A Game Theoretic Framework for Data Privacy Preservation in Recommender Systems

Maria Halkidi[1] and Iordanis Koutsopoulos[2]

[1] Dept. of Digital Systems, University of Piraeus
mhalk@unipi.gr
[2] Dept. of Computer and Communication Engineering,
University of Thessaly and CERTH
jordan@uth.gr

Abstract. We address the fundamental tradeoff between privacy preservation and high-quality recommendation stemming from a third party. Multiple users submit their ratings to a third party about items they have viewed. The third party aggregates the ratings and generates personalized recommendations for each user. The quality of recommendations for each user depends on submitted rating profiles from all users, including the user to which the recommendation is destined. Each user would like to declare a rating profile so as to preserve data privacy as much as possible, while not causing deterioration in the quality of the recommendation he would get, compared to the one he would get if he revealed his true private profile.

We employ game theory to model and study the interaction of users and we derive conditions and expressions for the Nash Equilibrium Point (NEP). This consists of the rating strategy of each user, such that no user can benefit in terms of improving its privacy by unilaterally deviating from that point. User strategies converge to the NEP after an iterative best-response strategy update. For a hybrid recommendation system, we find that the NEP strategy for each user in terms of privacy preservation is to declare false rating only for one item, the one that is highly ranked in his private profile and less correlated with items for which he anticipates recommendation. We also present various modes of cooperation by which users can mutually benefit.

Keywords: privacy preservation, recommendation systems, game theory.

1 Introduction

The need for expert recommender systems becomes ever increasing in our days, due to the massive amount of information and abundance of choices available for virtually any human action involving decision making, which is connected explicitly or implicitly to the Internet. Recommender systems arise with various contexts, from providing personalized search results and targeted advertising, to making social network related suggestions, up to providing personalized suggestions on various goods and services. Internet users become more and more

D. Gunopulos et al. (Eds.): ECML PKDD 2011, Part I, LNAI 6911, pp. 629–644, 2011.

dependent on efficient recommender systems in order to expedite purchase of goods, selection of movies, places to dine and spend their vacation, and even to decide whom to socialize with and date. In general, users rely on recommendation systems so as to obtain quick and accurate personalized expert advice and suggestions, which will aid them in decision making. While the final decision at the user end depends on various psychological and subjective factors, it is beyond doubt that recommendation systems will enjoy accelerating penetration to users.

The efficiency of a recommender system amounts to high-quality personalized recommendations it generates for different users. Recommendation systems are fundamentally user-participatory. The quality of recommendations for an individual user relies on past experience and participation of other users in the rating process. Furthermore, even if not immediately apparent, each individual user can to a certain extent affect the quality of recommendations for himself by his own ratings about what he has experienced.

To see this, consider the following simple example. Suppose there exist two users, 1 and 2. User 1 has viewed and rated item A, while user 2 has viewed and rated item B. Suppose that the recommendation system recommends item B to user 1 if a certain metric exceeds a threshold, otherwise it does not. This metric will depend on (i) how high was the rating of user 2 for item B, (ii) how similar is item B to A, (iii) how high was the rating of user A for item 1. Clearly, *whether or not item B will be recommended to user 1 depends on the ratings of both users for the items they have viewed.*

Since recommendation systems involve data exchange between the users and the third party that performs recommendations, *privacy concerns of users are inescapable.* Users prefer to preserve their privacy by not revealing much information to the third party about their private personal preferences and ratings. On the other hand, users would like to receive high-quality recommendation results. Namely, the recommendation they would get as a result of not declaring their true private ratings should be as close as possible to the one they would get if they revealed their private data. In this paper, we attempt to understand this fundamental tradeoff between privacy preservation and good recommendation quality. Towards this end, we explicitly capture the mode of user interaction towards shaping the tradeoff above. Specifically we pose and attempt to answer the following questions:

- How can we quantify privacy preservation and recommendation quality?
- What is the resulting degree of privacy preservation of users if each user determines his strategy in terms of revealing his private ratings in a selfish way that takes into account only his personal objective?
- How can we characterize the stable operating points in such a system in terms of rating profile revelation?
- Can users coordinate and jointly determine their rating revelation strategy so as to have mutual benefit?

1.1 Related Work

Recommender systems automate the generation of recommendations based on data analysis techniques [9]. Recommendations for movies on Netflix or books on Amazon are some real-world examples of recommender systems. The approaches that have been proposed in the literature can be classified as follows: (i) Collaborative filtering, (ii) Content-based ones, and (iii) hybrid approaches.

In *collaborative filtering* (CF) systems, a user is recommended items based on past ratings of other users. Specifically, neighborhood-based CF approaches assume that users with correlated interests will most likely like similar items. The Pearson's correlation coefficient is the most widely used measure of similarity between ratings of two users [14]. However there exist several other measures that are used in the literature [17]. Based on a certain similarity measure, these approaches select k users (referred to as a user's neighbors) which have the highest similarity with the user considered for recommendation. Then, a prediction is computed by properly aggregating the ratings of selected neighbors. An extension to neighborhood-based CF is the item-to-item collaborative filtering approach [7], [15]. This approach matches a user's rated items to similar items rather than similar users.

