

Influence Maximization in Complex Networks

SANJAY KUMAR



**Computer Science and Application group
DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY DELHI
October 2021**

©Indian Institute of Technology Delhi (IITD), New Delhi, 2021

Influence Maximization in Complex Networks

by

SANJAY KUMAR

Department of Mathematics

Submitted

**In fulfillment of the requirements of the degree of DOCTOR OF
PHILOSOPHY**

to the



**Indian Institute of Technology Delhi
October 2021**

Dedicated to
My Family

Certificate

This is to certify that the thesis entitled “**Influence Maximization in Complex Networks**” submitted by “**Mr. Sanjay Kumar**” to the **Indian Institute of Technology Delhi**, for the award of the Degree of **Doctor of Philosophy**, is a record of the original bona fide research work carried out by her under my supervision and guidance. The thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted in part or full to any other university or institute for the award of any degree or diploma.

New Delhi
October 2021

Prof. Bhawani Sankar Panda
Professor
Department of Mathematics
IIT Delhi

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my thesis advisor, Prof. B. S. Panda, who introduced me to this world of algorithms and graph theory. His invaluable advices and research discussions helped me to grow as a researcher. I want to thank him for his constant encouragement, understanding, and support in various ways.

I acknowledge my current employer, Department of Computer Science & Engineering, Delhi Technological University, New Delhi, India for providing all support during my research period. I would also like to thank my SRC (Student Research Committee) members: Prof. Niladri Chatterjee (Department of Mathematics), Prof. Vikas Vikram Singh (Department of Mathematics), and Prof. Sumeet Agarwal (Department of Electrical Engineering) for their useful suggestions. I am grateful to the Department of Mathematics and IIT Delhi authorities to provide me with the facilities and a positive environment to complete my Ph.D. A special thanks to my senior colleagues, Dr. Arti Pandey and Dr. Shaily Verma, who helped me in every possible way in my research. I would not want to leave this opportunity to thank all my school and college teachers who were the initial building blocks of my Ph.D journey.

I would like to use this platform to thank the wonderful people I met during my Ph.D. and who made this entire learning experience in IIT Delhi productive and fun. I want to thank my friends Juhi, Pooja, Priyamvada, Vaibhav, Navneet, Deepak for their friendly assistance and constant motivation during hard times.

Words cannot completely express my love and gratitude towards my family for their affection,

encouragement, support, and unfailing faith in me. I want to express my love and appreciation to my parents for their sacrifices, patience, unconditional support, and encouragement.

Thank you all for your encouragement!

New Delhi

October, 2021

Sanjay Kumar

Abstract

The study of complex networks provides an enormous scope of graph mining and data analytics and renders a wide range of research problems. Influence maximization is a widely studied research problem in complex network analysis because of its business values in promoting and advertising the products. Influence Maximization (IM) is the task of identifying a constant number (k) of seed users with a high spreading capability known as influential nodes such that if a piece of information originates from them, the information can reach the optimal number of vertices in the system through a diffusion cascade.

In this thesis, we extensively study the influence maximization problem in complex networks and strengthen its literature.

We propose a novel influence maximization algorithm named NCVoteRank using the notion of neighborhood coreness and voting approach. NCVoteRank identifies spreaders with suitable localities, maximum coverage, and a minimum overlap of the neighborhood.

We propose a novel algorithm named Modified Degree with Exclusion Ratio (MDER) to identify influential nodes in complex networks through the concept of modified degree centrality and mutual exclusion. We evaluate the performance of the proposed MDER algorithm using complex real-life networks of varying applications, sizes, and complexity. The extensive experimental validation using different

performance matrices concludes that the proposed MDER algorithm outperforms many popular methods in the literature.

We propose a method based on Hooke's Law of Elasticity in complex weighted networks to find the influential spreaders. The edges present in the network are modeled as springs, which are connected in series and parallel. The proposed HookeRank model finds relevant influential nodes, on the basis of their Hooke rank distances from other nodes in the network.

We introduce a novel community detection method using network embedding, k -means clustering, and gravitational search algorithm (GSA). We devise a generic framework to apply various network embedding techniques on graph datasets to uncover community structure in complex networks. Using network embedding techniques, we consider nodes as points in d -dimensional vector space such that two adjacent nodes are closer in feature space and non-adjacent nodes are away. In the proposed community detection algorithm, the outcome of the k -means clustering as one of the inputs enhances the quality of the initial populations of GSA, which ensures a globally optimal solution.

