Journal of Information Processing Vol.21 No.4 617–623 (Oct. 2013)

[DOI: 10.2197/ipsjjip.21.617]

Regular Paper

Privacy-preserving Collaborative Filtering Using Randomized Response

Hiroaki Kikuchi[1,2,a)] Anna Mochizuki[3,b)]

Received: December 5, 2012, Accepted: June 14, 2013

Abstract: This paper proposes a new privacy-preserving recommendation method classified into a randomized perturbation scheme in which a user adds a random noise to the original rating value and a server provides a disguised data to allow users to predict the rating value for unseen items. The proposed scheme performs a perturbation in a *randomized response* scheme, which preserves a higher degree of privacy than that of an additive perturbation. To address the accuracy reduction of the randomized response, the proposed scheme uses a *posterior probability distribution function*, derived from Bayes' estimation for the reconstruction of the original distribution, to revise the similarity between items computed from the disguised matrix. A simple experiment shows the accuracy improvement of the proposed scheme.

Keywords: collaborative filtering, privacy-preserving data mining, cryptographic protocol

1. Introduction

Collaborative Filtering (CF) [10], [11], [12] is a useful method to predict rating values for unseen items based on the preference of communities who have evaluated the target items and have a similar preference with the users who wish to get a recommendation. CF has potential applications in the age of Internet where a huge number of items are available on-line and hence users worry about the best choice out of them. These personalized recommendations, however, raise the privacy concerns, too [6]. The personal preference database has a risk to be disclosed by malicious insiders.

In order to address the privacy concern in collaborative filtering, there have been several attempts so far. The first one, called the *cryptographic approach*, is made by Canny [7], [8], using an additive homomorphic cryptosystem for performing the Singular Value Decomposition (SVD) of a matrix between items and users. The scheme requires a number of iterative vector additions, which suffers from intensive computational costs to perform. In Ref. [14], Ahmad and Khokhar studied the modified version of Canny's protocol using the modified ElGamal cryptosystem instead of the Paillier [9]. These approaches suffer from high computation and communication costs for cryptographic operations.

The second approach of privacy-preserving CF is a randomized algorithm, often referred to as a *random perturbation*, introduced by Agrawal and Srikant in Ref. [1]. The idea is to add a random noise to an original data to preserve the privacy of the original rating against the server, called *disguised data*, and then estimate the original distribution from many collected disguised vectors. Polat and Du proposed CF schemes for the Pearson correlation-based algorithm [2]. Their idea was based on a hypothesis that the random noise uniformly chosen from the range $[\alpha, +\alpha]$ decreases to zero, namely, letting $A = (a_1, \dots, a_n)$, $B = (b_i, \dots, b_n)$, and $R = (r_1, \dots, r_n)$ be the original vectors and the random vector, the scalar product is estimated as

$$(A + R) \cdot (B + R) = \sum (a_i b_i + a_i r_i + b_i r_i + r_i^2) \approx \sum a_i b_i,$$

where $\sum a_i r_i$ and $\sum r_i^2$ will converge to zero as aggregated from many users. However, Zhang et al. pointed out that an additive perturbation does not preserve the privacy as much as had been believed by showing the experiment to derive an amount of the original data in Refs. [4], [5]. Huang, Du and Chen also applied the Principal Component Analysis (PCA) to the disguised data to retrieve partial original data [3].

In this paper, we suggest that the perturbation called *randomized response* preserves the privacy better than the additive perturbation. In the randomized response, we replace the original value by a uniformly chosen value in a pre-determined probability, which yields uniformly distributed disguised data, while the additive perturbation preserves the original distribution that contains significant information to recover the original information. Hence, the randomized response is expected to be robust against attacks such as PCA.

In the cost of privacy degree, the randomized response based perturbation may suffer from the reduction of the accuracy of the prediction. We claim that Polat's hypothesis does not hold in a randomized response since the average of the random noise does not necessary amounts to zero any more. To improve the accuracy in CF, we suggest to use *the posterior probability distribution function* $P(X|Y)$, which is computed in Bayes' estimation to

[1] Department of Frontier Media Science, School of Interdisciplinary Mathematical Sciences, Meiji University, Nakano, Tokyo 164–8525, Japan

[2] School of Information and Telecommunication Engineering, Tokai University, Minato, Tokyo 108–8619, Japan

[3] Graduate School of Science and Technology, Tokai University, Hiratsuka, Kanagawa, 259–1292 Japan

[a)] kikn@meiji.ac.jp

[b)] cream_18_puff@cs.dm.u-tokai.ac.jp

reconstruct the original distribution X from the disguised Y. In this paper, we propose a reconstruction scheme based on Bayes' rule and the scheme by Agrawal and Srikant so that we apply to the perturbation based on the randomized response. We show that the similarity between items computed from the disguised data Y can be revised based on the posterior probability distribution $P(X|Y)$. Our simple experiment demonstrates the accuracy improvement of our proposed CF method.

