference between our work and theirs is our relaxation of perfect rationality assumptions in the interest of a more realistic bounded rationality model of human behavior.

Also related to our work is the body of work on the economics of privacy. Researchers from many communities have noted the trade-off between privacy and utility (e.g., [6, 9, 28, 43, 45]). The majority of this prior work tends to view the privacy/utility trade-off as mutually exclusive: an increase in privacy (resp. utility) results in an immediate decrease in utility (resp. privacy). While this is certainly the case in some applications, (e.g. data anonymization [27]), it is not always such a straightforward relationship. In particular, the interplay of multiple users in any access control/privacy decision, in which privacy and utility are unevenly distributed among the players, and context-dependent, results in a more complicated relationship between these concepts.

Finally, our research overlaps with work on decision support systems (DSS) [5], and in particular, Group-centric and Model-driven DSS. Group-centric DSS focus on communication related activities of team members engaged in computer-supported cooperative work [11]. Some group-oriented work has emphasized the importance of optimizing group objectives, and developed standardized objectives for problem solving [29]. Model-driven DSS concentrate on formulation of quantitative (primarily business-related) problems (e.g., minimize stock or maximize revenue), and overlap to some extent with our envisioned approaches to modeling complex multi-party access control. However, DSS research has concentrated on how to solve group centric problems faster, especially through mixed-integer-linear programming, rather than on rigorous models and concrete case studies on access control (or security at large) [3, 5].

3. PROBLEM STATEMENT

Consider a social network (SN) site, wherein users share pieces of personal content freely within the network and possibly with selected subgroups of network users, according to a set of privacy settings for shared content made available by the site to its users. Examples of these settings in practice may include "visible to only me", "share with specific individuals", "share with friends", "share with my network" and "public". We abstract away from the details of how privacy options are presented by a site to its users, and map them to real values on the interval $[0, 1]$. In practice, the granularity of these options aims to be fine enough to meet users' needs, but coarse enough to be manageable in implementation for both the users and the SN site.

We focus here on the problem of assigning a privacy setting to a piece of content shared among multiple stakeholders. Assume one user in the network is the leading stakeholder, or equivalently, the poster of the content. This user chooses the privacy setting he is *most comfortable with* and posts the content to the site at this setting. This user constrains the action space of all subsequent users. Unlike in a Stackelberg game, we do not assume this user is necessarily maximizing an objective function, but simply imposing an upper-bound based on his/her comfort with data sharing.

In posting, we assume owner bounds the set of privacy options available for the content. In practice, this may mean that the owner lays out a discrete set of options from which followers may choose. Or, more simply, the leader sets an upper- or lower- bound on the sharing level for the given piece of content and followers choose from amongst options laid out by the SN site within those bounds. It may even be possible that the poster is not a personal stakeholder in the content and that he may choose not to constrain the space of privacy settings at all. The content stakeholders would then have the full range of available settings as options. From the standpoint of the model, these manifestations are equivalent as they leave followers in the position of jointly selecting one from amongst a fixed, finite set of access control settings. Following, we assume that the owner sets an upper-bound on sharing and remaining stakeholders have the ability to amend the privacy selection downward (toward more private options) but not to determine that the content should be shared more publicly.

Given this assumption, we assume all users' objective functions are composed of

1. a personal comfort term describing how comfortable they are with sharing a piece of content and
2. a peer pressure term making them more interested in coming to consensus with their peers. Peer pressure here also loosely includes the trust level users have toward each other when sharing content and establishing privacy settings.

Objective functions are expressed in terms of *energy* where a user seeks a minimum (social) energy state at any given time. The intuition is that users have an inherent degree of disclosure they feel most comfortable with, but are also influenced by their peers when making sharing decisions.

Since these two dimensions may not be considered equally for all users, we introduce weights to capture interpersonal differences in susceptibility to peer pressure. Precisely, we offer the option of including weights on either the peer pressure or personal comfort components of the user's utility function allowing customization of the model for non- homogeneous users and an opportunity to strengthen the model in the presence of additional information on user behavior, which the site may learn through observation. An overview of the general problem is provided in Figure 1.

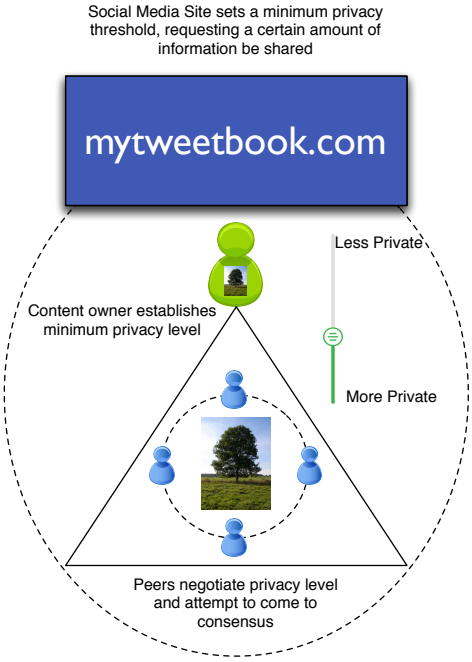

Figure 1: A representation of the multi-party access control problem for collective sharing in OSNs.

