ERCIM "Alain Bensoussan" Fellowship Scientific Report

Fellow: **Georgios Pitsilis**
Visited Location: **University of Luxembourg, FNR, Luxembourg**
Duration of Visit: **12 Months, 5th May 2010 – 4th May 2011**

I - Scientific activity

During my fellowship at University of Luxembourg I explored new directions in my areas of knowledge, that is Recommender Systems and Trust management, and also looked into other related fields of interest. Having already established collaboration with researchers from my previous employment, one of the objectives for this year was to continue work that has been successful. Actually, the new post was also found to be a nice opportunity to extend my research to new directions. In that respect, collaborations with new colleagues, experts in the area of data mining, led to producing interesting research outcome.

One part of the research activity includes extension of work done during my previous employment, with PhD Student P.H.Chia from Q2S-NTNU. In the present work we investigated the involvement of user's trusting behaviour into the selection of neighbours in a Recommender System, for improving the prediction accuracy. More specifically, we reasoned the use of new metrics for expressing distance between users, as the main criterion for selecting neighbours. This work led to a journal publication. (See publications).

The subject of collaboration with PhD student M-El.Hadedy from Q2S-NTNU was security in cloud applications. More particularly a system and a protocol that makes use of hash functions for protecting the authorship of digital media from attackers were proposed. The core of the idea we developed was to watermark secret messages inside digital media in a copy-resistant way. That makes the real owner of the content identifiable, no matter if the media has been copied and modified by some attacker who may claim the authorship afterwards. The practical benefit as well as the easy deployment of this concept onto today's infrastructures was demonstrated in a paper we finally submitted for publication to a conference.

It is interesting to note that I had also been given the opportunity to collaborate with security experts from the Interdisciplinary Center of Security and Trust (SnT Centre) of the University of Luxembourg. More particularly, the collaboration with Posdoc Dr. Wei Wang, expert in intrusion detection, led to the publication of two full papers. In the first one, performed a comparison analysis between mainstream and new clustering algorithms used for detecting attackers from http requests. The results demonstrated that the newly developed clustering algorithms are more effective for the above task. The second one concerned with the performance advantages received by the use of clusters into the user communities of a Recommender System. In this direction, the use of new clustering algorithms, such as *Affinity Propagation* was explored. Despite the notable impact on coverage, the proposed technique was found to improve quite significantly the accuracy of predictions.

Furthermore, the interesting results received gave rise to further exploring how to overcome the weaknesses of clustering in Recommender Systems. For this reason we developed a technique called *Cluster Abstraction* for improving coverage. In addition, we developed and tested a novel algorithm for clustering the users of a Recommender System using minimal information, such as the explicit trust obtained from Social Connectivity graphs.

Other research activities include co-authoring of a paper with a Master student from KAUST University and his supervisor. In this work a clustering technique was applied to a Recommender System to reduce the impact of Shilling Attackers to the prediction performance.

During my fellowship I had also been allocated tasks of reviewing papers for conferences and journals. In total I reviewed 2 papers for the following events:

- STM'10, Sixth International Workshop on Security and Trust Management, 1 paper allocated by Prof. Mauw.
- JTAER. Journal of Theoretical and Applied Electronic Commerce Research. 1 Paper allocated by the chief editor.

II- Publication(s) during your fellowship

Wang, W., **Zhang, X**., **Pitsilis, G.**, "Abstracting Audit Data for Efficient Anomaly Intrusion Detection" in proc ICISS 2010,Sixth International Conference on Information Systems Security, 15-19 Dec. 2010,DA-IICT, Gandhinagar Gujarat, India.

Abstract: High speed of processing massive audit data is crucial for an anomaly Intrusion Detection System (IDS) to achieve real-time performance during the detection. Abstracting audit data is a potential solution to improve the efficiency of data processing. In this work, we propose two strategies of data abstraction in order to build a lightweight detection model. The first strategy is exemplar extraction and the second is attribute abstraction. Two clustering algorithms, Affinity Propagation (AP) as well as traditional *k*-means, are employed to extract the exemplars, and Principal Component Analysis (PCA) is employed to abstract important attributes (a.k.a. features) from the audit data. Real HTTP traffic data collected in our institute as well as KDD 1999 data are used to validate the two strategies of data abstraction. The extensive test results show that the process of exemplar extraction significantly improves the detection efficiency and has a better detection performance than PCA in data abstraction.