On the other hand, *Content-based* approaches provide recommendations by comparing the content of an item to the content of items of potential interest a user. There are several approaches that treat the content-based recommendation problem as an information retrieval task. Balabanovic *et al.* [1] consider that user preferences can be treated as a query, and unrated objects are scored based on their similarity to this query. An alternative approach is proposed in [11], which treats recommendation as a classification problem. *Hybrid approaches* aim to leverage advantages of both content-based and collaborative filtering ones. Cotter *et al.* [4] propose a simple approach that collects results of both content-based and collaborative filtering approaches, and merges these results to produce the final recommendation. Melville *et al.* [8] propose a framework that uses content-based predictions to convert a sparse user ratings matrix into a full ratings matrix, and subsequently it employs a CF method to provide recommendations.

Since recommender servers need to have access to user preferences in order to predict other items that may be of interest to users, privacy of users is put at risk. A number of different techniques has been proposed to address privacy issues in recommender systems. Polat and Du [13] propose a randomized perturbation technique to protect user privacy in CF approaches. Although randomized perturbation techniques modify the original data to prevent the data collector from learning user profiles, the proposed scheme turns out to provide recommendations with decent accuracy. Another category of works refers to approaches that store user profiles locally and run the recommender system in a distributed fashion. Miller *et al.* [10] propose the PocketLens algorithm for CF in a distributed environment. Their approach requires only the transmission of similarity measures over network, and thus it protects user privacy by keeping their profiles secret. The work [2] addresses the problem of protecting user privacy through substituting the centralized CF system by a virtual peer-to-peer one. Also, user

profiles are partially modified by adding some degree of uncertainty. Although these methods almost eliminate user privacy losses, they require high degree of cooperation among users so as to achieve accurate recommendations.

Lathia *et al.* [6] introduce a new measure to estimate the similarity between two users without breaking user privacy. A randomly generated set of ratings is shared between two users, and then users estimate the number of concordant, discordant and tied pairs of ratings between their own profiles and the randomly generated one. An alternative method for preserving privacy is presented in [3]. Users create communities, and each user seeks recommendations from the most appropriate community. Each community computes a public aggregation of user profiles without violating individual profile privacy, based on distributed singular value decomposition (SVD) of the user rating matrix. A distributed mechanism that focuses on obfuscating user-item connection is proposed in [16]. Each user arbitrarily selects to contact other users over time and modifies his local profile off-line through an aggregation process. Users periodically synchronize their profiles at server (online) with their local ones.

In our work, we develop a game theoretic framework for addressing the privacy preserving challenge in recommender systems. Game theory has recently emerged as a mathematical tool for modeling the interaction of multiple selfish rational agents with conflicting interests, and for predicting stable system points (equilibrium points) from which no agent can obtain additional benefit by unilaterally moving away from them. While game theory has been extensively used in various contexts [12], very few works have used game theory for privacy related issues. The work [5] proposes a formulation of the privacy preserving data mining (PPDM) problem as a multi-party game. It relaxes many of the assumptions made by existing PPDM approaches,thus aiming to develop new robust algorithms for preserving privacy in data mining.

1.2 Our Contribution

In this work we address the fundamental tradeoff between privacy preservation and high-quality recommendation. We assume that multiple users submit their ratings to a third party about the items they have viewed. The third party aggregates these ratings and generates personalized recommendations for each user. The quality of recommendations for each user depends on submitted rating profiles from all users, including the user to which the recommendation is destined. Each user would like to declare a rating profile so as to preserve data privacy as much as possible, while not causing deterioration in the quality of the recommendation he would get, compared to the one he would get if he revealed his true private profile.

The contributions of our work to the literature are as follows: (i) We develop a mathematical framework for quantifying the goal of privacy preservation and that of good quality recommendations, and we define a user's strategy in terms of deciding about his declared rating profile to the third party; (ii) we employ game theory to model and study the interaction of multiple users and we derive conditions and expressions for the Nash Equilibrium Point (NEP). This consists

of the rating strategy of each user, such that no user can benefit in terms of improving its privacy by unilaterally deviating from that strategy. User strategies converge to the NEP after an iterative best-response strategy update; (iii) for a hybrid recommendation system, we find that the NEP strategy for each user in terms of privacy preservation is to declare false rating only for one item, the one that is highly ranked in his private profile and less correlated with items for which he anticipates recommendationq; (iv) We present various modes of cooperation by which users can mutually benefit. To the best of our knowledge, this is the first work that applies the framework of game theory to address the arising user interaction in privacy preserving recommendation systems.

The rest of the paper is organized as follows. In section 2 we present the model and assumptions for our approach. Section 3 elaborates on the case of a hybrid recommendation system. In section 4 we obtain valuable insights for conflict and cooperation in user interaction by analyzing the case of two users. Section 5 includes numerical results and section 6 concludes our study.

2 Model and Problem Definition

2.1 Ratings and Recommendation

Consider a set of $\mathcal{U}$ of N users and a set of items $\mathcal{I}$ available for recommendation. Each user i has already viewed, purchased, or in general it has obtained experience for a small subset of items $\mathcal{S}_i \subset \mathcal{I}$. Usually it is $|\mathcal{S}_i| << |\mathcal{U}|$, where $|\mathcal{A}|$ denotes the cardinality of set $\mathcal{A}$. Denote by $\mathbf{p}_i = (p_{ik} : k \in \mathcal{S}_i)$ the vector of ratings of user i for the items it has viewed, where p_{ik} is the rating of user i for item $k \in \mathcal{S}_i$. Without loss of generality, we assume that p_{ik} takes positive values in a continuous set which is upper bounded, i.e. it is $0 \leq p_{ik} \leq P$. The vector of ratings $\mathbf{p}_i$ is private information for each user i, and we refer to that as the *private profile* or *private ratings vector* of user i. Clearly, the private profile consists of the identities of viewed items and their ratings.