4. FULLY RATIONAL MODEL

Assume a SN is represented by a graph $G = (V, E)$, where V is a set of users (represented by vertices in the graph) and E is the set of social connections (edges) between them. For the remainder of this paper, assume $|V| = N$. The poster-leader (henceforth, poster) shares a piece of content on the SN site with privacy setting x_0. The other user stakeholders (henceforth, users) engage in an iterative process of determining the final privacy setting of the content x^*. At round k:

1. Each user i chooses value $x_i^k \leq x_0$.
2. If $|x_i^k - x_j^k| \leq \epsilon$ for all pairs i, j, then iteration stops and $x^* = \bar{x}$. For $\epsilon = 0$, iteration stops when the users have converged on a common value.

For a fixed piece of content, we define the optimality privacy selection of each user in terms of:

1. Peer Pressure (and Reputation)
2. Comfort level

Comfort level in the context of privacy and information disclosure refers to the degree of disclosure users feel comfortable with. This notion, often used to characterize information sharing in online sites (e.g. [1, 13]), is also adopted in our model. Users reaching their optimal comfort level wish not to change any of their information sharing practices. Previous work has indicated that reputation are highly correlated [44], and so we consider peer pressure here as a single

measure representing the social considerations of an individual user with respect to sharing.

To model peer pressure, we assume that individuals are encouraged to behave in accordance with the norms of their social group. In particular as stakeholders in the same piece of shared content, we assume that there is an additional pressure to come to agreement and select the privacy setting for the respective content. We can assume that if a consensus is not reached by the users some default setting is selected with which no one is happy (e.g., the content is not published to the site).

We assume that as time passes without consensus, the peer pressure felt by all members to come to agreement increases [7]. For User i, we define, let the peer-pressure function be:

$$P_i(x_i, \mathbf{x}_{-i}, k) = \rho^{(k)} \cdot \sum_{j \in N(i)} f_P(x_i - x_j). \quad (1)$$

Here f_P is a differentiable quasi-convex function with minimum at 0, $\mathbf{x}_{-i}$ is the vector of strategies of all other users in the network (other than User i), and $N(i)$ is the immediate neighborhood of User i in a social network. The coefficients increase monotonically so that $1 \leq \rho^{(1)} \leq \rho^{(2)} \leq \rho^{(3)} \leq \cdots$, thus increasing the influence of peer pressure. This is notationally consistent with [16].

Let the personal comfort function for User i be:

$$C_i(x_i) = f_C(x_i - x_i^+). \quad (2)$$

Here, x_i^+ is the preferred privacy level for a specific piece of user content according to User i and f_C is a differentiable quasi-convex function with minimum at 0. Since the privacy level is constrained by the content owner, we assume that if x_i^+ is more permissive than this level, then x_i^+ is reset to the constrained level set by the content owner. This assures that all privacy decisions will force the information exposure level downward (toward a more private choice).

In practice x_i^+ may be difficult to determine for an unknown User i. However, we assume that based on user demographics, as well as observed overall user behavior for a mass of users, either at the individual or group level, it is possible to infer of x_i^+, or at least an expected value $E[x_i^+]$ within a tolerated window of error.

Thus, the total objective function for User i is:

$$J_i(x_i, \mathbf{x}_{-i}, k) = P_i(x_i, \mathbf{x}_{-i}, k) + C_i(x_i) = \rho^{(k)} \sum_{j \in N(i)} f_P(x_i - x_j) + f_C(x_i - x_i^+). \quad (3)$$

Proposition 1 (Rosen 1963)**.** *If f_P and f_C are convex and $x_i \in \Omega_i$, where Ω_i is a convex set for $i = 1, \ldots, N$, then there is an equilibrium solution $x_1^*, \ldots, x_N^*$.* □

Rosen's result informs us that there is a (Nash) equilibrium assuming simultaneously play, it neither ensures the uniqueness of the equilibrium, nor does it ensure flocking or consensus behavior; i.e., that $|x_i - x_j| < \epsilon$ for some small ϵ.

For the case when $f_P(z) = f_C(z) = z^2$ and we have a complete graph, it is relatively easy to see that consensus occur for some value $\rho^{(k)}$. To see this, note that at equilibrium, each agent must satisfy the equations (derived the first order necessary conditions of optimality):

$$\rho^{(k)} \sum_{j \neq i} (x_i - x_j) + (x_i - x_i^+) = 0 \quad \forall i \quad (4)$$

This system of equations admits the solution:

$$x_i^{(k)} = \left[\frac{\rho^{(k)}}{1 + N\rho^{(k)}} \sum_{j \neq i} x_j^+ \right] + \left[\frac{(1 + \rho^{(k)}) x_i^+}{1 + N\rho^{(k)}} \right] \quad (5)$$

The following is immediately clear:

Proposition 2. *Suppose $\rho^{(1)} < \rho^{(2)} < \cdots < \rho^{(k)} < \cdots$ and $\lim_{k \to \infty} \rho^{(k)} = \infty$, then:*

$$\lim_{k \to \infty} x_i^* = \frac{1}{N} \sum_j x_j^+ \quad (6)$$

Therefore, as long as the *pressure* to come to consensus increases on each round, the players converge to their mean comfort levels.