Pitsilis, G., Zhang, X., **Wang, W**., "Clustering Recommenders in Collaborative Filtering Using Explicit Trust Information", to appear in proc. IFIPTM 2011 International Conference on Privacy, Trust Management and Security, Copenhagen, Denmark, 28th June- 1st-July 2011.

Abstract: In this work, is explored the benefits of combining clustering and social trust information for Recommender Systems. We demonstrate the performance advantages of traditional clustering algorithms like k-Means and we explore the use of new ones like Affinity Propagation (AP). Contrary to what has been used before, we investigate possible ways that social-oriented information like explicit trust could be exploited with AP for forming clusters of high quality. We conducted a series of evaluation tests using data from a real Recommender system Epinions.com from which we derived conclusions about the usefulness of trust information in forming clusters of Recommenders. Moreover, from the results we conclude that the potential advantages in using clustering can be enlarged by making use of the information that Social Networks can provide.

Chia, H.P., Pitsilis, G., "Exploring the Use of Explicit Trust Links for filtering Recommendations: A Study on Epinions.com", to appear in Special Issue on Trust management, Journal of Information Processing, Vol.19., pp.1-13, July-2011, Information Processing Society of Japan, 2011.

Abstract: The majority of recommender systems predict user preferences by relating users with similar attributes or taste. Prior research has shown that trust networks improve the accuracy of recommender systems, predominantly using algorithms devised by individual researchers. In this work, omitting any specific trust inference algorithm, we investigate how useful it might be if explicit trust relationships are used to select the best neighbors or predictors, to generate accurate recommendations. We conducted a series of experiments using data from

Epinions.com, a popular recommender system. We conducted a series of experiments using data from Epinions.com, a popular recommender system. We find that for highly active users, using trusted sources as predictors does not give more accurate recommendations compared to the classical similarity-based collaborative filtering scheme. This cautions against the intuition that inputs from trusted sources would always be more accurate or helpful. The use of explicit trust links, however, provides a slight gain in prediction accuracy when it comes to the less active users. These findings highlight the need to better understand the properties of trust when employing it in recommender systems.

Another 3 papers have been submitted to conferences and are still under review.

III -Attended Seminars, Workshops, and Conferences

I attended (will attend) the following scientific conferences at which I also gave (will give) presentation of my work.

- IFIPTM 2011, Fifth IFIP WG 11.11 International Conference on Trust management, June 28th- July 1st , 2011, Technical University of Denmark, Copenhagen, Denmark.

I also participated in one local event:

- LPWST, 2010. 1st Luxembourg-Polish Workshop on Security and Trust, May 20-22, 2009. May 6-7, Castle of Bourglinster, Luxembourg.

Also attended all the weekly seminars of the security group in the SaToSS group, where the work of the group members was presented. (phd-students and postdocs.) I also had the chance to present once the progress of my work in this event. I also participated in the Term meetings.

IV – Research Exchange Programme (12 month scheme)

First Period: One week at Q2S, NTNU, Trondheim, Norway. (14th – 19th February 2011).
Project group: Network Security headed by Prof. Svein Knapskog.
Scientific contact: Prof. Svein Knapskog. (email: knapskog@q2s.ntnu.no)
During my visit at NTNU I had the opportunity to collaborate with Mohamed El-hadedy, a PhD student from the security group. In this collaboration we developed an idea for a system that can be used for protecting the authorship digital multimedia. The outcome of this collaboration was captured in a paper which has now been submitted to a conference.
During this visit I also had the chance to present my current work to the group and discuss my ideas with the members.

Second Period: One Week at ICS-FORTH, Heraklion, Greece. (11th April – 16th April 2011).
Project Group: Distributed Systems group headed by Prof. Evangelos Markatos.
Scientific contact: Prof. Evangelos Markatos. (email: markatos@ics.forth.gr).
During my stay in FORTH I gave a presentation in the distributed systems group of my recent work and findings. I also attended the weekly meetings and seminars of the group at the time of my stay. Ideas for possible collaboration in the field of securing digital artefacts were also discussed.

Acknowledgments

I would like to thank my scientific contact Prof. Mauw for welcoming me in his research group and the administrative and technical stuff for the help they provided in many occasions.