After viewing or experiencing items $k \in \mathcal{S}_i$, user i has to submit a rating to a third party, which will be a recommendation server. Let $\mathbf{q}_i = (q_{ik} : k \in \mathcal{S}_i)$ be the vector of *declared* ratings from user i to the server. This can in general be different from $\mathbf{p}_i$. We refer to $\mathbf{q}_i$ as the *declared profile* or the *declared ratings vector* of user i. The declared profile consists of the identities of viewed items and their declared rating. In this work, we assume that the user will always declare all items it has viewed. Hence, $\mathbf{q}_i$ will include only items $k \in \mathcal{S}_i$ and only these, and the user may only alter the ratings for these items.

The recommendation server is the repository of all ratings submitted by all users. It collects declared user profiles and is responsible for issuing the different, personalized recommendations to different users. Let $\mathbf{P} = (\mathbf{p}_i : i \in \mathcal{U})$ be the ensemble of private ratings of users. Let $\mathbf{Q} = (\mathbf{q}_i : i \in \mathcal{U})$ be the ensemble of declared ratings of all users to the server. When the server receives a recommendation request from a user i, it takes into account the ensemble of ratings $\mathbf{Q}$ to compute a recommendation vector with ratings for items that user i has not yet viewed. Let $\mathbf{r}_i = (r_{i\ell} : \ell \notin \mathcal{S}_i)$ be the *recommendation vector* for user i.

We will assume that the recommendation server employs a generic mapping $f_i(\cdot)$ to compute the recommendation vector for each user i. Hence, we denote the dependence of the recommendation for user i on declared ratings of all users as $\mathbf{r}_i = f_i(\mathbf{Q}) = f_i(\mathbf{q}_1, \ldots, \mathbf{q}_N)$. Here, we have implicitly assumed that ratings of all users in the system are taken into account. However, in general the recommendation server may take into account just a subset of users and their ratings in order to compute the recommendation for a user i. In this work, we are not concerned with designing a recommendation mapping; we will assume that a given mapping is employed by the server, and this mapping is known to all users. Note that $\mathbf{r}_i$ depends on rating vector $\mathbf{q}_i$ that user i has provided about the items he has viewed. A hint about that dependence was provided in the introduction, and it will be revisited in the sequel.

Next, the recommendation vector is fed back to user i in some way that is intrinsic in the specific recommendation system. In general, a part of vector $\mathbf{r}_i$ is returned to user i. For instance, the server may return just one item, the one with the highest rating out of those in set $\{\ell : \ell \notin \mathcal{S}_i\}$, or in general it may return the L highest rated items from the set above. In this work, without loss of generality, we will assume that the entire vector of ratings $\mathbf{r}_i$ is returned to user i, possibly reordered, such that the highest rated components appear first.

2.2 Privacy Metric

For each user i we define a metric that *quantifies the degree at which privacy is preserved* for user i. Intuitively, the degree of privacy preservation depends on the private profile and the declared profile of user i. We denote this dependence by a continuous function $g_i(\mathbf{p}_i, \mathbf{q}_i)$. In general, different users may value their privacy differently, hence functions $g_i(\cdot)$ in general are different for different users. Here, without loss of generality, we assume that all users are characterized by the same privacy preservation function $g(\cdot)$. Thus, the privacy preservation for user i is quantified as $g(\mathbf{p}_i, \mathbf{q}_i)$.

2.3 Recommendation Quality

The users would like to get good quality recommendations for items that have not been viewed yet. The recommendation that each specific user receives depends on declared profiles of other users to the server, but also on the declared profile of this specific user. Even if a user declares his true private profile, the ratings he would get would still depend on the declared profiles of other users. Thus the user may still receive suboptimal ratings, while at the same time compromising its privacy. The problem for each user i is to specify its declared profile so as to maximize the degree of preserved privacy, while at the same time not affecting much the quality of the recommendation. The latter means the user wants calibrate its declared profile so as to receive recommendations *close to the ones he would receive if he would have declared his true private profile*, regardless of the declaration policy of other users.

Let us denote by $\mathbf{q}_{-i} = (\mathbf{q}_1, \ldots, \mathbf{q}_{i-1}, \mathbf{q}_{i+1}, \ldots, \mathbf{q}_N)$ the declared rating vector of all users except user i. Thus, $\mathbf{r}_i = f_i(\mathbf{q}_i, \mathbf{q}_{-i})$. Now, let $\tilde{\mathbf{r}}_i = f_i(\mathbf{p}_i, \mathbf{q}_{-i})$ be the resulting recommendation vector if user i declared its true profile. Then, the goal above is quantified by the following constraint for user i:

$$(\mathbf{r}_i - \tilde{\mathbf{r}}_i)^2 \leq D \;\Leftrightarrow\; [f_i(\mathbf{q}_i, \mathbf{q}_{-i}) - f_i(\mathbf{p}_i, \mathbf{q}_{-i})]^2 \leq D\,, \tag{1}$$

where D is an upper bound that denotes the maximum distortion that can be tolerated in the recommendation by user i. We assume that all users are characterized by the same such maximum tolerable distortion amount D.