Corollary 3. *For all $\epsilon > 0$ there exists a $K > 0$ so that for all $k > K$, $|x_i^{(k)} - x_j^{(k)}| < \epsilon$ for all i, j. Thus the system converges after a finite number of rounds.*

If all players realize this, then rational play at each round results in behavior equivalent to a one shot game where users intuitively *know* a K that ensures fast convergence.

An alternate update rule, and more in keeping with human behavior is to assume a *recency model* wherein individual users take their peers' strategy from the last round of play and construct a weighted average based on this value rather than their peers' initial selections. Specifically, the objective function for User i is then:

$$J_i(x_i, \mathbf{x}_{-i}, k) = \rho^{(k)} \sum_{j \in N(i)} f_P(x_i - x_j^{(k-1)}) + f_C(x_i - x_i^+). \quad (7)$$

In this case:

$$x_i^{(k)} = \frac{\rho^{(k)} \sum_{j \neq i} x_j^{(k-1)}}{1 + (N-1)\rho^{(k)}} + \frac{x_i^+}{1 + (N-1)\rho^{(k)}} \quad (8)$$

with $x_i^{(0)} = x_i^+$ for all i. We show that this play also results in the rational Nash equilibrium. Let $F_k : [0,1]^N \to [0,1]^N$ be the function whose i^{th} component is given by Equation 8. Then: F_k has at least one fixed point by Brouwer's Fixed Point theorem and we can show that:

$$x_i^*(k) = \frac{x_i^+ + \rho^{(k)} \sum_j x_j^+}{1 + N\rho^{(k)}} \quad (9)$$

is a (unique) fixed point of F arising from the solution of N linear fixed point equations.

Lemma 4. *$F_k : [0,1]^N \to [0,1]^N$ is a contraction for $N \geq 4$.*

Proof. For notational simplicity, $\rho = \rho^{(k)}$ and let $\mathbf{x}, \mathbf{y} \in [0,1]^N$ where (e.g.) $\mathbf{x} = \langle x_1, \ldots, x_N \rangle$. Then:

$$\frac{|F_k(\mathbf{x}) - F_k(\mathbf{y})|^2}{|\mathbf{x} - \mathbf{y}|^2} = \frac{\rho^2}{(1 + (N-1)\rho)^2} \frac{\sum_i \left(\sum_{j \neq i} (x_j - y_j) \right)^2}{\sum_i (x_i - y_i)^2} \quad (10)$$

Let:

$$\beta = \frac{\rho^2}{(1 + (N-1)\rho)^2} \quad (11)$$

Then we can re-write Expression 10 as:

$$\beta\frac{(N-1)\sum_i(x_i-y_i)^2+N\sum_i\sum_{j>i}(x_i-y_i)(x_j-y_j)}{\sum_i(x_i-y_i)^2}=$$
$$\beta\left((N-1)+\frac{N\sum_i\sum_{j>i}(x_i-y_i)(x_j-y_j)}{\sum_i(x_i-y_i)^2}\right)$$

The expression

$$(N-1)+\frac{N\sum_i\sum_{j>i}(x_i-y_i)(x_j-y_j)}{\sum_i(x_i-y_i)^2}$$

has maximum value when $x_i - y_i = \frac{1}{2}$ for all i (this can be proved using the Karush-Kuhn-Tucker conditions on an ancillary maximization problem). Evaluating at this point we have:

$$(N-1)+\frac{N\sum_i\sum_{j>i}(x_i-y_i)(x_j-y_j)}{\sum_i(x_i-y_i)^2}\leq$$
$$(N-1)+\frac{N(N-1)}{2}=(N-1)\left(1+\frac{N}{2}\right)$$

Thus:

$$\frac{|F_k(\mathbf{x})-F_k(\mathbf{y})|^2}{|\mathbf{x}-\mathbf{y}|^2}\leq$$
$$\frac{\rho^2(N-1)\left(1+\frac{N}{2}\right)}{(1+(N-1)\rho)^2}<\frac{\rho^2(N-1)\left(1+\frac{N}{2}\right)}{((N-1)\rho)^2}=$$
$$\frac{1+N/2}{N-1} \quad (12)$$

For $N \geq 4$, $(1+N/2)/(N-1) \leq 1$. Since the inequality is strict in Expression 12, it follows F_k is a contraction for $N \geq 4$. □

The following is an immediately consequence of the lemma and the Banach Fixed Point Theorem.

Corollary 5. *The fixed point given by Expression 9 is unique and is the contraction point for F_k.*

Remark 1. In reality, it may be the case we can do better than $N \geq 4$ and show that F_k is a contraction for $N \geq 1$. Certainly for $N = 1$, F_k is just a constant function and thus a contraction.

