Survey of Confidentiality and Trust in Recommender Systems

Seyed Hossein HosseiniNazhad, Morteza Abdi Reyhan
Department of computer science, Faculty of engineering, Payame Noor University , Iran
E-mail: s.hosseininejad@gmail.com, abdi@yahoo.com

Abstract

Recommender systems has an important role in social networks. With the growth and development of social networks, this issue is becoming more and more important. Recommending systems try to predict the user's interests and then suggest the closest items to the user's tastes. Recommender systems analyze the user's behavior and suggest the most appropriate items. By collecting user information, the system categorizes and summarizes them, allowing users to access more relevant information in less time. Recommender system is an intelligent system that creates appropriate suggestions for each person by discovering and analyzing user information. In this paper, we will investigate recommending systems in three sections: types of recommending systems, information confidentiality and trust in recommender systems. We will refer to the related works in each section, review the challenges of them, and present our results and evaluation on these methods.

Keywords: Simulation, network, network simulation software, general purpose simulator, comparison

1. Introduction

Internet is growing rapidly and has become an opportunity to share knowledge as well as create social networks. The main purpose of recommendation systems is to generate meaningful recommendations to a group of users who are interested in that group of products or items.

Recommender systems try to guess the user's interests and suggest the closest and most suitable product to the user's tastes. Recommending systems, by analyzing the user behavior, suggest the most appropriate items (data, information, goods, etc.). This system is an approach to deal with the large and growing volume of information problems. Recommending systems offers a personalized offer to users who are looking for a specific type of information related to their priorities among a large amount of information. This system, gather users' behavior and movements, categorize and interpret them, and has made it possible for users to access more relevant information in less time [1,2].

2. What is a Computer Network

Simulation technology and software are one of the most powerful methods and tools available to managers, industry engineers, system analysts, and so on, which enables them to make systems, in hands, before making any decision about any production system, service, Modeling and simulating them, performing or working them, and making necessary statistical surveys in all its dimensions in order to make better decisions, with the goal of reducing costs and increasing profit (or efficiency). Using simulation, a wide range of dynamic (dynamic) issues can be analyzed in the areas of manufacturing, support, and services.

The simulation allows for modeling the flow of materials and goods, human resources and information in the organization, and analyzing and analyzing the system by simulating and adjusting different scenarios, 3D animations, and ... It was concerned with potential improvements.

3. Simulation

In this section some technical approaches that can reduce the risks of confidentiality and facilitate the realization of confidentiality are provided.

These techniques are not a sufficient condition for controlling the privacy risks of recommender systems, but these technologies should be used as a necessary condition in the design of user-friendly systems that also take into account the legal aspects as enhanced privacy protections.

Define users by nicknames

In recommender systems, it is possible for users to be anonymous and at the same time receive all personalization facilities. In an infrastructure with changing identity of users, that supports personalization, users tend to have the following capabilities (using a set of terms [5, 6]):

1- Unidentifiable: Neither the system nor third parties should be able to identify users with nicknames.

2- Linkable for the personalized system: The recommender system should be able to link each interaction to a specific user.

3- Unlinkable for third parties: A third party is not allowed to link the two stages of a user's interaction.

4- Unobservable: A third party cannot detect that a recommender system is being used by a given user.

Client-side personalization

Some researchers [7, 8, 9, 10, 11] worked on recommendation systems in which user data is deployed on the client side instead of on the server side. In addition, all personalization processes that depend on this data are performed only on the client side. In terms of confidentiality, this approach has two main advantages:

1-The size of the privacy issue becomes smaller because a small amount of users' personal data will be stored on the server. In fact, if a website with user-side personalization does not have control over the data intended to identify users with acceptable intentions, this will generally not be subject to the rules of confidentiality.

2- If personalization is done on locally stored data instead of using it remotely, users may want to share their information. The reason is that they know they have more control over their local physical environment.

However, client-side personalization has some challenges:

1- Conventional methods of user modeling and personalization, such as participatory modification that relies on the analysis of data from the entire user population, cannot be used or must be fundamentally redesigned.

2- Customization processes must also be performed on the client side, as temporary and partial transfer of personal data to the server is likely to negate the benefits of client-side personalization. But the program code used for personalization often

includes confidential business rules and methods, and precautions should be taken to avoid reverse-engineered disclosure. For this purpose, reliable computing platforms similar to those described by Kuroama and Langrinrich must be developed to ensure the integrity of their client-side set of private data [12, 13].

If these disadvantages do not pose a problem in a particular application area, the web-based consulting system designer should choose the client-side personalization as soon as the appropriate tools are provided. This, takes a big approach to data minimization and possibly increases user reliability.