Our contributions of this work are as follows.

(1) We propose a new perturbation method based on a randomized response to preserve the privacy of a rating value.

(2) We present a new reconstruction method for a randomized response based perturbed data.

(3) We present a new CF method to predict a rating value based on the posterior probability distribution given from the revised similarity between items.

(4) We present experimental results showing the accuracy improvement.

The rest of the paper is organized as follows. In Section 2, we show the fundamental model and definitions of primitives used to construct our protocol. We show our protocol in Section 3 with simple numerical examples. The performance evaluations are given in Section 4 and finally, we conclude the paper.

2. Preliminaries

2.1 Model

Let $U = \{u_1, u_2, \ldots, u_n\}$ be a set of users, where n is the number of users. Let $I = \{i_1, i_2, \ldots, i_m\}$ be a set of items, where m is the number of items. Let $r_{i,j}$ be a rating given by user u_i for item j, for $i = 1, \ldots, n$, and $j = 1, \ldots, m$. Users do not evaluate all items. We denote a missing rating by $r_{i,j} = \phi$. We assume that the matrix of ratings contains many missing elements, that is, a sparse matrix.

The goal of CF is to predict a missing rating based on the other users' preference to the given item. Our model supposes that users are willing to get recommendations for items that they have not seen before, but at the same time they are concerned about the privacy of ratings made by themselves.

2.2 User-based Collaborative Filtering Algorithm

Collaborative Filtering is an algorithm to estimate missing ratings based on the preference database. The prediction of user i for item k is given by a weighted average of neighbor users whose ratings are expected to be similar to those of the target user:

$$P^U_{i,k} = \overline{r_{i,\cdot}} + \frac{\sum_{j \in U_k} s(u_i, u_j)(r_{j,k} - \overline{r_{j,\cdot}})}{\sum_{j \in U_k} |s(u_i, u_j)|} \tag{1}$$

where U_k is the set of users who have rated the k-th item, i.e., $U_k = \{i \in U \mid r_{ik} \neq \phi\}$, and $\overline{r_{j,\cdot}}$ is the mean of all ratings made by user j. The weight $s(u_i, u_j)$ is the similarity between users u_i and u_j. Although there exist many definitions of the similarity including the Pearson correlation coefficient, or the Euclidean distance, in this work, we use the simplest Cosine correlation defined by

$$s(u_i, u_j) = \frac{\sum_{k=1}^{n} r_{i,k} r_{j,k}}{\sqrt{r_{i,1}^2 + \cdots + r_{i,n}^2}\sqrt{r_{j,1}^2 + \cdots + r_{j,n}^2}}. \tag{2}$$

Table 1 Original probability distribution of A, $P(A)$.

a	0	1	2	3
$P(A = a)$	0.1	0.3	0.1	0.5

Table 2 Conditional probability distribution $P(B|A)$, where the probability to remain same as the original, $p = 0.4$.

$B \backslash A$	0	1	2	3
0	0.4	0.2	0.2	0.2
1	0.2	0.4	0.2	0.2
2	0.2	0.2	0.4	0.2
3	0.2	0.2	0.2	0.4

Table 3 Probability distribution of the disguised vector $P(B)$.

b	0	1	2	3
$P(B = b)$	0.22	0.26	0.22	0.3

Table 4 The first and second estimations for the posterior distribution of $A = a$ in the reconstruction algorithm.

a	0	1	2	3
$P^1(A = a)$	0.22	0.26	0.22	0.31
$P^2(A = a)$	0.21	0.26	0.21	0.33

2.2.1 Randomized Response

We demonstrate how a randomized response works with a simple instance. Let A be a random variable chosen from a probability distribution defined in **Table 1**.

A *randomized response* B of A with the conditional probability $P(B|A)$ given by **Table 2**, in which an original value is replaced by a randomly chosen data from the range, $\{0, 1, 2, 3\}$, with a probability of $1 - p$, i.e., p is the probability to be left as it was. **Table 3** shows the result of the probability distribution of B, modified from A in the randomized response with $p = 0.4$. The randomized value is almost uniformly distributed around 0.3, while the original value was skewed at $A = 3$ as the highest probability.

