International Journal of Information Processing, 9(3), 106-116, 2015
ISSN : 0973-8215
IK International Publishing House Pvt. Ltd., New Delhi, India

Recommender Systems to Improve Response Time of Computationally Intensive Tasks Distributed on Social Networking Platform

Varsha G Maragi[a], Prathibha M V[a], Lavanya B[a], Annappa B[a]

[a]Computer Science and Engineering, National Institute of Technology Karnataka, Surathkal, Mangalore-575025, India Contact: varshamaragi@gmail.com, prathibhamv92@gmail.com, whitesnowfall50@gmail.com, annappa@ieee.org

Recommendation systems are being harnessed in most of the e-commerce websites, thereby proving that they are beneficial assets that help the buyers as well as the sellers. Social networks on the other hand are slowly turning out to be platforms for resource sharing. In this paper, we present a novel idea of harnessing social networks to perform various computationally intensive tasks and improve the response time of these tasks by using recommender systems in order to target the right set of service providers among social network users. The paper describes the techniques used to build such a recommendation system and reflects upon the various issues that concern it and further elaborates on the solution for each of these issues. The idea of using recommendation systems may be further extended to any crowdsourcing Internet systems and for open source middleware systems that make use of idle CPU cycles of Internet users.

Keywords : Crowd-Sourcing System, Distributed Computing, Recommendation Systems, Resource-Discovery, Social Networks

1. INTRODUCTION

The need for resources in order to process computationally intensive tasks has been on the rise. A distributed system proves to be a more cost-efficient, reliable and effective solution to processing of these tasks than a single computer. The process of distributing these tasks can be taken to a new level by harnessing the resources in social networks and we hope that this would bring some major advancement in the field of distributed computing systems. We extend the idea described in [1], thereby reducing the need for infrastructure for setting up a distributed system and enhancing response time in crowd sourcing systems. It is known that the friends in social networks are more open to accept your computationally intensive task, ready to process and give back results when compared to any unknown Internet user because of the small-world property that characterizes any social network. Hence targeting the right social network user proves beneficial. This technique will significantly help projects such as SETI@Home, Folding@Home, Einstein@Home, BOINC which gather a gigantic pool of resources on the Internet, by using computers from any Internet user, allowing them to process data much quicker than in traditional supercomputers.

Recommender systems are widely used nowadays in order to help the e-commerce sites to select the right set of customers for their products. Recommender systems can be used in the field of distributed computing systems to suggest service providers whose idle CPU cycles can be harnessed to perform any computationally intensive task. We use the collaborative filtering technique in order to suggest a set of social network users who are capable of accepting and processing any computationally intensive task like video processing, image processing, climate prediction *etc.* The paper elaborates the various issues that need to be considered which includes the cold-start proboccupation. However, while distributing tasks, the Trans-SocialDP mentioned in [1] makes use of the messaging module as the main module to communicate with all social networks. All the task related results and queries are sent as messages. Hence, the chat logs play an important role while determining the area of interest of these users. Hence, we propose to make use of these chat logs of social network users in order to determine the similarity between the requestor and his friends.

For each user, we can generate a document of the users entire conversation history. TF-IDF, which is the Term-Frequency-Inverse Document Frequency (reflects the importance of a word in a document), can be made use of in order to calculate the cosine-similarity between the chat log documents of the requestor and the service provider. This factor can prove to be very effective and can help the recommender system to a large extent.

5.5. Time of Chat

We observed that in many chat log datasets, the time of chatting is also specified. This detail can be put to use thereby helping the recommendation system. The work pattern of the social network users can be deduced from this information *i.e.*, we can deduce whether the user logs into the social network during the day time or at night. This detail can then be used to send request messages or task related messages at the particular time of the day that the user is online. Also, many other deductions like the amount of time that the user is online, the mood of the user *etc.*, can be acquired from these chat datasets and they can be incorporated as various factors that can determine the neighbourhood dataset in the recommender system.

We have restricted ourselves to finding the chat time pattern. We used the dataset from an online community for students at University of California, Irvine to find the average timings at which each user of this social network chats. This dataset contains number of chat messages and time at which these messages were passed for a particular user over a period of 11 days.We averaged out the chat timings for each user of this network. The results were plotted. ,

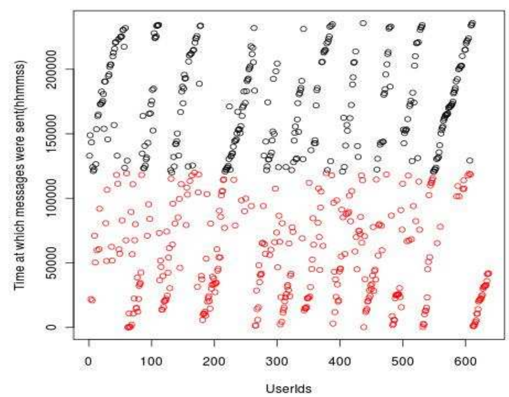

Figure 8. User Id Vs. Time of Chat

The x-axis represents the user-ids and the y-axis represents the time at which the messages were sent(hhmmss). The red plots are the users who generally chat during day-time while the black ones are the users who generally prefer to chat during the night time during this particular 11 day period. Hence, this information can be useful for the recommender system to target tasks to the target users at a particular time of the day.