We propose a Community structure with an Integrated Features Ranking (CIFR) algorithm to find influential spreaders in the network. CIFR algorithm utilizes the community detection algorithm to find communities in the system. Then, we rank the nodes of the network based on three factors- local ranking, gateway ranking, and community rank. In general, we term these three features of nodes as integrated features. CIFR algorithm finds seed nodes that give rise to an improved performance based on various evaluation criteria and outperforms many well-known algorithms, including recently introduced local information dimensionality (LID) and gateway local rank (GLR) method.

We propose an Influence Maximization technique named IM-ELPR using Extended h -index and Label Propagation with Relationship matrix. The proposed IM-ELPR method uses extended h -index centrality to detect the seed nodes and

then uses the label propagation technique to detect communities. The proposed algorithm performs the merging of relatively smaller and related communities into a larger community for a better perspective and to maximize the modularity. The experimental results obtained exhibit improved performance of IM-ELPR over many contemporary influence maximization algorithms.

We propose a novel approach of influence maximization using transfer learning via graph-based LSTM. The proposed approach considers the influence maximization problem as a classical regression task. We implement a graph-based long short-term memory (GLSTM) model using a quasi-embedding based approach derived from three classical node centrality methods. We perform intensive experiments on several real-life networks, which reveal the commendable performance of the GLSTM based model as compared to other algorithms for influence maximization.

सार

जटिल नेटवर्क का अध्ययन ग्राफ खनन और डेटा विश्लेषण का एक विशाल दायरा प्रदान करता है और अनुसंधान समस्याओं की एक विस्तृत श्रृंखला प्रस्तुत करता है। उत्पादों को बढ़ावा देने और विज्ञापन देने में अपने व्यावसायिक मूल्यों के कारण जटिल नेटवर्क विश्लेषण में प्रभाव अधिकतमकरण व्यापक रूप से अध्ययन की गई शोध समस्या है। प्रभाव अधिकतमकरण (आईएम) एक उच्च प्रसार क्षमता वाले बीज उपयोगकर्ताओं की एक स्थिर संख्या की पहचान करने का कार्य है, जिसे प्रभावशाली नोड्स के रूप में जाना जाता है, जैसे कि यदि जानकारी का एक टुकड़ा उनसे उत्पन्न होता है, तो जानकारी सिस्टम में अधिकतम संख्या तक पहुंच सकती है।

इस थीसिस में, हम जटिल नेटवर्क में प्रभाव अधिकतमकरण समस्या का व्यापक अध्ययन करते हैं और इसके साहित्य को मजबूत करते हैं। हम पड़ोस की कोरनेस और वोटिंग दृष्टिकोण की धारणा का उपयोग करते हुए NCVoteRank नामक एक नया प्रभाव अधिकतमकरण एल्गोरिथम का प्रस्ताव करते हैं। NCVoteRank उपयुक्त इलाकों, अधिकतम कवरेज और पड़ोस के न्यूनतम ओवरलैप वाले स्प्रेडर्स की पहचान करता है।

हम संशोधित डिग्री केंद्रीयता और पारस्परिक बहिष्करण की अवधारणा के माध्यम से जटिल नेटवर्क में प्रभावशाली नोड्स की पहचान करने के लिए संशोधित डिग्री के साथ बहिष्करण अनुपात (एमडीईआर) नामक एक नया एल्गोरिथम प्रस्तावित करते हैं। हम अलग-अलग अनुप्रयोगों, आकारों और जटिलता के जटिल वास्तविक जीवन नेटवर्क का उपयोग करके प्रस्तावित MDER एल्गोरिथम के प्रदर्शन का मूल्यांकन करते हैं। विभिन्न प्रदर्शन मैट्रिक्स का उपयोग करते हुए व्यापक प्रयोगात्मक सत्यापन का निष्कर्ष है कि प्रस्तावित MDER एल्गोरिथम बेहतर प्रदर्शन करता है।

हम प्रभावशाली स्प्रेडर्स को खोजने के लिए जटिल भारित नेटवर्क में हुक के लोच के नियम पर आधारित एक विधि का प्रस्ताव करते हैं। नेटवर्क में मौजूद किनारों को स्प्रिंग्स के रूप में तैयार किया गया है, जो श्रृंखला और समानांतर में जुड़े हुए हैं। प्रस्तावित हुकरैंक मॉडल नेटवर्क में अन्य नोड्स से उनके हुक रैंक की दूरी के आधार पर प्रासंगिक प्रभावशाली नोड्स पाता है।