2.2.2 Reconstructing the Original Distribution

Given the probability distribution of a randomized response, $P(B)$, and the conditional probability to be used to disguise the original distribution, $P(B|A)$, we would like to estimate the original distribution of $P(A)$. The solution to the problem was given by Agrawal and Srikant in the algorithm known as "reconstruction algorithm" [1].

We begin with an initial estimation as $P^0(A) = P(B)$, and use Bayes' rule to estimate the i-th posterior distribution function as

$$P^i(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{3}$$

$$= \frac{P(B|A)P^{i-1}(A)}{\sum_{a \in A} P(B|A = a)P^{i-1}(A = a)} \tag{4}$$

which follows the corresponding posterior probability of A given by the average for possible values in the range as

$$P^1(A) = \sum_{b \in B} P^0(A|B = b)P(B = b).$$

After a sufficient number of iterations of estimation by $P^{i+1}(A) = P^i(A)$, we have the convergence of the estimation, indicated as P^*. The estimated distribution would be close to the original one as illustrated in **Table 4**.

2.3 Item-based Collaborative Filtering

Collaborative filtering (CF) is a method to predict the ratings of

Table 5 Matrix of original rating values (R^X).

	i_1	i_2	i_3	i_4	i_5
u_1	2	2	3	1	
u_2	1	3	2		3
u_3	2		3	3	2
u_4	3	2	$*$	2	2

an unseen item based on the similarities between users, or items, known as user-based and item-based CFs, respectively. In the paper, we study the latter one as for the base algorithm to be extended toward privacy preservation.

Let V be the range of the rating value, and $r_{u,i} \in V$ be a rating value for item i evaluated by user u. Let R be an n by m matrix of rating values, where n and m are the numbers of users, and items, respectively. **Table 5** is an example matrix with $n = 4$ and $m = 5$. Note that users don't evaluate all items and empty cells indicate *missing values*.

According to the item-based CF algorithm, a rating value indicated as $*$ is predicted by [*1]

$$r_{u,i} = \frac{\sum_j^m s_{j,i} r_{u,j}}{\sum_j \mid s_{j,i} \mid}, \qquad (5)$$

where $s_{i,j}$ is a similarity between items i and j. An arbitrary definition of the similarity can be used here. Because of the simplicity, we use the cosine similarity defined by

$$s_{j,i} = \frac{\sum_{k=1}^n r_{k,i} r_{k,j}}{\sqrt{r_{1,i}^2 + \cdots + r_{n,i}^2}\sqrt{r_{1,j}^2 + \cdots + r_{n,j}^2}}. \qquad (6)$$

3. Proposed Method

3.1 Idea

The goal of privacy-preserving collaborative filtering is to predict rating values as accurately as possible, given the probability p to randomize the original matrix R^X and the disguised matrix R^Y.

The privacy of a rating value is preserved since the original values are disguised by the randomized values. The degree of privacy in the randomized response is expected to be higher than that of the additive perturbation in which the original value X is randomized by a random noise R as $Y = X + R$ because the distribution of Y is skewed as evenly in the randomized response, while the random variable Y in the additive perturbation is distributed almost identically to the original X.

Without learning the original rating matrix R^X, we may estimate the rating value from the disguised matrix R^X but we have to compromise the accuracy of prediction in the cost of privacy. The accuracy is not as high as that of the prediction from the original matrix R^X.

Our approach for improving the accuracy is to use the posterior probability distribution function $P(X|Y)$ given via the reconstruction processes. We use the additional information not only for reconstructing the original distribution but also for predicting the rating for missing items. Namely, our formalized problem is to predict rating values from $P(X|Y)$ in addition to the disguised matrix R^Y and the probability p of retention a value in the randomization. The accuracy is expected to be improved since we have the estimation of the probability distribution closer to the original one, which must be useful to approximate the similarity between items more accurately.

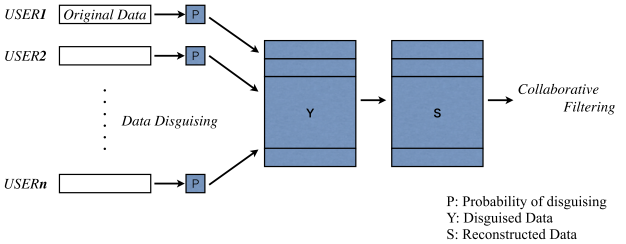

Fig. 1 The proposed scheme for privacy-preserving collaborative filtering.