6. CONCLUSIONS

The objective of this paper is to enhance response time of distributed computing applications by harnessing social network users idle CPU cycles and using the recommendation system in order to select the service providers among social network users. The recommendation system uses the CF technique and further modifications are suggested to this system in order to address cold-start problem, malicious nodes and diversification of the recommendation list.

It should be noted that such a recommendation system is implemented in the centralized servers of a social network and the local client applications query them in order to obtain a list of service providers. Also, the algorithms are tested against Advogato dataset and Epinions dataset (used for evaluating the general CF system) which are trust scores and product ratings.

Overall, the results that are achieved by implementing the proposed algorithms and constraints prove that the modified CF recommender system is beneficial for any crowdsourcing system. Further improvements in terms of considering various other factors that help the recommender system choose the right set of service providers and rating them, need to be considered. We hope that our module will enhance and benefit any novel cycle-sharing models.

REFERENCES

1. Nuno Apolonia, Paulo Ferreira and Luis Veiga. Trans-Social Networks for Distributed Processing *In Proceeding of the* 11^{th} *International IFIP TC6 conference on Networking*, 2012.
2. David P Anderson, Jeff Cobb, Eric Korpela, Matt Lebofsky and Dan Werthimer. SETI@Home *Magazine Communications of the ACM,* 45:56-61, November 2002.
3. Luis Veiga, Rodrigo Rodrigues and Paulo Ferreira. GiGi: An Ocean of Gridlets on a Grid-for- the-Masses *In Proceedings of* 7^{th} *IEEE International Symposium on Cluster Computing and the Grid, IEEE Computer Society*, 2007.
4. David P. Anderson BOINC: A System for Public-Resource Computing and Storage *In Proceedings of the* 5^{th} *IEEE/ACM International Workshop on Grid Computing, IEEE Computer Society*, pages 4–10, 2004.
5. F Boldrin, C Taddia and G Mazzini. Distributed Computing Through Web Browser, *in IEEE* 66^{th} *Vehicular Technology Conference (VTC)*, pages 2020–2024, 2007.
6. Jianming He, Wesley W Chu. A Social Network-based Recommender System *University of California Los Angeles, CA, USA*, 2010.
7. Kyle Chard, Simon Caton and Omer Ranaand Kris Bubendorfer. Social Cloud: Cloud Computing in Social Networks, *in IEEE 3rd International Conference on Cloud Computing (CLOUD),* 2010.
8. Alan Mislove, Krishna P Gummadi and Peter Druschel. Exploiting Social Networks for Internet Search, *In Proceedings of the* 5^{th} *Workshop on Hot Topics in Networks (HotNets-V)*, pages 79–84, Irvine, CA, 2006.
9. Jennifer Ann Golbeck. Computing and Applying Trust in Web-Based Social Networks, *Digital Repository at the University of Maryland*, 2005.
10. Paolo Massa and Paolo Avesani, Trust-aware Collaborative Filtering for Recommender Systems, *In Proceedings of the 2007 ACM Conference on Recommender Systems*, pages 17–24, New York, USA, 2007.
11. Daire ODoherty, Salim Jouili Eura Nova and Peter Van Roy. Trust-Based Recommendation:An Empirical Analysis, *In Proceedings of the* 2^{nd} *ACM SIGMOD Workshop on Databases and Social Networks*, 2012.
12. Sonnek J D, Weissman J B. A Quantitative Comparison of Reputation Systems in the Grid *In 6th IEEE/ACM International Workshop on Grid Computing*, 2005.
13. Martijin C Willemsen, Bart P Knijnenburg, Mark P Graus. Using Latent Features Diversification to Reduce Choice Difficulty in Recommendation Lists *In Proceedings of the* 4^{th} *ACM Conference on Recommender Systems*, 2010.

Varsha G Maragi is currently working as a Software Engineer in Samsung R&D Bangalore. She recieved her BTech Degree in Computer Engineering from National Institute of Technology Karnataka, Surathkal, Mangalore, India in the year 2014. Her Her research interests include Distributed Computing Systems, Social Networking, Machine Learning and Cloud Computing. She has worked on numerous projects and internships in these areas, including a winterschool internship from CMU,IISc and an idustrial internship at a start-up incubated in IIMB. She has a paper publication in Elsevier and received "Best-Paper Award" for the same in the Conference.

Annappa B is currently working as Associate Professor and Head of the Department in the Department of Computer Science and Engineering, National Institute of Technology Karnataka, Surathkal, Mangalore, India. He has more than 20 years of experience in teaching and research. He holds Ph.D and M.Tech in Computer Science and Engineering from National Institute of Technology Karnataka Surathkal and BE from UBDT College of Engineering, Davangere affiliated to Mysore University, Karnataka. He is a Senior member of IEEE, ACM and Life member of Computer Society of India, Institution of Engineers (INDIA) and Indian Society of Technical Education. His research interests include Cloud Computing, Big Data Analytics, Distributed Computing, Software Engineering and Process Mining. He has published more than 75 papers in International conferences and Journals. He was the Organizing Chair of International conference ADCONS-2013 and he is in the TPC of many International conferences and reviewer of Journals