2.4 Problem Formulation

Intuitively, the user would like to submit rating profiles that are sufficiently far away from its real private profile so as to preserve as much privacy as possible, by hiding its private profile. On the other hand, he would like to make the declaration above such that the recommendation to him will not be affected too much, and in that sense he would like to maintain the recommendation vector close enough in distance, at most D to the one he would get if he declared the true private profile. The challenge arises because the constraint (1) above includes the strategies $\mathbf{q}_{-i}$ of other users. The objective above can be formulated from the point of view of each user i as follows:

$$\max_{\mathbf{q}_i} g(\mathbf{p}_i, \mathbf{q}_i) \tag{2}$$

subject to:

$$[f_i(\mathbf{q}_i, \mathbf{q}_{-i}) - f_i(\mathbf{p}_i, \mathbf{q}_{-i})]^2 \leq D \tag{3}$$

In other words, user i has to select its declared profile vector out of a set of feasible profile vectors which satisfy (1). Nevertheless, this set of feasible vectors is determined by declared profiles $\mathbf{q}_{-i}$ of other users. Denote by $F(\mathbf{q}_{-i})$ this feasible set of vectors.

Notice that the problem stated above involves only the point of view of user i which behaves in a selfish but rational manner. That is, he cares only about its own maximum privacy conservation, subject to keeping the quality of the recommendation good enough, and he does not take into account the objectives of other users. In his effort to optimally address and resolve this tradeoff, and in particular to ensure that the recommendation vector will be close enough to the one he would get under full privacy compromise, other users' strategies matter. These other users also act in the same rational way since they strive to fulfill their own privacy preservation objectives, while trying to maintain good quality recommendation for themselves.

Definition of NEP: A strategy profile $\mathbf{Q}^* = (\mathbf{q}_1^*, \ldots, \mathbf{q}_N^*)$ is called Nash Equilibrium Point (NEP) for the privacy preservation problem above if for each user $i = 1, \ldots, N$, the following property holds:

$$g(\mathbf{p}_i, \mathbf{q}_i^*) \geq \max_{\mathbf{q}_i \in F(\mathbf{q}_{-i}^*)} g(\mathbf{p}_i, \mathbf{q}_i)\ \forall \mathbf{q}_i \neq \mathbf{q}_i^* \tag{4}$$

The NEP denotes the point that comprises strategies of all users, from which no user will benefit if it deviates from his strategy unilaterally. In our problem, in the NEP $(\mathbf{q}_1^*, \ldots, \mathbf{q}_N^*)$, no user i can further increase its privacy preservation metric $g(\cdot)$ by altering its declared profile to $\mathbf{q}_i \neq \mathbf{q}_i^*$, provided that all other users stay with their NEP declared profiles.

Cooperative user Strategies: Agents may coordinate among themselves in an effort to mutually benefit from a cooperative approach. A global objective $G(\mathbf{P}, \mathbf{Q})$ needs to be defined for the system first. For instance, $G(\mathbf{P}, \mathbf{Q}) = \sum_{i \in \mathcal{U}} g(\mathbf{p}_i, \mathbf{q}_i)$ denotes the total amount of preserved privacy in the system. Or, $G(\mathbf{P}, \mathbf{Q}) = \min_{i \in \mathcal{U}} g(\mathbf{p}_i, \mathbf{q}_i)$, denoting the user with the least-preserved privacy. In a coordinated approach, users act jointly so as to optimize the global objective. A *feasible cooperation regime* for the N users in $\mathcal{U}$ is a joint profile declaration strategy $\mathbf{Q}^0 = (\mathbf{q}_1^0, \ldots, \mathbf{q}_N^0)$ such that:

$$\mathbf{q}_i^0 \in F(\mathbf{q}_{-i}^0), \quad \text{and} \quad g(\mathbf{p}_i, \mathbf{q}_i^0) \geq g(\mathbf{p}_i, \mathbf{q}_i^*), \qquad \forall\, i \in \mathcal{U}, \tag{5}$$

where $\mathbf{Q}^*$ is the NEP. Namely, a cooperation regime is feasible if: (i) belongs to the set of feasible vectors as specified by constraint (1) for all users, (ii) each user has a privacy at least as much as the one he receives at the NEP. This latter requirement renders cooperation meaningful for the user and provides the incentive to the user so as to participate in the coordinated effort.

A first goal of cooperation is to jointly find the set of feasible cooperation regimes, call it $\mathcal{F}_c$. If $\mathcal{F}_c \neq \emptyset$, there exists at least one joint strategy $\mathbf{Q}^0$ such that all users are privacy-wise better off compared to the NEP, and this strategy can be found from solving the set of inequalities in (5). Out of the set of feasible cooperation regimes, a further goal could be to select one that maximizes the global privacy objective $G(\mathbf{P}, \mathbf{Q})$ or one that guarantees certain properties of the privacy preservation vector $(g(\mathbf{p}_1, \mathbf{q}_1), \ldots, g(\mathbf{p}_N, \mathbf{q}_N))$.

3 The Case of a Hybrid Recommendation System

We consider a specific instance of recommendation system as case study to demonstrate our game theoretic model and analysis and derive various important insights.

3.1 Model Specifics

First, we present the specifics of our model in terms of the recommendation metric computed by the server, the specific privacy preservation and recommendation quality metrics.

Recommendation Metric: In this subsection, we discuss the model we adopt for functions $f_i(\cdot)$ that signify the recommendation metrics that are computed for each user i. Each user declares its profile $\mathbf{q}_i$ for items $k \in \mathcal{S}_i$. For each user i,

the recommendation server applies the following measure to compute metrics $r_{i\ell}$ for items $\ell \notin \mathcal{S}_i$, $\ell \in \mathcal{S}_j$ for $j \neq i$, so as to rate them and include them in the recommendation vector that is sent to each user i:

$$r_{i\ell} = \frac{1}{N-1} \sum_{\substack{j \neq i: \\ \ell \in \mathcal{S}_j}} q_{j\ell} \cdot \frac{1}{|\mathcal{S}_i|} \sum_{k \in \mathcal{S}_i} \rho_{k\ell} q_{ik} \,, \tag{6}$$

where $\rho_{k\ell} \in [0,1]$ is the correlation between items k and ℓ. The server computes the metric above for all $\ell \notin \mathcal{S}_i$ and forms vector $\mathbf{r}_i$. We will assume that the $|\mathcal{I}| \times |\mathcal{I}|$ correlation matrix that contains the pairwise correlations between any two items in the system is computed a priori, it is fixed, it is preloaded to the server and is known by user agents. For example, if the items are movies, the correlation between two movies could be directly related to the common theme of the movie, common starring actors, the director or other attributes.