Table 6 Matrix of disguised rating values R^Y.

	i_1	i_2	i_3	i_4	i_5
u_1	2	3	1	1	
u_2	1	1	2		1
u_3	1		3	3	
u_4	3	2	$*$	2	3

Table 7 Posterior probability distribution $P(X|Y)$ with $p = 0.4$.

$Y \backslash X$	0	1	2	3
0	0.37	0.18	0.23	0.22
1	0.19	0.36	0.23	0.22
2	0.18	0.17	0.44	0.21
3	0.18	0.17	0.22	0.43

3.2 New Scheme – Collaborative Filtering with Expected Similarities

We illustrate our proposed scheme in **Fig. 1**. Each user randomizes his/her original rating vectors X according to the common probability p before submitting to a server. The randomization processes are performed independently from users. The randomized response y of the original rating value $r_{u,i}$ is defined by

$$y = \begin{cases} r_{u,i} & \text{w./p.} = p, \\ v \in V - \{r_{u,j}\} & \text{otherwise}, \end{cases} \qquad (7)$$

where p is a probability that the y is equal to the original value.

The server collects the disguised vectors Y from all users to form the matrix R^Y and then applies the Bayes' reconstruction to obtain the posterior probability distribution of Z. For instance, **Table 7** shows the posterior probability distribution of $P^*(X|Y)$, estimated from the disguised matrix of R^Y in **Table 6** and $P(X|Y)$ in Table 5, which is given from the probability that the randomized value is the same to the original value as $p = 0.4$.

Finally, any party (users and server) allows to predict the rating value for an arbitrary item from the published disguised matrix R^Y and the posterior probability distribution of Z, which is close to the original one of X, in Algorithm 1.

4. Evaluation

4.1 Accuracy

In order to evaluate the improvement in accuracy in the proposed scheme, we compare it with a prediction using the similar-

[*1] We use a simplified equation for prediction without z-score assuming rating values are distributed with mean of zero and variance of 1 for all items. The simplification makes our modification easier and can be extended if we wish to improve accuracy.

```
Algorithm 1 PPCF ($R^Y$,$P(Y|X)$)
Input: disguised matrix $R^Y$, conditional probability of $P(Y|X)$
Out: predicted rating value $r^E_{u,i}$
Step 1 Let $W = Y_1 \cdot Y_2$ be a random variable of product of two disguised
    rating values $Y_1$ and $Y_2$. Compute the probability distribution of $W$ by
        $P(W|Y_1, Y_2) = \sum_{W=\alpha\cdot\beta} P(X = \alpha|Y_1) \cdot P(X = \beta|Y_2).$
Step 2 Compute the expected value of $W$ as
        $E[W|Y_1, Y_2] = \sum_{\gamma\in V_2} P(W = \gamma|Y_1, Y_2),$
    where $V_2$ is a range of product of two $V$s, i.e., letting $V = \{1, \dots, v\}$, we
    have $V_2 = \{1, \dots, v^2\}$.
Step 3 Compute the expected value of Cosine similarity between item $i$ and
    $j$, $s^E_{i,j}$, defined by Eq. (6), by
        $s^E_{i,j} = E[S_{i,j}|R^Y] = \frac{E[\sum_u^n r^X_{u,i} \cdot r^X_{u,j}|R^Y]}{E\left[\sqrt{\sum_u (r^X_{u,i})^2}\sqrt{\sum_u (r^X_{u,j})^2}\right]}$
        $= \frac{\sum_u^n E[W|Y_1 = r^Y_{u,i}, Y_2 = r^Y_{u,j}]}{\sqrt{\sum_u (r^Y_{u,i})^2}\sqrt{\sum_u (r^Y_{u,j})^2}},$
    where we assume that the perturbation does not change the mean of the
    squared sum of norm of the original matrix $R^X$, which is replaced by
    that of the disguised matrix $R^X$ at the denominator.
Step 4 Predict the rating of user $u$ for item $i$ using the revised similarity $s^E_{ii,j}$
    and the collaborative filtering prediction, Eq. (5),
        $r^E_{u,i} = \frac{\sum_j^m S^E_{i,j} \cdot r^X_{u,j}}{\sum_j^m S^E_{i,j}}.$
```

ity computed from a disguised matrix R^Y without correcting.