The following lemma is clear from the work done so far:

Lemma 6. *The following holds:*

$$\lim_{k\to\infty} x_i^*(k)=\bar{x}^+=\frac{1}{N}\sum_j x_j^+$$

further convergence of the sequence of fixed points $x^*(k)$ to $\bar{x}^+$ is uniform.

It remains to show this is an attracting fixed point of the recurrence relations. From Theorem 2 of [15] and the fact that the F_k are contractions, we have:

Theorem 7. *Let $F_0(\mathbf{x})$ return $\langle x_1^+,\ldots,x_N^+\rangle$. If $G_k(\mathbf{x}) = F_k \circ F_{k-1} \circ \cdots \circ F_0(\mathbf{x})$ and $N \geq 4$, then:*

$$G=\lim_{k\to\infty} G_k \quad (13)$$

is a constant function whose value is the limit of the fixed points of F_k. Furthermore, convergence of the function sequence $\{G_k\}_k$ is uniform.

Having shown that the recency update rule with increasing peer pressure leads to a constant fixed point and that each individual's privacy choice approaches this fixed point, we have the following corollary:

Corollary 8. *Assume the latency updating rule and $N \geq 4$. For all $\epsilon > 0$, there exists a $K > 0$ so that if $k > K$, then $|x_i^{(k)} - x_j^{(k)}| < \epsilon$. Thus the system converges after a finite number of rounds under the latency rule.*

5. BOUNDED RATIONALITY

In practice, as users negotiate sharing levels at each iteration of the game, they express their preferences not in terms of the continuous value $x_i^*(t)$ but as a selection among the discrete set of options $\{l_1,\ldots,l_n\}$. In practice these discrete values map to settings specifying the audience for posted content, e.g., "friends" or "my network".

We could expect that perfectly rational users, observing complete information without error will, at any given time t, choose the option l_i closest to $x_i^{(k)}$. That is, each user will select $l_i = \min_{\{l_1,\ldots,l_n\}} |l_i - x_i^{(k)}|$.

However, [23, 42] suggests that realistic models of human behavior should relax perfect-rationality assumptions in favor of *bounded rationality* models accounting for limited time and information, as well as cognitive limitations. Let:

$$U_{i,j}^{(k)}=J_i\left(l_j,\mathbf{x}_{-i}^{(k)}\right).$$

This is the social stress experienced by User i making choice l_j when all other players make choice $\mathbf{x}_{-i}^{(k)}$ at round k.

One well-established model of bounded rationality is the Quantal Response model [30], which gives the probability that user i will choose setting l_j at epoch k:

$$q_i^{(k)}(\lambda|\mathbf{x})=\frac{\exp\left(-\lambda U_{i,j}^{(k)}\right)}{\sum_r\exp\left(-\lambda U_{i,r}^{(k)}\right)}$$

Recent work has indicated that a Quantal Response model is a promising model of human behavior in the context of security games [50]. In our consensus-building formulation, the probabilities that each player will select each option converge to a single probability distribution (see Proposition 9) under the assumption of *rational play*.

Proposition 9. *Suppose $\rho^{(1)} < \rho^{(2)} < \cdots < \rho^{(k)} < \cdots$ and $\lim_{k\to\infty}\rho^{(k)} = \infty$. If users engage in rational play, then there is a probability distribution with support $\{l_1,\ldots,l_k\}$, $q^*(\lambda)$ so that:*

$$\lim_{k\to\infty} q_i^{(k)}(\lambda|\mathbf{x})=q^*(\lambda). \quad (14)$$

Proof. As $k \to \infty$, $x_i^* = x_j^*$ for all i, j. As a result, $U_{i,r}^{(k)} = U_{j,r}^{(k)}$ for all i, j. Thus, $q_i^{(k)}(\lambda|\mathbf{x}) = q_j^{(k)}(\lambda|\mathbf{x}) = q^*(\lambda)$. □

Given perfectly rational players, that is as $\lambda \to \infty$, this distribution will tend to 1 for the choice nearest to $x_i^*(t)$ and 0 everywhere else. In this case, players are guaranteed to come to agreement on an access control policy. For decreasing values of λ and more uniform corresponding probability distributions, players' actual selections are non-deterministic and therefore non-convergent. This is *particularly* true when users engage in the recency model of play in which previous

players' decisions are weighted against the users' preferred privacy level.

To gain a better understanding of the consensus-building game with bounded rationality, we simulated play among players in various population sizes for a varying rationality parameter λ.

Specifically, consider a game among human users acting with bounded rationality and using the recency model. The uploader shares a piece of content with privacy setting x_0 and the remaining stakeholders (followers) determine their individual $x_i^k \leq x_0$ according the user objective function (3). We (randomly) assign users a discrete choice from the fixed set of privacy settings at each iteration of play based on the probabilities given by the Quantal Response model (5). We iterate this procedure until we obtain convergence of the expected value of each player's strategy vector within some small interval ϵ. We consider games with $3, 4, 5, 6, 7, 8, 9$ and 10 players, and vary $0 \leq \lambda \leq 10$. Simulations are run 100 times for each population size/lambda pairing, and averages are reported.