Introducing four strategies to deal with privacy threats

Recommender systems gather a large amount of information about their users to find rules that allow future recommendations**.** Without confidentiality techniques, such databases may not always be reliable and may be an attractive target for unauthorized access. Client side personalization is not the solution to these privacy attacks. Here are some strategies to cope these risks:

1- Distribution

One possible strategy for better protection of individual data is to avoid using a central database that contains all users' data. Distributed clusters that contain only the information of some users can be used for this purpose. Distribution can also improve the performance and availability of the recommender system. In the Yenta system [14], for example, this method has improved the issue of confidentiality of information. PocketLens Distributed Collaborative Algorithm has gone a step further in avoiding data disclosure.

2- Aggregation of encrypted data

Kenny [15, 16] has proposed the use of a secure multidimensional computational model that, using homomorphic coding and P2P communications, allows users to confidentially maintain their private ratings and allows a community of such users to calculate the composition of their private information without disclosing.

For the future works, it will be possible to generate recommendation systems through the user's own rating on the client side. This method is prone to statistical vulnerabilities. The PocketLens system [17] also allows a community of users to calculate a similarity model without revealing their rates.

3- Perturbation

In this approach, user rates are provided to a central server that runs the entire algorithm. But these rates are changed before being assigned to the server, so that the real values of the rates are hidden from the server. Plott and Dow [18, 19] showed that by adding random numbers to user rates, acceptable recommendations can be result. If the number of items and users increases and also the standard deviation of the perturbation function decreases (the latter explicitly reduces confidentiality) then the quality of recommendation based on confusing data improves.

4- Obfuscation

In this approach, proposed by Berkowski et al., a certain percentage of user rates are replaced with different values before being presented to the central server to run the algorithm. It is assumed that users are free to choose the part of their data that should be ambiguous and to be able to deny the correctness of their exposing part of data**.**

4. Network Simulator

Some research [20, 21] has shown in the initial interactions of the user with the system, the number of rates is zero or low and similarity-based algorithms are not able to produce quality suggestions. Therefore using the trust network which each user has, can be a good alternative in early interactions with the user (in case of cold start problems and dispersion of the item / user matrix).

Therefore, it can be concluded that the discussion of trust in the design and construction of recommending systems can be used in three positions:

1- To give weight to users and assign higher weight to users who trust the same opinions.
2- Use in participatory refinement algorithm along with calculating similarity between users.
3- In sorting and refining user feedback and suggesting it to the user, using trust to prioritize trusted users.

Trust definition

Trust is generally a complex concept. There are many factors involved in trust, including personal background, activity background, similarity, credibility, and position. According to sociologists, trust is associated with belief and commitment.

In the recommending systems literature, the trust scale is used to measure the similarity of people's opinions to each other. Definitions of trust are placed in several categories, and it is not easy to provide a constant definition. Some researchers tried to give some computational form of trust and for this purpose, both the social and technical aspects were considered [22]. In recent years, some research has used trust in improving the quality of recommendations in recommender systems. In [25], trust relationships are modeled automatically from rates of users. Some other researches has used communication between users to model the trust.

Paolo Massa and Bobby Mucharji [23] have built a model of trust directly from the data of the Epinen.com website. In this study, based on the degree to which visitors have been found useful and reliable in the past, they are assigned a trust score. In a similar work on Epineen.com [24] data, a trust-aware recommendation architecture has been proposed that relies on a network of trust in which a user can trust other users in the system.Traditional participatory refining systems have relied on similarities between user rating profiles as a way of scoring the predictive share of different profiles. But research has shown that profile similarity alone may not be enough, and also other factors have impact. "Trust" is defined as the degree to which a person trusts a particular profile when predicting a score. However, as trust-based data usually has a level of confidentiality, therefore evaluating trust-based systems is a challenging problem. On the other hand, there are very few databases available to most researchers. Another problem with trust-based systems is the complexity of their algorithms, as well as the modeling of trust and weight dispreading are time consuming. However, trust is a new topic in research that has many challenges and a lot of research needs to be done to address the relevant topics.

Conclusion

This article reviews and compares the tools and simulators of the network. The research shows that this kind of simulator is due to network control to determine whether the network is capable of working in real time or not and has the capacity to reduce the time and cost required to test the functionality of the network. In this paper, we tried to compare and compare tabular characteristics and application of the 23 common simulators. Since the use of a network simulator can be effective in the performance of a project and other important issues in a laboratory research, depending on the specific characteristics of each simulator, the use of each of its types can vary depending on the application.

References

[1] Borking, J. J. and Raab, C. D., ―Laws, PETs and other Technologies for Privacy Protection.‖ Journal of Information, Law and Technology, (2001).

[2] Burkert, H. ―Privacy-Enhancing Technologies: Typology, Critique, and Vision.‖ In: Technology and Privacy: The New Landscape, Agre, P. E. and Rotenberg, M., Eds. Boston, MA: MIT Press (1997).