हम नेटवर्क एम्बेडिंग, k -साधन क्लस्टरिंग, और गुरुत्वीय खोज एल्गोरिथम (GSA) का उपयोग करके एक नवीन समुदाय पहचान पद्धति का परिचय देते हैं। हम जटिल नेटवर्क में सामुदायिक संरचना को उजागर करने के लिए ग्राफ डेटासेट पर विभिन्न नेटवर्क एम्बेडिंग तकनीकों को लागू करने के लिए एक सामान्य ढांचा तैयार करते हैं। नेटवर्क एम्बेडिंग तकनीकों का उपयोग करना, हम नोड्स को डी-डायमेंशनल वेक्टर स्पेस में बिंदुओं के रूप में मानते हैं जैसे कि दो आसन्न नोड्स फीचर स्पेस में करीब हैं और गैर-आसन्न नोड्स दूर हैं। प्रस्तावित कम्युनिटी डिटेक्शन एल्गोरिथम में, एक इनपुट के रूप में k -साधन क्लस्टरिंग का परिणाम GSA की प्रारंभिक आबादी की गुणवत्ता को बढ़ाता है, जो सामान्य रूप से एक इष्टतम समाधान सुनिश्चित करता है।

हम नेटवर्क में प्रभावशाली स्प्रेडर्स को खोजने के लिए एक एकीकृत फीचर रैंकिंग (सीआईएफआर) एल्गोरिथम के साथ एक सामुदायिक संरचना का प्रस्ताव करते हैं। सीआईएफआर एल्गोरिथम सिस्टम में समुदायों को खोजने के लिए कम्युनिटी डिटेक्शन एल्गोरिथम का उपयोग करता है। फिर, हम तीन कारकों के आधार पर नेटवर्क के नोड्स को रैंक करते हैं- स्थानीय रैंकिंग, गेटवे रैंकिंग, और समुदाय रैंक। सामान्य तौर पर, हम नोड्स की इन तीन विशेषताओं को एकीकृत सुविधाएँ कहते हैं। सीआईएफआर एल्गोरिथम बीज नोड्स ढूँढता है जो विभिन्न मूल्यांकन मानदंडों के आधार पर बेहतर प्रदर्शन को जन्म देता है और कई प्रसिद्ध एल्गोरिथम से बेहतर प्रदर्शन करता है।

हम विस्तारित एच-इंडेक्स और रिलेशनशिप मैट्रिक्स के साथ लेबल प्रचार का उपयोग करके आईएम-ईएलपीआर नामक एक प्रभाव अधिकतमकरण तकनीक का प्रस्ताव करते हैं। प्रस्तावित आईएम-ईएलपीआर विधि बीज नोड्स का पता लगाने के लिए विस्तारित एच-इंडेक्स केंद्रीयता का उपयोग करती है और फिर समुदायों का पता लगाने के लिए लेबल प्रसार तकनीक का उपयोग करती है। प्रस्तावित एल्गोरिथम अपेक्षाकृत छोटे और संबंधित समुदायों को एक बेहतर परिप्रेक्ष्य के लिए और प्रतिरूपकता को अधिकतम करने के लिए एक बड़े समुदाय में विलय करता है। प्राप्त प्रायोगिक परिणामों ने कई समकालीनों की तुलना में IM-ELPR के बेहतर प्रदर्शन को प्रदर्शित किया।

हम ग्राफ-आधारित LSTM के माध्यम से स्थानांतरण सीखने का उपयोग करके प्रभाव को अधिकतम करने के एक नए दृष्टिकोण का प्रस्ताव करते हैं। प्रस्तावित दृष्टिकोण प्रभाव को अधिकतम करने पर विचार करता है एक शास्त्रीय प्रतिगमन कार्य के रूप में समस्या। हम तीन शास्त्रीय नोड केंद्रीयता विधियों से प्राप्त अर्ध-एम्बेडिंग आधारित दृष्टिकोण का उपयोग करके एक ग्राफ-आधारित दीर्घकालिक अल्पकालिक स्मृति (जीएलएसटीएम) मॉडल लागू करते हैं। हम कई वास्तविक जीवन नेटवर्क पर गहन प्रयोग करते हैं, जो प्रभाव को अधिकतम करने के लिए अन्य एल्गोरिथम की तुलना में GLSTM आधारित मॉडल के सराहनीय प्रदर्शन को प्रकट करते हैं।