Of primary interest is the expected value of the final probability vector q^*. Figure 2 shows a histogram of the average differences $\frac{1}{N}\sum_j E[q_j] - \frac{1}{N}\sum_j x_j^+$ obtained for simulations run for each population size/lambda pairing (88 total). We observe that $E[q^*] \to \bar{x}^+$. This result is consistent with Equations 6 and 9, despite using the recency update rule with bounded rationality. We will make this result formal, proving convergence to the mean comfort in future work.

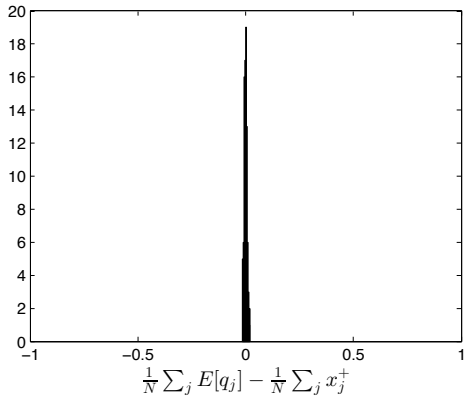

Figure 2: A histogram of average difference between average expected value and average comfort obtained over 100 simulations for 88 parameter pairs.

Despite the convergence of $E[q_i]$ for all players, we have argued that in the presence of bounded rationality, each user's actual choice from among the discrete set of options will vary depending on λ. That is, less rational actors are more likely to deviate from optimal play. Formally,

$$\lim_{\lambda \to 0} q_i^{(k)} = \frac{1}{n}$$

for $l_j \in \{l_1, \ldots, l_n\}$. This relationship between decreasing lambda and increasing variance is illustrated in Figure 3 for our simulations.

An additional practical consideration for real applications of multi-party access control is the number of iterations of revisions required for group consensus, or at least sufficient convergence of $E[q_i]$. Figure 4 indicates the number of rounds of play until convergence within an interval of $\epsilon = 0.1$. Here the number of iterations is increasing in lambda, respecting the greater difficulty in aligning the expected values of non-uniform probability vectors. In other words, more rational users or users more acutely aware of their optimal choice may take longer to reach consensus. Access control decisions for a single piece of shared content in an OSN are typically not subject to multiple revisions, but are rather posted by the uploader and potentially revised once (e.g., a user may remove a tag of him or herself in an image or remove a tagged image from his or her profile page, newsfeed, or similar). However, we suggest that multiple revisions may be considered an abstraction of a process which spans the course of multiple instances of collective sharing decisions for a given group of users. Similarly, we can interpret ϵ as the granularity of a privacy policy, where a wider set of options would require a smaller ϵ for group consensus in real-life scenarios.

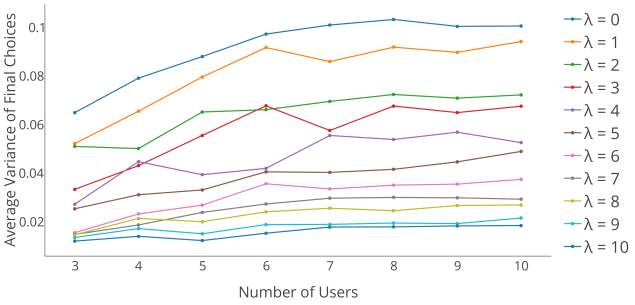

Figure 3: A plot of the average statistical variance of users' final selected privacy settings, obtained over 100 simulations for given number of players and rationality parameter λ.

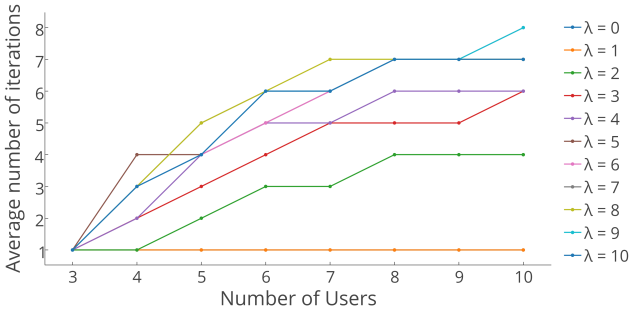

Figure 4: Average number of iterations of play until group consensus within $\epsilon = 0.1$, obtained over 100 simulations for given number of players and rationality parameter λ.

6. EXPERIMENTAL RESULTS

In order to validate the mathematical results discussed for multi-party privacy decisions, we collected real-world data

on privacy setting selections in a small group setting. Our small groups (of 3 or 4 individuals) represent a set of user-stakeholders determining access control settings for a piece of shared content, within a set of constraints, namely given a fixed set of privacy options.

Specifically, we created a fictitious SN site using Drupal as a backbone and invited users to generate a social group within the site. Participants were asked to work in teams of 3 or 4 to jointly set up fictitious privacy options for a newly formed social group, and select, for three different content types (images, documents, posts/event notifications) the privacy settings the group should offer as well as default settings for each content type.