Our experiment uses the original matrix R^X and the disguised matrix R^Y using probability $p = 0.4$, shown in Tables 5 and 6, respectively. The matrix consists of rating values in the range $V = \{1, 2, 3\}$ for $n = 4$ users, and $m = 5$ items. The value $r_{i,u} = 0$ indicates a missing value.

Figure 2 shows the result of the reconstruction scheme with distributions of rating values in the original matrix X, the disguised matrix Y with $p = 0.4$, and the reconstructed one Z, which is estimated by Bayes rule in 50 iterations. We observe that the reconstruction successfully makes the distribution of Z close to the original X, while the disguised values in Y are distributed nearly uniformly (with about 0.25 probability) and hence the original value is hard to be guessed from a single submitted rating value. The number of iterations to converge depends on the original distribution and the probability to randomize.

We revise the disguised similarity with the posterior probability distribution $P(X|Y)$ obtained through the reconstruction processes. For instance, given two disguised values $Y_1 = 2$ and $Y_2 = 3$, we can compute the probability distribution of the product $W(= Y_1 \cdot Y_2)$ for all possible combinations of X_1 and X_2, as follows

$$P(W|Y_1, Y_2) = \sum_{W=\alpha_1\cdot\alpha_2} P(\alpha_1 \cdot \alpha_2|Y_1, Y_2)$$
$$= \sum_{W=\alpha_1\cdot\alpha_2} P(X_1 = \alpha|Y_1)P(X_2 = \alpha_2|Y_2),$$

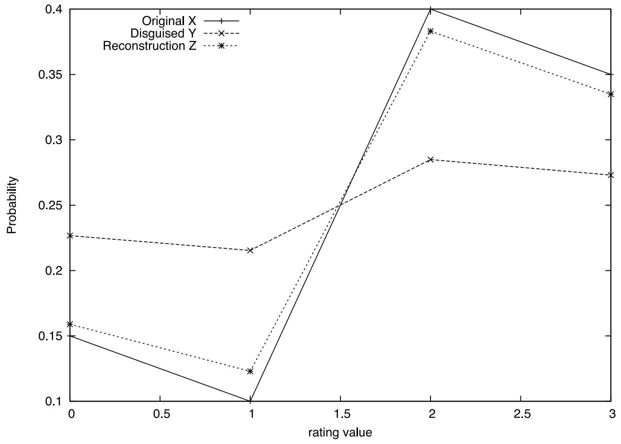

Fig. 2 Distribution of original, disguised and reconstructed values.

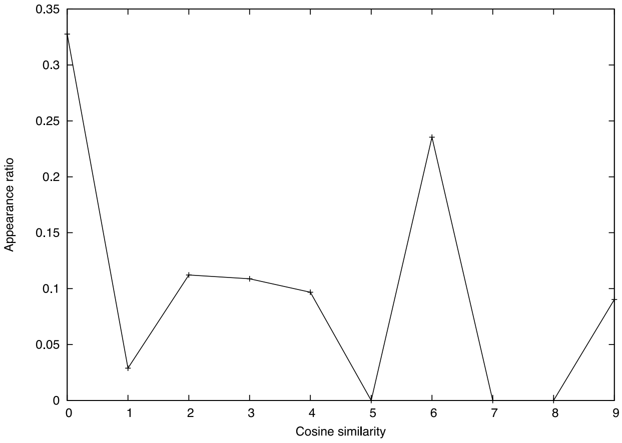

Fig. 3 Probability distribution of $W(= Y_1 \cdot Y_2)$ given $Y_1 = 2$ and $Y_2 = 3$, $P(W|Y_1 = 2, Y_2 = 3)$.

Table 8 Expected value of product of rating values $E[W|Y_1, Y_2]$.

$Y_2\backslash Y_1$	0	1	2	3	sum
0	1.69	1.92	2.18	2.47	8.26
1	1.92	2.19	2.49	2.81	9.41
2	2.18	2.49	2.82	3.19	10.68
3	2.47	2.81	3.19	3.16	11.63

Table 9 Expected value of similarity between items $E[S_{i,j}|R^Y]$.

	i_1	i_2	i_3	i_4	i_5
i_1	–	0.60	0.46	0.66	0.54
i_2	0.60	–	0.46	0.65	0.56
i_3	0.46	0.46	–	0.50	0.44
i_4	0.66	0.65	0.50	–	0.62
i_5	0.54	0.56	0.44	0.62	–

which follows the distribution shown in **Fig. 3**. The most likely value for W is 0, followed by the 2nd highest $W = 6 = 2 \cdot 3$. Taking average for $W = 1, \dots, 9$, we have the revised product $E[W|Y_1, Y_2] = 3.19$, which will be the primary element to evaluate the similarity for two items in Eq. (2). If we just use the disguised matrix R^Y, the product of $2 \cdot 3$ gives $6 > 3.19$. The difference would work to improve the prediction accuracy.