The recommendation metric above pertaining to user i can be viewed as an instance of a *hybrid* recommendation. Indeed, the first term above implies a *collaborative filtering* approach, in which, for each item ℓ under tentative recommendation to user i, the ratings of all other users are aggregated. On the other hand, the second term can be viewed as representative of a content-based recommendation approach, since it involves a correlation metric that connects item ℓ (candidate for recommendation) with other items that user i has viewed.

Here, the aggregation function in the first part is taken to be simply the mean rating of all other users $j \neq i$ which have already viewed the item. Clearly various modes of aggregating the ratings of other users can be employed. For example, different weights may be applied in the aggregation. Or, only the ratings from a subset of users are taken into account, e.g. K users which have viewed common items with user i, where K is a parameter of the recommendation server. These K users are denoted by set $\mathcal{U}_i$. In this case the first term would be equal to:

$$\frac{1}{K} \sum_{\substack{j \in \mathcal{U}_i : |\mathcal{U}_i| = K \\ \mathcal{S}_i \cap \mathcal{S}_j \neq \emptyset}} q_{j\ell}$$

In our model, we adopt (6) as the recommendation metric in order to have analytical tractability and expose our approach. We note that a similar treatment and game theoretic results hold for other types of functions $f_i(\cdot)$.

Privacy Preservation: The function $g(\cdot)$ that quantifies privacy preservation for user i is taken to be equal to:

$$g(\mathbf{p}_i, \mathbf{q}_i) = \sum_{k \in \mathcal{S}_i} p_{ik} (p_{ik} - q_{ik})^2 \tag{7}$$

The metric above reflects the intuitive fact that privacy preservation increases as the Euclidean distance $\sum_{k \in \mathcal{S}_i} (p_{ik} - q_{ik})^2$ between the declared and the private profiles increases. This distance is weighted by the private rating p_{ik} so as to capture the fact that, among items whose private and declared rating have the

same distance, it is preferable from a privacy preservation perspective to change the rating of items that are higher rated in reality. Note also that other types of metrics that include various measures of distance between vectors other than the Euclidean one can be used.

Recommendation Quality: Since users modify their private ratings when they declare them to the server in an effort to increase their privacy, they affect the quality of the recommendation they get from the server. For user i, we measure this effect in terms of the difference between the recommendation user i gets if he declares profile $\mathbf{q}_i$ and the one he would get if he declared the real private rating $\mathbf{p}_i$, regardless of what other users do. Other users $j \neq i$ make in general declarations $\mathbf{q}_j$. Thus, the constraint that needs to be fulfilled for acceptable recommendation quality for user i is derived by using (3) and (6):

$$\frac{1}{|\mathcal{S}_i|} \sum_{k \in \mathcal{S}_i} \sum_{\ell \notin \mathcal{S}_i} \frac{1}{N-1} \sum_{\substack{j \neq i: \\ \ell \in \mathcal{S}_j}} q_{j\ell} \cdot \rho_{k\ell} (q_{ik} - p_{ik})^2 \leq D \tag{8}$$

Information Exchange between users and the Server: An iterative process of data exchange between users and the recommendation server takes place. We envision a software agent at the side of each user i which acts on behalf of the user. The agent is responsible for preserving privacy of each user i and to deliver good recommendation quality results to each user i. The agent continuously sends queries for recommendation to the server. The steps of data exchange are summarized as follows:

- **STEP 0:** An initial or default rating vector $\mathbf{q}_i^{(init)}$ is used by each user.
- For each user $i = 1, \ldots, N$:
- **STEP 1:** At each iteration cycle t, the server passes to each agent i the ratings from other users that refer to items that user i has not viewed yet. These ratings are based on the information that agents of other users have sent to the server at the same iteration cycle. That is, the server passes to user i the quantity $\frac{1}{N-1} \sum_{j \neq i} q_{j\ell}^{(t)}$ which is an aggregated version of ratings $\{q_{j\ell}^{(t)}\}$ for each item $\ell \notin \mathcal{S}_i$ and users $j \neq i$.
- **STEP 2:** The agent of each user i gets to observe the ratings of other users, namely it observes the first part of (8). It then solves the optimization problem (**P**):

$$\max_{\mathbf{q}_i^{(t)}} g(\mathbf{p}_i, \mathbf{q}_i^{(t)}) = \sum_{k \in \mathcal{S}_i} p_{ik} (p_{ik} - q_{ik}^{(t)})^2, \tag{9}$$

  subject to constraint (8), which includes $\{\mathbf{q}_j^{(t-1)}\}_{j \neq i}$ from the previous iteration, and thus it computes his own declared rating vector $\mathbf{q}_i^{(t)}$ for the current iteration t.
- **STEP 3:** Each agent declares its rating $\mathbf{q}_i^{(t)}$ to the server.