Once the available privacy options were established, participants were asked to individually set their privacy preferences for each piece of content shown to them. All users within the group simultaneously entered their preferences for that piece of content. As individual selections were made, they were displayed to the rest of the members of the group. Upon observation of their peers' selected privacy settings, group members were allowed to revise their options as desired, for as many iterations as they chose. Users were encouraged to come to a consensus, being told that if they did not, the content would not be shared at all.

Monitoring the iterated individual revisions of selected privacy settings, we compare the influence of users' peers on individual behavior with the behavior anticipated by our theoretical results. In particular, we aim to determine whether the mean-consensus model (6,9) well-describes user observed behavior.

Overall, we had 95 participants (74% females and 26% males). Participants were college-aged students, and reported a high frequency of use of SNing sites (87% declared accessing SN once or more a day). Further, 93% of responders indicated changing their privacy settings on any of their profile items at least once, therefore indicating some degree of privacy awareness. We partitioned participants into 25 fictional social groups and recorded each group's data for 9 pieces of content - 3 photos, 3 blog posts and 3 events.

Each available privacy setting (specifically, "everyone", "friends", "colleagues", "family", "self") is mapped into the interval $[0, 1]$, where 0 indicates most restrictive ("only self") and 1 least restrictive ("everyone"). Each team represents a small complete friendship graph on three or four nodes. For a given piece of content, let the initial privacy setting selection for User i represent his comfort level x_i^+, since this selection is made before any information is obtained about peers' preferences.

As User i receives feedback from his peers, we argue that the iterative revisions he makes to his own preferences should tend toward $x_i^* = \frac{1}{N}\sum_j x_j^+$ if his objective function follows similarly to one we have proposed.

Consider each group's decision process for each piece of content as negotiation amongst followers to come to agreement on an access control policy. Amongst 225 of these games represented in our dataset, 109 involved at least one user revision. In 85 of these 109, the revision process led to group consensus. If we assume that each user's initial privacy setting, before seeing his peers' selections, represents his inherent comfort level for sharing that particular content, we expect that the final group choice is equal to the mean of the users' initial selections (6). Let x^* be the final collaborative access control decision of a particular group for a particular piece of content. Figure 5 gives a histogram of values $x^* - \frac{1}{N}\sum_j x_j^+$ for the 85 games in which consensus was reached after revision. As expected, these center very near 0, with mean 0.006 and standard deviation 0.099. Notice Figure 5 is qualitatively similar to Figure 2. In future

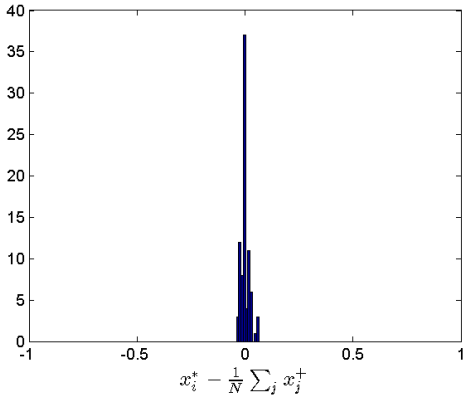

Figure 5: A histogram of values $x_i^* - \frac{1}{N}\sum_j x_j^+$ for 85 pieces of shared content.

work, we plan to collect more data and compare these two results using (e.g.) a χ^2 Goodness of Fit test to determine whether the simulated user behaviors provides a reasonable model of true human behavior.

7. CONCLUSION

We present a behavioral game-theoretic model of multi-party access control in OSNs. Specifically, we address the problem where one user shares a piece of content online and other users with a personal stake in the given content respond within the constraints established by the uploader. We discuss the convergence of user decisions given a consensus-building game with fully-rational users, and simulate play with a quantal response model to better approximate the bounded rationality of human agents. Results of experimental validation with real users in a mock-social network indicate that these models closely approximate real user behavior.

We suggest that this work meets a growing need for foundational research to understand and facilitate the privacy requirements of multiple users for collaboratively managing shared data in OSNs, and may serve as the basis for extended validation. Future work includes incorporating asymmetrical peer pressure, alternate rationality parameters, penalties for iterated revision, and dynamic comfort which may represent memory and allow users to learn from previous games.

The model we have developed here does not rely on the structure of the OSN and we propose is generalizable in the context of collaborative sharing more widely. Other variations of this model may need to consider context-dependent amendments to the peer pressure function, provided the nature of peer pressure remains quantifiable and can be assumed to be increasing over time with lack of consensus.

Acknowledgments

Portions of Dr. Griffin's, Dr. Squicciarini's and Dr. Rajtmajer's work were supported by the Army Research Office under grant W911NF-13-1-0271. Portions of Dr. Squicciarini's work were additionally supported by a National Science Foundation grant 1453080.

REFERENCES

[1] M. S. Ackerman, L. F. Cranor, and J. Reagle. Privacy in e-commerce: Examining user scenarios and privacy preferences. In *Proceedings of the 1st ACM Conference on Electronic Commerce*, EC '99, pages 1–8, New York, NY, USA, 1999. ACM.

