Mona Maghraby

بسم اللّه الرحمن الرحيم

مركز الشبكات وتكنولوجيا المعلومات

قسم التوثيق الإلكتروني

Mona Maghraby

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها

علي هذه الأقراص المدمجة قد أعدت دون أية تغيرات

Ain Shams University
Faculty of Computer and Information Sciences
Information Systems Department

Effective Approaches for Influence Maximization in Social Networks

A thesis submitted to the Department of Information Systems
Faculty of Computer and Information Sciences
Ain Shams University, Cairo, Egypt

In partial fulfillment of the requirements for the Degree Doctor of Philosophy in Computer and Information Sciences

BY

Ahmed Mohamed Samir Ali Gamal Eldin
MSc. of Computers and Information, Information Systems Department
Assistant Lecturer at Software Engineering Department, Faculty of Computers and Information, Kafrelsheikh University

Under the Supervision of

Prof. Dr. Tarek Fouad Gharib
Professor, Head of Information Systems Department,
Faculty of Computer and Information Sciences,
Ain Shams University

Prof. Dr. Osama Mohamed Abo-Seida
Professor, Dean of Faculty of Computers and Information,
Faculty of Computer and Information,
Kafrelsheikh University.

Dr. Sherine Rady Abdel Ghany
Associate Professor,
Information Systems Department,
Faculty of Computer and Information Sciences,
Ain Shams University

February 2022

Acknowledgement

First and foremost, all thanks and praise are due to Allah, who has granted me this great success in my PhD.

I would like to express my gratitude to numerous people who have helped me to complete my work over the last number of years. I would like to express my utmost gratitude to my supervisor **Prof. Dr Tarek Fouad Gharib** for all his guidance, patience, feedback, encouragement, and for his great support personally and professionally throughout my PhD journey. I could not have done it without him.

I would like to express my utmost gratitude to my supervisor **Dr Sherine Rady** for all her great efforts and great support throughout my PhD journey. Really I have learned so much from her, and all her feedback have added great value personally and professionally.

I would like to express my utmost gratitude to my supervisor **Prof. Dr. Osama Mohamed Abo-Seida** for all his great help and great support throughout my PhD journey.

I would like to express my utmost gratitude to my Wife **Eman Farid** for her great support and her encouragement throughout my PhD journey.

I would like to express my utmost gratitude to the soul of my father **Mohamed Samir**, he told me from more than 12 years that I should fight for my dreams, May Allah bless his soul. I would like to express my utmost gratitude to all my family and my friends especially my best friend **Kareem Saeed** for their encouragement throughout my PhD journey

And I've saved the best for last: My utmost gratefulness and sincerest thanks to my mother (**Rokaia Abdel Moneim**), whom I wish that she is with me in this moment, I will never ever forget your face, smile, help and your support in all my life, and especially in my PhD journey, May Allah bless your soul.

Abstract

The detection of the top influential users is well-known scientifically as the social influence maximization. The current existing solutions suffer from several limitations, such as the highly required computations and the running time to find the top influential seed set. Therefore, finding an effective and efficient solution is still a challenging task. In order to solve the current scientific gap, this thesis proposes an effective and scalable community-based approach for the influence maximization problem called Louvain-k-shell Generalization (LKG). LKG is a fast and scalable community-based hybrid approach to detect top influential users in social networks. The LKG hybrid approach consists of three phases: 1) Community detection, in which the complete social network is partitioned into related communities using the Louvain algorithm; 2) Community top nodes detection that applies the k-shell decomposition locally in each portioned community; and finally 3) Selection generalization, in which the prior obtained results are generalized over the whole network for maximizing the global spread of influence. The results of the LKG approach have been shown to achieve better results for the spread of influence using incomplete social networks than the existing related work approaches and with far much less processing time.

An efficient method is presented to enhance the selection criteria for generated communities. The improved community-based approach is called Louvain CRANK-Select (LCS), which is based on CRANK algorithm for better ranking the generated communities that will be used to select the top influential seed set.

The proposed LKG approach has been also applied on a practicable application for the problem of the influence maximization to analyze a sample of Twitter data concerning Covid-19 epidemic. The effective positive influential users' identification is suggested to help health organizations to share and publish a useful and a helpful information about the latest update of the epidemic.

Table of Contents

List of Abbreviations

ACO-IM	Ant Colony Optimization- Influence Maximization
APIs	Application programming interfaces
ASIM	A Scalable Algorithm for Influence Maximization
CD	Credit Distribution
CDIM	Continuous-Time Markov Chain into the Independent Cascade Model.
CELF	Cost-Effective Lazy Forward
CELF++	Improved Cost-Effective Lazy Forward
CFIN	Community Finding Influential Node
CIM	Community based Influence Maximization
CoFIM	Community based For Influence Maximization
COND-MAT	Condense Matter Physics
Covid-19	Coronavirus disease of 2019
CPSP-Tree	Community-aware Partial Shortest Path Tree
CRANK	Community prioritization model
CTMC-ICM	Continuous-Time Markov Chain into the Independent Cascade Model.
DAG	Directed Acyclic Graph
DDSE	Degree Descending Search Strategy
DSE	Descending Search Strategy
DLIM	Degree discount and Local Improvement Method
DomIM	Dominating Set for Influence Maximization
GNA	Genetic New Greedy Algorithm
GRASP	Greedy Randomized Adaptive Search Procedure
GWIM	Gray Wolf based Influence Maximization

HC	Heuristic clustering
IC	Independent Cascade
IM	Influence Maximization
IMM	Influence Maximization via Martingales
IPA	parallel algorithm for influence maximization problem
IRIE	Influence Ranking (IR) and Influence Estimation (IE)
LCS	Louvain CRANK-Select
LDAG	Local Directed Acyclic Graph
LGIEM	Global and local node influence based community detection
LIE	Local Influence Estimation
LKG	Louvain K-shell Generalization
LT	Linear Threshold
MCDM	Multi-Criteria Decision Making
MCS	Monte Carlo Simulations
MIA	Maximum Influence Arborescence
MLIM	Maximum likelihood-based scheme under the Independent Cascade(IC) model
NAV	Node Approximate Influence Value
NP-hard	Non-deterministic polynomial-time hardness
PGP	Pretty Good Privacy
PMIA	Prefix excluding MIA
SA	Simulated Annealing
SIMPATH	Simple Paths in the neighborhood for influence maximization
SIR	Susceptible Infected Recovered
SKIM	Sketch-Based Influence Maximization
SLPA	Speaker-Listener Label Propagation Algorithm
SNAP	Large Net- work Dataset Collections

SPIN	Shapley Value-Based Discovery of Influential Nodes
SSA	Stop-and-Stare Algorithm
WHO	World Health Organization
WIC	Weighted Independent Cascade

List of Figures

List of Tables

List of Publications

[1] A. M. Samir, S. Rady, and T. F. Gharib, "LKG: A fast scalable community-based approach for influence maximization problem in social networks," Physica A: Statistical Mechanics and its Applications, vol. 582, p. 126258, July 2021. Elsevier, IF: 3.263.

[2] A. M. Samir, S. Rady, and T. F. Gharib, "An efficient community-based approach for the influence maximization problem in social networks," IEEE Tenth International Conference on Intelligent Computing and Information Systems (ICICIS), pp. 335-340, 2021.

[3] A. M. Samir, S. Rady, and T. F. Gharib, "The Identification of the Top Positive Influential Users of the Social Networks to Help in the Control of Covid-19 Spread", Submitted to International Journal of Intelligent Computing and Information Sciences, 2021.