In the same way, we have the expected value of products of rating values for $Y_1, Y_2 \in V$ and the revised similarity matrix in **Tables 8** and **9**, respectively. We illustrate how the revised similarity is distributed as close as the original one in **Fig. 4**, where the revised similarity S^Z and the disguised similarity (without revision process) S^Y are plotted with respect to the original similarity S^X.

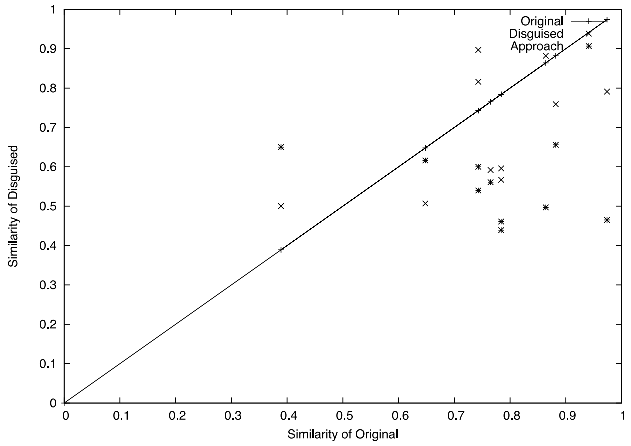

Fig. 4 Revised similarities with regard to original values, (S^X, S^Y, S^E).

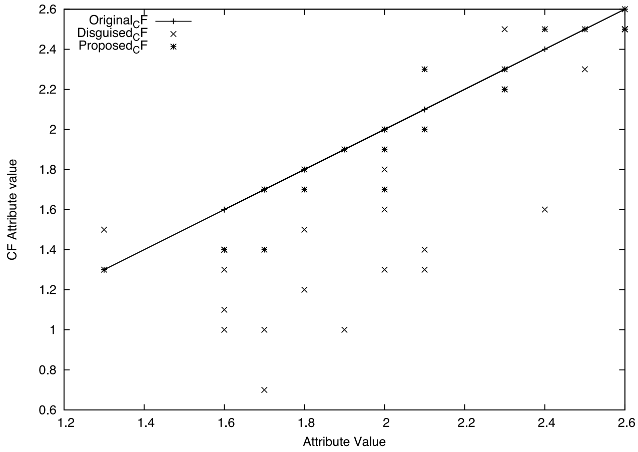

Fig. 5 Predicted rating values in collaborative filtering from original, disguised and revised similarities, r^X, r^Y and r^E.

Table 10 Mean absolute error.

	MAE	Standard deviation
Original	0.968	1.171
Disguised	1.033	1.204
Proposed	1.009	1.228

Based on the revised similarity, we apply the CF to predict the rating value for unseen items and show the distribution of predicted values with respect to the original values in **Fig. 5**. We show distributions from the original rating matrix X, which lies along the linear line, from the disguised matrix Y, and the reconstructed values Z, which are plotted closely to the original one.

We summarize the Mean Absolute Error (MAE) of prediction for the proposed scheme in **Table 10**, defined by $MAE^E = \sum_{u,i} |r^X_{u,i} - r^E_{u,i}|$ for the proposed scheme, and the MAE for the disguised matrix as $MAE^Y = \sum_{u,i} |r^X_{u,i} - r^Y_{u,i}|$. In the summary, the proposed scheme reduces the error in the prediction. The reason of the failure of prediction for some items includes the skew in the experimental matrix and the negative effect of a missing value.

4.2 Hypothesis Testing

The difference of MAE between the proposed scheme and the disguised one is 0.024 and may not be significant for other datasets. In order to verify the reliability of the experiment, we conduct a statistical hypothesis test.