[2] B. An, D. Kempe, C. Kiekintveld, E. Shieh, S. Singh, M. Tambe, and Y. Vorobeychik. Security games with limited surveillance. *Ann Arbor*, 1001:48109, 2012.

[3] D. Arnott and G. Pervan. Eight key issues for the decision support systems discipline. *Decision Support Systems*, 44(3):657–672, 2008.

[4] L. Backstrom, D. Huttenlocher, J. Kleinberg, and X. Lan. 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*, pages 44–54. ACM, 2006.

[5] H. K. Bhargava, D. J. Power, and D. Sun. Progress in web-based decision support technologies. *Decision Support Systems*, 43(4):1083–1095, 2007.

[6] B. Bhumiratana and M. Bishop. Privacy aware data sharing: balancing the usability and privacy of datasets. In *Proceedings of the 2nd International Conference on PErvasive Technologies Related to Assistive Environments*, page 73. ACM, 2009.

[7] R. Böhme and S. Pötzsch. Collective exposure: Peer effects in voluntary disclosure of personal data. In *Proceedings of the 15th International Conference on Financial Cryptography and Data Security*, FC'11, pages 1–15, Berlin, Heidelberg, 2012. Springer-Verlag.

[8] M. Brown, B. An, C. Kiekintveld, F. Ordóñez, and M. Tambe. An extended study on multi-objective security games. *Autonomous Agents and Multi-Agent Systems*, 28(1):31–71, Jan. 2014.

[9] A. Brush, J. Krumm, and J. Scott. Exploring end user preferences for location obfuscation, location-based services, and the value of location. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 95–104. ACM, 2010.

[10] J. Chen, M. R. Brust, A. R. Kiremire, and V. V. Phoha. Modeling privacy settings of an online social network from a game-theoretical perspective. In *Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom), 2013 9th International Conference Conference on*, pages 213–220. IEEE, Oct. 2013.

[11] Y.-L. Chen, L.-C. Cheng, and C.-N. Chuang. A group recommendation system with consideration of interactions among group members. *Expert systems with applications*, 34(3):2082–2090, 2008.

[12] P. C. Cheng, P. Rohatgi, C. Keser, P. A. Karger, G. M. Wagner, and A. S. Reninger. Fuzzy multi-level security: An experiment on quantified risk-adaptive access control. In *IEEE Symposium on Security and Privacy.*, pages 222–230. IEEE, 2007.

[13] L. F. Cranor, J. Reagle, and M. S. Ackerman. *Beyond concern: Understanding net users' attitudes about online privacy*. Cambridge, MA: MIT Press, 2000.

[14] F. Fu, C. Chen, L. Liu, and L. Wag. Social dilemmas in an online social network: The structure and evolution of cooperation. *Physics Letters A*, 371(1-2):58–64, 2007.

[15] J. Gill. The use of the sequence $f_n(z) = f_n \circ \cdots \circ f_1(z)$ in computing fixed points of continued fractions, products, and series. *Applied Numerical Mathematics*, 8(6):469 – 476, 1991.

[16] C. Griffin, A. Squicciarini, S. Rajtmajer, M. Tentilucci, and S. Li. Site-constrained privacy options for users in social networks through stackelberg games. In *Proc. of Sixth ASE International Conference on Social Computing. May 2014.*, 2014.

[17] C. Griffin, A. C. Squicciarini, S. Rajtmajer, M. Tentilucci, and S. Li. Site-constrained privacy options for users in social networks through stackelberg games. In *Sixth ASE International Conference on Social Computing (SocialCom)*, 2014.

[18] R. Gross and A. 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.

[19] H. Hu, G.-J. Ahn, and J. Jorgensen. Multiparty access control for online social networks: model and mechanisms. *Knowledge and Data Engineering, IEEE Transactions on*, 25(7):1614–1627, 2013.

[20] H. Hu, G.-J. Ahn, Z. Zhao, and D. Yang. Game theoretic analysis of multiparty access control in online social networks. In *Proceedings of the 19th ACM Symposium on Access Control Models and Technologies*, SACMAT '14, pages 93–102, New York, NY, USA, 2014. ACM.

[21] P. Hui and S. Buchegger. Groupthink and Peer Pressure: Social Influence in Online Social Network Groups. In *2009 International Conference on Advances in Social Network Analysis and Mining (ASONAM)*, pages 53–59, Los Alamitos, CA, USA, July 2009. IEEE.

[22] N. Immorlica, B. Lucier, and B. Rogers. Emergence of cooperation in anonymous social networks through social capital. In *In Proceedings of the 11th ACM Conference on Electronic Commerce (EC*, 2010.

[23] D. Kahneman. Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, 93(5):1449–1475, 2003.

[24] P. A. Kirschner and A. C. Karpinski. Facebook® and academic performance. *Computers in Human Behavior*, 26(6):1237–1245, Nov. 2010.

[25] J. Kleinberg and E. Tardos. Balanced outcomes in social exchange networks. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, STOC '08, pages 295–304, New York, NY, USA, 2008. ACM.