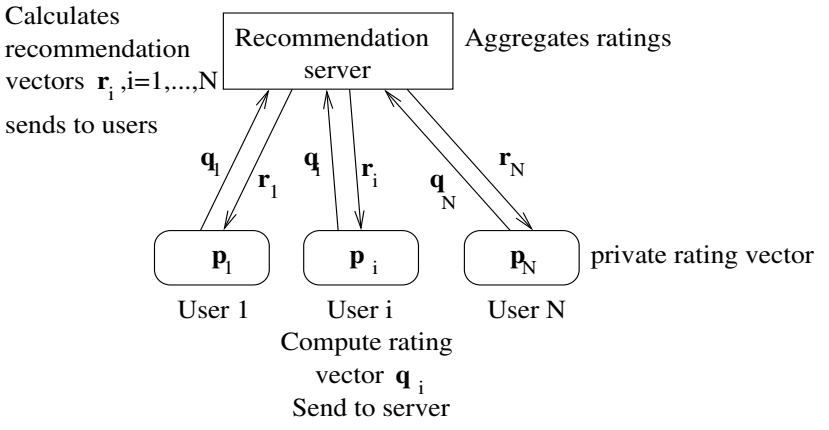

Fig. 1. Overview of system architecture and of the data exchange process at each iteration

– **STEP 4:** The server uses these ratings to compose aggregated quantities for items that have not been viewed by other users. Go to Step 1. Repeat until convergence.

The system is depicted in Figure 1. Agents in general update their rating vectors in subsequent iterations. At each iteration cycle t, each user i solves its own optimization problem based on the ratings of other users $\{\mathbf{q}_j^{(t-1)}\}_{j\neq i}$, that have been passed to i at the end of the previous iteration cycle. User i declares its ratings to the server. The server collects rating vectors from all users and it announces the relevant parts to different users in order for the new iteration to start. The procedure continues as above at each iteration.

Convergence to the NEP: The procedure described above involves a *Linear Programming* problem that is solved by each user. The submitted ratings of other users appear in the constraint (8). The iterative procedure above is an instance of iterative *best-response update* from each user. It is known from fundamental game theory that the sequence of best-response updates for linear problems converges to the NEP, starting from any initial vector $\mathbf{q}_i^{(init)}$.

4 Game Theoretic Analysis

By setting $x_{ik} = (p_{ik} - q_{ik})^2$, and $\mathbf{x}_i = (x_{ik} : k \in \mathcal{S}_i)$, it can be seen that problem $(\mathbf{P})$ that is solved by each user i at each iteration is written as:

$$\max_{\mathbf{x}_i} \sum_{k\in\mathcal{S}_i} p_{ik}x_{ik}, \qquad \text{subject to:} \qquad \sum_{k\in\mathcal{S}_i} \beta_{ik}x_{ik} \leq D(N-1)\,, \qquad (10)$$

with

$$\beta_{ik} = \frac{1}{|\mathcal{S}_i|} \sum_{\ell \notin \mathcal{S}_i} \sum_{j \neq i : \ell \in \mathcal{S}_j} q_{j\ell} \rho_{k\ell} \tag{11}$$

and it is a *Linear Programming* problem. The solution to this problem is found among the extreme points of the feasible set. Each user finds item,

$$k^* = \arg\min_{k \in \mathcal{S}_i} \frac{\beta_{ik}}{p_{ik}} \tag{12}$$

and it sets

$$x_{ik^*} = D(N-1)|\mathcal{S}_i| \frac{p_{ik^*}}{\beta_{ik^*}} \tag{13}$$

For all other items $k \neq k^*$, it is $x_{ik} = 0$. Moving back to the initial variables, we deduce that for item k^*, the rating declaration should be:

$$q_{ik^*} = p_{ik^*} \pm \sqrt{\frac{D(N-1)|\mathcal{S}_i| p_{ik^*}}{\beta_{ik^*}}} \tag{14}$$

while $q_{ik} = p_{ik}$ for other items $k \neq k^*$. It can be observed that agent i maximizes its preserved privacy if it declares its true private profile for all viewed items, except one, k^*, for which the quantity

$$\frac{\beta_{ik}}{p_{ik}} = \frac{\sum_{\ell \notin \mathcal{S}_i} \rho_{k\ell} \sum_{j \neq i : \ell \in \mathcal{S}_j} q_{j\ell}}{p_{ik}} \tag{15}$$

is the smallest among items it has viewed. The denominator implies that an item has more chances to be the selected one k^*, if it is highly rated in its private rating vector. The numerator implies that this item should have low correlation with items that it has not viewed for which the average declared rating of other users for this item is low. This is clearly meaningful for privacy preservation. We note that the above result about privacy maximization emerges because we quantify privacy and recommendation quality with Euclidean distance metric. A metric based on vector norms other than Euclidean would alter the nature of the solution.

4.1 Special Case: N=2 Users

In order to demonstrate properties of the equilibrium, consider the simplest nontrivial case of $N = 2$ users, each of which has viewed two items. User 1 has viewed items in set $\mathcal{I} = \{A, B\}$ and has private rating vector (p_{1A}, p_{1B}), while user 2 has viewed items in $\mathcal{I}_2 = \{B, C\}$ with private rating vector (p_{2B}, p_{2C}). Let $x_{ik} = (p_{ik} - q_{ik})^2$ for $i = 1, 2$ and $k = A, B$.