The null hypothesis is

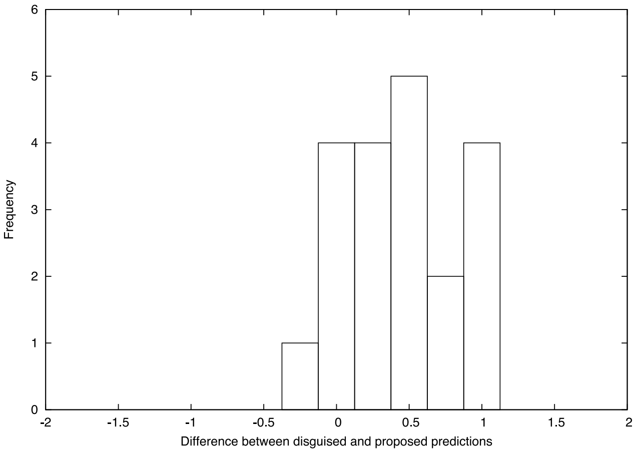

Fig. 6 Distribution of differences of prediction between the disguised and the proposed scheme.

$$H_0 : MAE^E \geq MAE^Y$$

and the alternative is

$$H_A : MAE^E < MAE^Y.$$

In our study, a rating value is predicted from a simple disguised similarity (disguised) and from the revised similarity (proposed). Hence, we examine the paired t-test for the analysis, rather than considering the two sets of observations to be distinct samples.

The mean of the set of differences is

$$\overline{d} = \frac{1}{nm} \sum_{u,i} r^y_{u,i} - r^e_{u,i} = 0.34$$

and the standard deviation of the difference is

$$s_d = \sqrt{\frac{r^y_{u,i} - r^e_{u,i} - \overline{d}}{nm - 1}} = 0.39789.$$

As **Fig. 6** shows, the differences can be considered to be approximately normally distributed, with the center being shifted toward positive. Therefore, H_0 can be tested by computing the statistic

$$t = \frac{\overline{d} - 0}{s_d / \sqrt{nm}} = 3.8214.$$

If the null hypothesis is true, the quantity has t distribution with $nm - 1 = 19$ degree of freedom and $p = 0.000576$. Rejecting the null hypothesis at 0.001 level, we conclude that there is a significant difference of prediction between the proposed scheme and the disguised one [2].

4.3 Robustness

Huang, Du and Chen claimed that the additive perturbation is not secure since using the PCA, the original data can be reconstructed and a private information can be disclosed [3]. In an additive perturbation such as Ref. [2], random numbers are independent for each attribute. Their correlations are zero. Hence, their variance is evenly distributed. The PCA transformation allows an attacker to remove the random numbers' variance as follows.

Let C be the covariance matrix derived from the disguised Y.

(1) Conduct PCA to get $C = Q \Lambda Q^T$, where Λ is a diagonal matrix of eigenvalues.

(2) Let $\hat{Q}$ be the first κ columns of Q.

Table 12 Comparison of privacy preserving collaborative filtering schemes.

scheme	Canny [7]	Polat and Du [2]	Proposed
randomization	cryptographic	additive perturbation	randomized response
accuracy	high (no error)	low (inaccurate if sufficient randomness provided)	better (low but can be improved via the expected value)
robustness	secure under computational assumption	weak (vulnerable by the PCA-based reconstruction)	better
performance	heavy (suffered by the encryption cost for each value)	lightweight and scalable	lightweight and scalable

Table 11 PCA-based reconstruction.

	MAE	$\kappa = 2$	$\kappa = 3$
additive perturbation [2]	0.47	1.7623	0.57666
randomized response	0.65	1.7744	1.4186

(3) Reconstruct the original data by $\hat{X} = Y\hat{Q}\hat{Q}^T$.

Their attack assumes that the data is disguised as $y_i = x_i + r_i$, where r_i is a random number chosen from a certain distribution. Hence, the assumption does not hold in the proposed scheme where the y is determined by a randomized response with the probability p.

We have verified the robustness of the randomized response against the PCA-based reconstruction. **Table 11** shows the MAEs of the reconstructed matrix from the randomized matrix in Table 6, where $\kappa = 2, 3 < m = 5$ are used. In comparison to the scheme [2] where random numbers in a uniform distribution from 0 to 1 are added to the original X, the randomized response is robust against the PCA-based reconstruction.

4.4 Comparison to Related Works

We compared the proposed scheme to the existing schemes in terms of accuracy, security and performance as summarized in **Table 12**. In comparison to the cryptographic approach, e.g., Refs. [7], [14], the proposed scheme is free from the computational cost to perform an additive homomorphic cryptosystem such as Ref. [9]. The randomized approaches, e.g., Ref. [2], can be performed quickly but the accuracy is lost. The proposed scheme is classified as a randomized approach, too. It is more robust against the PCA-based reconstruction than the simple additive perturbation.