[26] J. M. Kleinberg. Challenges in mining social network data: processes, privacy, and paradoxes. In *KDD '07: Proceedings of the 13th ACM SIGKDD international*

conference on Knowledge discovery and data mining, pages 4–5, New York, NY, USA, 2007. ACM.

[27] K. LeFevre, D. J. DeWitt, and R. Ramakrishnan. Mondrian multidimensional k-anonymity. In *Proceedings of the 22nd International Conference on Data Engineering, ICDE'06*, pages 25–25. IEEE, 2006.

[28] K. Liu and E. Terzi. A framework for computing the privacy scores of users in online social networks. *ACM Trans. Knowl. Discov. Data*, 5(1):6:1–6:30, Dec. 2010.

[29] J. Lu, G. Zhang, and D. Ruan. *Multi-objective group decision making: methods, software and applications with fuzzy set techniques*. Imperial College Press, 2007.

[30] R. D. McKelvey and T. R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6 – 38, 1995.

[31] A. Muise, E. Christofides, and S. Desmarais. More Information than You Ever Wanted: Does Facebook Bring Out the Green-Eyed Monster of Jealousy? *CyberPsychology & Behavior*, 12(4):441–444, Aug. 2009.

[32] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502–505, May 2006.

[33] N. Park, K. Kee, and S. Valenzuela. Being Immersed in Social Networking Environment: Facebook Groups, Uses and Gratifications, and Social Outcomes. *CyberPsychology & Behavior*, 12(6):729–733, Dec. 2009.

[34] J. Pita, M. Jain, F. Ordóñez, M. Tambe, S. Kraus, and R. Magori-Cohen. Effective solutions for real-world stackelberg games: When agents must deal with human uncertainties. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 369–376. International Foundation for Autonomous Agents and Multiagent Systems, 2009.

[35] J. Pita, M. Jain, M. Tambe, F. Ordóñez, and S. Kraus. Robust solutions to Stackelberg games: Addressing bounded rationality and limited observations in human cognition. *Artificial Intelligence*, 174(15):1142–1171, 2010.

[36] J. Pita, M. Tambe, C. Kiekintveld, S. Cullen, and E. Steigerwald. Guards: game theoretic security allocation on a national scale. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 37–44. International Foundation for Autonomous Agents and Multiagent Systems, 2011.

[37] S. Rajtmajer, C. Griffin, D. Mikesell, and A. Squicciarini. An evolutionary game model for the spread of non-cooperative behavior in online social networks. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, SAC '15, pages 1154–1159, New York, NY, USA, 2015. ACM.

[38] S. M. Rajtmajer, C. Griffin, D. Mikesell, and A. Squicciarini. A cooperate-defect model for the spread of deviant behavior in social networks. *CoRR*, abs/1408.2770, 2014.

[39] D. G. Rand, S. Arbesman, and N. A. Christakis. Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Sciences*, 108(48):19193–19198, 2011.

[40] A. Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, Nov. 1979.

[41] Y. Sharma and D. P. Williamson. Stackelberg thresholds in network routing games or the value of altruism. *Games and Economic Behavior*, 67(1):174–190, 2009.

[42] H. Simon. A behavioural model of rational choice. In H. Simon, editor, *Models of man: social and rational; mathematical essays on rational human behavior in a social setting*, pages 241–260. J. Wiley, New York, 1957.

[43] D. J. Solove. *Nothing to hide: The false tradeoff between privacy and security*. Yale University Press, 2011.

[44] A. Squicciarini and C. Griffin. An informed model of personal information release in social networking sites. In *2012 ASE/IEEE Conference on Privacy, Security, Risk and Trust*, Amsterdam, Netherlands, September 2012.

[45] F. Stutzman and W. Hartzog. Boundary regulation in social media. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, pages 769–778. ACM, 2012.

[46] J. Suler. The Online Disinhibition Effect. *Cyberpsychology & Behavior*, 7(3):321–326, June 2004.

[47] M. Tambe. *Security and game theory: algorithms, deployed systems, lessons learned*. Cambridge University Press, 2011.

[48] Z.-X. Wu, Z. Rong, and H.-X. Yang. Impact of heterogeneous activity and community structure on the evolutionary success of cooperators in social networks. *Phys. Rev. E*, 91:012802, Jan 2015.

[49] R. Yang, F. Fang, A. X. Jiang, K. Rajagopal, M. Tambe, and R. Maheswaran. Designing better strategies against human adversaries in network security games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pages 1299–1300. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[50] R. Yang, C. Kiekintveld, F. Ordóñez, M. Tambe, and R. John. Improving resource allocation strategies against human adversaries in security games: An extended study. *Artificial Intelligence*, 195:440 – 469, 2013.

[51] Z. Yin, D. Korzhyk, C. Kiekintveld, V. Conitzer, and M. Tambe. Stackelberg vs. Nash in security games: Interchangeability, equivalence, and uniqueness. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 1139–1146. International Foundation for Autonomous Agents and Multiagent Systems, 2010.