[3] Goldberg, I. A., ―Privacy-Enhancing Technologies for the Internet, II: Five Years Later.‖ In: Privacy Enhancing Technologies – Second International Workshop, PET 2002, Dingledine, R. and Syverson, P., Eds. Berlin - Heidelberg: Springer Verlag (2003).

[4] van Blarkom, G. W., Borking, J. J., and Olk, J. G. E., Eds., ― Handbook of Privacy and Privacy-Enhancing Technologies: The Case of Intelligent Software Agents.‖ The Hague, the Netherlands: TNO-FEL (2003).

[5] ISO: ISO/IEC 15408-2: Information Technology — Security Techniques — EvaluationCriteria for ITS Security: Part 2: Security Functional Requirements. (1999).

[6] Pfitzmann, A. and Köhntopp, M., ―Anonymity, Unobservability, and Pseudonymity: A Proposal for Terminology.‖ In: Anonymity 2000, Federrath, H., and Ed. Berlin-Heidelberg, Germany: Springer-Verlag (2001).

[7] GVU: GVU's 10th WWW User Survey. Graphics, Visualization and Usability Lab, Georgia Tech (1998).

[8] Cassel, L. and Wolz, U., ―Client Side Personalization. DELOS Workshop: Personalisation and Recommender Systems in Digital Libraries.‖ Dublin, Ireland (2001).

[9] Canny, J., ―Collaborative Filtering with Privacy.‖ IEEE Symposium on Security and Privacy, Oakland, CA, 45-57, DOI 10.1109/SECPRI.2002.1004361, (2002).

[10] Ceri, S., Dolog, P., Matera, M., and Nejdl, W., ―Model-Driven Design of Web Applications with Client-Side Adaptation.‖ In: Web Engineering: 4th International Conference, ICWE 2004, Koch, N., Fraternali, P., and MartinWirsing, Eds. Berlin – Heidelberg: Springer Verlag (2004).

[11] [Mulligan, D. and Schwartz, A., ―You're Place or Mine? Privacy Concerns and Solutions for Server and Client-Side Storage of Personal Information.‖ Computers, Freedom & Privacy Conference (1999).

[12] Coroama, V., ―The Smart Tachograph: Individual Accounting of Traffic Costs and Its Implications.‖ In: Pervasive Computing: 4th International Conference, PERVASIVE 2006,, Fishkin, K. P., Schiele, B., Nixon, P., and Quigley, A., Eds. Berlin – Heidelberg: Springer Verlag (2006).

[13] Coroama, V. and Langheinrich, M., ―Personalized Vehicle Insurance Rates: A Case for Client-Side Personalization in Ubiquitous Computing.‖ Proceedings of PEP06, CHI 2006 Workshop on Privacy-Enhanced Personalization, Montreal, Canada (2006).

[14] Foner, L. N., ―Yenta: A Multi-Agent Referral-Based Matchmaking System.‖ International Conference on Autonomous Agents, Marina del Rey, CA (1997).

[15] Canny, J., ―Collaborative Filtering with Privacy.‖ IEEE Symposium on Security and Privacy, Oakland, CA (2002).

[16] Canny, J., ―Collaborative Filtering with Privacy via Factor Analysis.‖ 25th Annual

International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2002), Tampere, Finland (2002).

[17] Miller, B. N., Konstan, J. A., and Riedl, J., ―PocketLens: Toward a Personal Recommender System.‖ ACM Transactions on Information Systems 22, (2004).

[18] Polat, H. and Du, W., ―Privacy-Preserving Collaborative Filtering.‖ International Journal of Electronic Commerce 9, (2003).

[19] Polat, H. and Du, W., ―SVD-based Collaborative Filtering with Privacy.‖ ACM Symposium on Applied Computing, Santa Fe, New Mexico (2005).

[20] Agrawal, D. and Aggarwal, C. C., ―On the Design and Quantification of Privacy Preserving Data Mining Algorithms.‖ 20th ACM SIGACT-SIGMOD-SIGART Symposium of Principles of Database System, Santa Barbara, CA (2001).

[21] Massa, P., and Avesani, P., ―Trust-aware recommender systems.‖ In Proceedings of RecommenderSystems, (2007).

[22] Marsh, S., ―Formalising trust as a computational concept.‖Ph.D. Thesis. Department of Mathematics and Computer Science, University of Stirling, (1994).

[23] Massa, P., Bhattacharjee, B., ―Using trust in recommender systems: an experimental analysis.‖ Proceedings of 2nd International Conference on Trust Managment, Oxford, England, (2004).

[24] Massa, P., Avesani, P., ―Trust-aware collaborative filtering for recommender systems. ― Proceedings of International Conference on Cooperative Information Systems, Agia Napa, Cyprus, (2004).

[25] O'Donovan, J., Smyth, B., ―Trust in recommender systems.‖ In IUI '05: Proceedings of the 10th international conference on intelligent user interfaces, pages 167–174. ACM Press, (2005).