Influential Users Detection in Online Social Networks

Nouran Ayman Roushdy Abd Al-Azim

A thesis submitted to the department of Information Systems,

Faculty of Computer and Information Sciences,

Ain Shams University.

In partial fulfilment of the requirements for the degree of Master of Science

in Computer and Information Sciences

Supervised by:

Prof. Tarek F. Gharib

Head of Information Systems Department

Faculty of Computer and Information Sciences

Ain Shams University

Assistant Prof. Mohamed Hamdy

Information Systems Department

Faculty of Computer and Information Sciences

Ain Shams University

Dr. Yasmine Afify

Information Systems Department

Faculty of Computer and Information Sciences

Ain Shams University

2020

Acknowledgements

I would like to thank Prof. Tarek F. Ghraib for his continuous guidance, motivation, support and for enlightening me the first glance of research.
I express my sincere gratitude to Dr.Mohamed Hamdy for his immense knowledge and fruitful supervision.
I would like to show by appreciation to my life-long role model Dr. Yasmine Afify for her continuous advices and for the sleepless nights.
I must express my very profound gratitude to papy and mummmy for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.
Mahy and Dido, you are the ones who pull me to earth when I feel so down, you are the ones who defend me, thank you, and I love you both.
I would like to show my full appreciation to my grandparents, I hope all of you are here today with me. Nanna and Daddi, you are my source of joy and delectation. Gedo and Toufa, you are my source of endurance. I am looking forward to making you proud. I love you unconditionally.
Mariam and Lobna, you are always here for my ups and downs. Our delightful moments are priceless. I love you.
FCIS-IT team, you are the most dependable, trustful and friendly team I have ever met. I want to express my sincere thanks to all of you.

Finally, I owe thanks to a very special person, my fiance', Faysal for his continued and unfailing love, support and understanding during my pursuit of master degree that made the completion of thesis possible. You are always around at times I thought that it is impossible to continue, you helped me to keep things in perspective. I greatly value your contribution and deeply appreciate your belief in me.

Abstract

Social networks are considered one of the main merits of this era. People worldwide use this online platform to build their own social ties. The key feature behind the success of social networks is *microblogging*. This feature facilitates the interactions between people around the globe. People use social networking platforms to share their ideas, populate their believes and find other people with the same preferences.

Social network users tend to interact with each other by sharing, commenting and reacting to disseminated content. These interactions help in the content spread across the network. The dynamics of user interest in the disseminated content leads to the clustering of social network users to varying groups (communities) called *"interest groups"*. The analysis of users behaviour raises some crucial questions about who is responsible for content spread, the roles played by users in an interest group, the user rank based on his/her role and the rank of the interest group as a whole.

Our research objective is to propose ranking models that take into consideration the dynamic nature of social networks topology and the users interest to tackle the previously mentioned limitations. In order to achieve this objective, four models are proposed. First, Influence Ranking Model (IRM) which aims to rank all the social network users based on their interactivities. It introduces the usage of weighted and directed graph with the classic kshell

decomposition methodology. The uniqueness of the obtained ranking list of IRM is on average equals 1 and network coverage is improved by 0.3%.
Second, Interest Group Identification (IGI) model which aims to cluster users based on their interest in the disseminated content. The quality of separation of IGI reaches 0.923.
Third, Influence Propagation (IP) model which aims to identify the role played by each member in the interest group to spread content to other members. Moreover, a new role called *"ultimate observer"* is introduced. Then these roles are used to rank interest group members based on their contribution in content dissemination. The distinction of ranking of IP is on average equals 1 and its network coverage is competing with the benchmark approaches.
Finally, UltRank model which aims to rank the interest group as a whole to reach the goal of ultimate rank using a new reachability metric. It takes in consideration: 1.Distance from interest group to the other groups. 2.Size of the interest group compared to the size of all the reachable interest groups. 3.Number of reachable interest groups with respect to the number of all interest groups in social network. Meanwhile, new role called *"bridge nodes"* is presented. The ranking capability and network coverage results of UltRank is improved by 1.6% and 4% respectively compared to the benchmark approach These promising results encourage the employment of the proposed models in different applications in social networks such as viral marketing, monitoring public opinion, event prediction and recommendation systems.

Contents

List of Figures

List of Tables