5. Conclusion

We have proposed a new scheme for privacy-preserving collaborative filtering based on the posterior probability distribution generated through a Bayes reconstruction process.

Our experiment shows that the proposed scheme allows to revise the similarity between items and to predict the rating value for unseen items more accurately than predicting from the disguised data. The advantage of the perturbation in a randomized response is the higher degree of privacy of a personal rating in terms of the entropy of the disguised vector.

Acknowledgments The work has been supported by the project on security technologies for cloud computing funded by the Japanese Ministry of Internal Affairs and Communications.

References

[1] Agrawal, R. and Srikant, R.: Privacy-Preserving Data Mining, *ACM SIGMOD 2000*, pp.439–450 (2000).

[2] Polat, H. and Du, W.: Privacy-Preserving Collaborative Filtering using Randomized Perturbation Techniques, *ICDM 2003*, pp.1–15 (2003).

[3] Huang, Z., Du, W. and Chen, B.: Deriving Private Information from Randomized Data, *ACM SIGMOD 2005*, pp.37–48 (2005).

[4] Zhang, S., Ford, J. and Makedon, F.: Deriving private information from randomly perturbed ratings, *SIAM-Data Mining Conference* (2006).

[5] Zhang, S., Ford, J. and Makedon, F.: A Privacy-preserving Collaborative Filtering Scheme with Two-way Communication, *ACM EC '06*, pp.316–323 (2006).

[6] Cranor, L.F.: I Didn't Buy it for Myself, Privacy and E-Commerce Personalization, *WPES 2003*, Washington, DC, USA, pp.111–117 (2003).

[7] Canny, J.: Collaborative Filtering with Privacy, *IEEE Conf. on Security and Privacy*, Oakland CA (May 2002).

[8] Canny, J.: Collaborative filtering with privacy via factor analysis, *Proc. 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, *SIGIR '02*, pp.238–245, ACM (2002).

[9] Paillier, P.: Public-Key Cryptosystems Based on Composite Degree Residuosity Classes, *Proc. EUROCRYPT '99*, LNCS 1592, pp.223–238, Springer (1999).

[10] Breese, J.S., Heckerman, D. and Kadie, C.: Empirical analysis of predictive algorithms for collaborative filtering, *UAI*, pp.43–52 (2004).

[11] GroupLens Data Sets, available from ⟨http://grouplens.org/⟩ (accessed 2008-08).

[12] Resnick, P., Iacovou, N., Sushak, M., Bergstrom, P. and Riedl, J.: GroupLens: An open architecture for collaborative filtering of netnews, *Proc. 1994 Computer Supported Collaborative Work Conference* (1994).

[13] Katzenbeisser, S. and Petkovic, M.: Privacy-Preserving Recommendation Systems for Consumer Healthcare Services, *Proc. 2008 3rd International Conference on Availability, Reliability and Security* (*ARES 2008*), pp.889–895, IEEE Computer Society (2008).

[14] Ahmad, W. and Khokhar, A.: An Architecture for Privacy Preserving Collaborative Filtering on Web Portals, *Proc. 3rd International Symposium on Information Assurance and Security*, pp.273–278, IEEE Computer Society (2007).

[15] Kikuchi, H., Kizawa, H. and Tada, M.: Privacy-Preserving Collaborative Filtering Schemes, *WAIS 2009*, *ARES 2009 Federated Workshop*, IEEE Press (2009).

[16] Sarwar, B., Karypis, G., Konstan, J. and Riedl, J.: Item-Based Collaborative Filtering Recommendation Algorithms, *ACM WWW10*, Hong Kong (May 2001).

Hiroaki Kikuchi was born in Japan. He received his B.E., M.E. and Ph.D. degrees from Meiji University in 1988, 1990 and 1994. After working in Fujitsu Laboratories Ltd. from 1990, in Tokai University since 1994, respectively, he joined Meiji University in 2013. He is currently a Professor at the Department of Frontier Media Science, School of Interdisciplinary Mathematical Sciences, Meiji University. He was a Visiting Researcher at the School of Computer Science, Carnegie Mellon University in 1997. His main research interests are fuzzy logic, cryptographic protocol, network security, and privacy-preserving data mining. He is a member of IEICE, the Japan Society for Fuzzy Theory and Systems (SOFT), IEEE and ACM. He is a fellow of IPSJ.

Anna Mochizuki was born in Japan. She received his B.E. and M.E. degrees from Tokai University in 2010 and 2012. Her research interest is a privacy-preserving data mining. She is a member of IPSJ.