Game Theoretic Interaction: If users 1 and 2 act autonomously and without coordination, each user will attempt to maximize its own privacy. Thus, the problem faced by user A is:

$$\max_{x_{1A}, x_{1B}} p_{1A} x_{1A} + p_{1B} x_{1B}, \quad \text{subject to:} \quad \rho_{AC} x_{1A} + \rho_{BC} x_{1B} \leq \frac{2D}{q_{2C}}, \tag{16}$$

where the factor of 2 comes due to the $1/|\mathcal{S}_1|$ factor in the left-hand side of inequality. Similarly, for user 2, the problem is:

$$\max_{x_{2B},x_{2C}} p_{2B}x_{2B} + p_{2C}x_{2C}, \quad \text{subject to:} \quad \rho_{AB}x_{2B} + \rho_{AC}x_{2C} \leq \frac{D}{q_{1A}}\,. \tag{17}$$

The NEP is the point $(\mathbf{x}_1^*, \mathbf{x}_2^*) = (x_{1A}^*, x_{1B}^*, x_{2B}^*, x_{2C}^*)$ that solves the two problems above. Depending on the private rating vectors $\mathbf{p}_1 = (p_{1A}, p_{1B})$, $\mathbf{p}_2 = (p_{2B}, p_{2C})$ and correlations $\rho_{AC}, \rho_{BC}, \rho_{AB}$, we distinguish four cases:

$$\frac{\rho_{AC}}{p_{1A}} \lesseqgtr \frac{\rho_{BC}}{p_{1B}}, \quad \text{and} \quad \frac{\rho_{AB}}{p_{2B}} \lesseqgtr \frac{\rho_{AC}}{p_{2C}} \tag{18}$$

Observe that user 1 will declare the true private profile for the item for which the fraction above is larger, and it declare a different rating for the item for which the fraction is the smaller. Thus, he prefers to declare different rating for the item that is higher ranked and least correlated to item C that is candidate for recommendation to him. For instance, if $\frac{\rho_{AC}}{p_{1A}} < \frac{\rho_{BC}}{p_{1B}}$, user 1 will maximize its privacy by setting $x_{1B} = 0$, thus declaring $q_{1B} = p_{1B}$, while $x_{1A} = \frac{2D}{\rho_{AC}q_{2C}}$. For each of the four cases above, we have a respective NEP. For example, if

$$\frac{\rho_{AC}}{p_{1A}} < \frac{\rho_{BC}}{p_{1B}} \quad \text{and} \quad \frac{\rho_{AB}}{p_{2B}} < \frac{\rho_{AC}}{p_{2C}} \tag{19}$$

then the NEP is:

$$\mathbf{x}_1 = (\frac{2D}{\rho_{AC}p_{2C}}, 0), \quad \text{and} \quad \mathbf{x}_2 = (\frac{2D}{\rho_{AB}(p_{1A} \pm \sqrt{\frac{2D}{\rho_{AC}p_{2C}}})}, 0) \tag{20}$$

or, equivalently

$$\mathbf{q}_1^* = (p_{1A} \pm \sqrt{\frac{2D}{\rho_{AC}p_{2C}}}, p_{1B})\ ,\ \mathbf{q}_2^* = (p_{2B} \pm \sqrt{\frac{2D}{\rho_{AB}(p_{1A} \pm \sqrt{\frac{2D}{\rho_{AC}p_{2C}}})}}, p_{2C})\,, \tag{21}$$

and the privacy metrics are $P_1 = p_{1A}x_{1A}$, $P_2 = p_{2B}x_{2B}$.

5 Numerical Results

In this section, we evaluate the performance of our game theoretic approach in terms of privacy preservation and recommendation quality. To assess the privacy preservation in our system, we have adopted the privacy metric presented in (7). We consider a recommender system consisting of $N = 50$ users and $|\mathcal{I}| = 20$ items. The content correlation of items is calculated a priori and announced to the agents that represent the users. The preferences of users for items are randomly selected for the sake of performance evaluation, and we assume that user

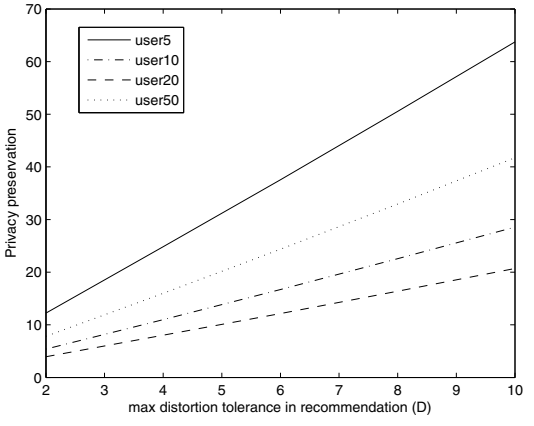

Fig. 2. Privacy preservation versus maximum distortion tolerance in recommendation

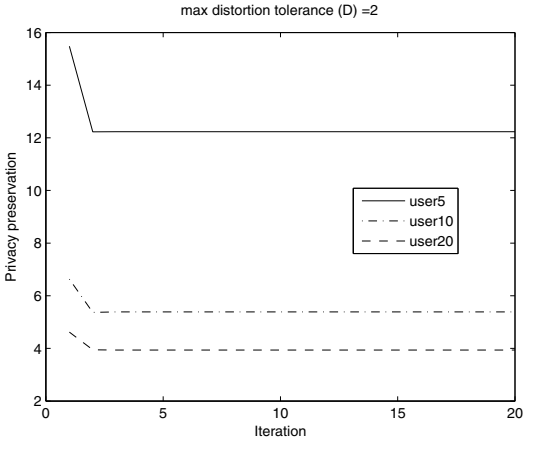

Fig. 3. Convergence of the iterative best response strategy for different users

ratings lie in the value interval $[1, 5]$. In Figure 2, we depict the privacy preservation metric for different users as a function of the maximum distortion tolerance D in recommendation. We observe that as user tolerance to recommendation quality increases, the privacy preservation metric also increases. This confirms the tradeoff between privacy preservation and quality of recommendation. Thus, users that are less tolerant to the error of recommendation quality they receive from the server, have to reveal more information about their preferences.

Our approach is based on the iterative best response process of data exchange between users and the recommendation server that was discussed in section 3.

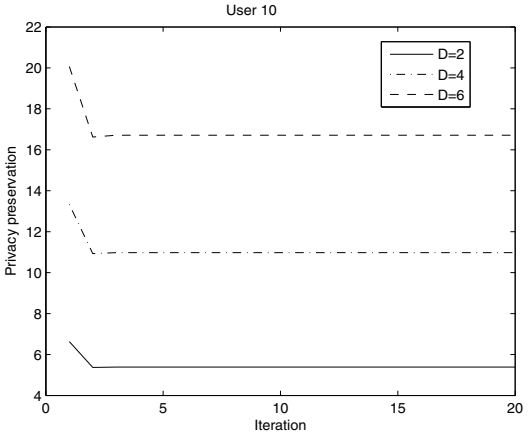

Fig. 4. Convergence of the iterative best response strategy for different values of D

Figure 3 depicts and verifies the convergence of the privacy preservation iteration for different users at the NEP as the rating vector exchange process progresses. Then we consider that a specific user chooses to vary the values of his/her maximum distortion tolerance in recommendation (D). Figure 4 also shows that the privacy preservation iteration of a user convergences at the NEP and this can also be verified for different values of D. It is clear that after a small number of iterations, usually no more than $2-3$, the system converges and the privacy preservation metric of a user at the NEP is determined. This fast convergence of the best response update is a direct consequence of the linear programming type of problem that each user solves.

6 Conclusion

In this work, we took a first step towards characterizing the fundamental tradeoff between privacy preservation and good quality recommendation. We introduced a game theoretic framework for capturing the interaction and conflicting interests of users in the context of privacy preservation in recommendation systems. Viewed abstractly from the perspective of each user, the privacy preservation problem that arises in the process of deciding about the declared profile reduces to that of placing the declared rating vector sufficiently far away from the actual, private vector. The constraint on having recommendation quality close enough to the one that would be achieved if the true profile was revealed, places a constraint on the meaningful distance between the actual and declared profiles. Nevertheless, the key challenge is that the extent to which this constraint is satisfied, depends on the declared profiles of other users as well, which in turn face a similar profile vector placement problem. We attempted to capture this interaction, we characterized the Nash Equilibrium Points, and we proposed various modes of cooperation of users.

References

1. Balabanovic, M., Shoham, Y.: Fab: Content-based collaborative recommendation. Communications of the Association for Computing Machinery 40(3) (1997)
2. Berkovsky, S., Eytani, Y., Kuflik, T., Ricci, F.: Enhancing privacy and preserving accuracy of a distributed collaborative filtering. In: Proc. of ACM RecSys (2007)
3. Canny, J.: Collaborative filtering with privacy. In: IEEE Symposium on Security and Privacy (2002)
4. Cotter, P., Smyth, B.: PTV: Intelligent personalized tv guides. In: Proc. of AAAI/IAAI (2002)
5. Kargupta, H., Das, K., Liu, K.: A game theoretic approach toward multi-party privacy-preserving distributed data mining. In: Proc. of PKDD (2007)
6. Lathia, N., Hailes, S., Capra, L.: Private distributed collaborative filtering using estimated concordance measures. In: Proc. of ACM RecSys (2007)
7. Linden, G., Smith, B., York, J.: Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing 7(1) (2003)
8. Mellville, P., Mooney, R.J., Nagarajan, R.: Content-boosted collaborative filtering for improved recommedations. In: Proc. of the National Conference on Artificial Intelligence (2002)
9. Melville, P., Sindhwani, V.: Recommender Systems, Encyclopedia of Machine Learning. Springer, Heidelberg (2010)
10. Miller, B., Konstan, J.A., Riedl, J.: Pocketlens: Toward a personal recommender system. ACM Transactions on Information Systems 22(3) (2004)
11. Mooney, R.J., Roy, L.: Content-based book recommending using learning for text categorization. In: Proc. of ACM Conf. on Digital Libraries (2000)
12. Nisan, N., Roughgarden, T., Tardos, E., Vazirani, V.V.: Algorithmic Game Theory, Cambridge (2007)
13. Polat, H., Du, W.: Privacy-preserving collaborative filtering using randomized perturbation techniques. In: Proc. of Inter. Conf. on Data Mining, ICDM (2003)
14. Resnick, P., Iacovou, N., Sushak, M., Bergstrom, M., Reidl, J.: Grouplens: An open architecture for collaborative filtering of netnews. In: Proc. of the Computer Supported Cooperative Work Conference (1994)
15. Sarwar, B., Karypis, G., Konstan, J., Reidl, J.: Item-based collaborative filtering recommendation algorithms. In: Proc. of the Inter. Conf. of WWW (2001)
16. Shokri, R., Pedarsani, P., Theodorakopoulos, G., Hubaux, J.P.: Preserving privacy in collaborative filtering through distributed aggregation of offline profiles. In: Proc. of ACM RecSys (2009)
17. Su, X., Khoshgoftaar, T.M.: A survey of collaborative filtering techniques. Advances in Artificial Intelligence (2009)