Contents

Certificate	i
Acknowledgements	iii
Abstract	v
List of Figures	xiii
List of Symbols	xix
1 Introduction	1
1.1 Introduction	1
1.2 Community Structure in Complex Networks	3
1.3 Popular existing Influence Maximization Algorithms	4
1.4 Information Diffusion	12
1.5 Performance Evaluation Metrics for Influence Maximization	14
1.6 Contribution of the thesis	16
1.7 Organization of the thesis	18
2 Influence Maximization using Voting Approach	23
2.1 Introduction	23
2.2 Proposed Work	24
2.2.1 Computational Complexity of NCVoteRank	27
2.3 Datasets and toy network simulation	28
2.4 Results and Analysis	32
2.4.1 Final infected scale ($F(t_c)$) Vs. time	33

2.4.2	Final infected scale Vs. fraction of spreaders	34
2.4.3	Kendall Tau (τ) Metric	37
2.4.4	Average Shortest Path Length between Spreaders (L_s)	38
3	Influence Maximization using MDER and HookeRank methods	41
3.1	Introduction	41
3.2	Influence Maximization using MDER Algorithm	43
3.2.1	Explanation on how our algorithm achieves mutual exclusion through a toy network	47
3.2.2	Time Complexity of MDER algorithm	49
3.2.3	Dataset used for MDER algorithm	50
3.2.4	Experimental results and analysis of MDER algorithm	52
3.3	Influence maximization using HookeRank algorithm for weighted networks	61
3.3.1	Distance Calculation	62
3.3.2	Proposed HookeRank algorithm	64
3.3.3	Datasets used in HookRank Algorithm	65
3.3.4	Results and Analysis of HookeRank Algorithm	65
3.3.5	Simulation of the proposed HookeRank algorithm on a toy network	66
3.3.6	Simulation of the proposed HookeRank algorithm on real-life networks	68
4	Uncovering Community Structure and Influence Maximization using Community structure and Integrated Feature Ranking	71
4.1	Introduction	71
4.2	Backgrounds	74
4.2.1	Network Embedding	74
4.2.2	Gravitational Search Algorithm	76

4.2.3	Evaluation Criteria for Community detection algorithm	
	78
4.3	Proposed Community Detection Algorithm	80
4.3.1	Datasets used for Community detection algorithm	85
4.3.2	Experimental results and analysis of proposed community de-	
	tection Algorithm	89
4.4	Proposed Influence Maximization using Community Structure and	
	Integetrated Features (CIFR) Algorithm	99
4.4.1	Time complexity analysis of CIFR Algorithm	107
4.4.2	Datasets used in CIFR Algorithm	108
4.4.3	Simulation of Proposed CIFR Algorithm on a Toy Network	110
4.4.4	Experimental Results and Analysis of CIFR Algorithm	111
5	Influence Maximization using Label Propagation based Community	
	Structure	121
5.1	Introduction	121
5.2	Proposed Algorithm: IM-ELPR	123
5.2.1	Seeding Phase	124
5.2.2	Label Propagation	126
5.2.3	Merging Communities by Relationship Matrix	128
5.2.4	Finding k Influential Nodes	130
5.2.5	Time Complexity	131
5.3	Datasets used	132
5.4	Experimental Results	133
6	Influence Maximization using Transfer Learning via Graph-based	
	LSTM	143
6.1	Introduction	143
6.2	Proposed work	146
6.2.1	Feature extraction	147

6.2.2	Label generation	149
6.2.3	Training the model using Graph-based LSTM	149
6.2.4	Transfer learning	151
6.2.5	Proposed model architecture	152
6.2.6	Algorithms	152
6.2.7	Time complexity analysis	157
6.3	Datasets used	159
6.4	Experimental results and analysis	161
6.4.1	Parametric analysis	161
6.4.2	Results	163
7	Conclusion and Future Research	173
7.1	Contributions	173
7.2	Future work	178
	Bibliography	181
	Bio-Data	193

List of Figures

1.1 K-Shell Centrality of various nodes	8
1.2 SIR Model	13
2.1 Toy network after the first iteration	29
2.2 Toy network after the second iteration	30
2.3 (a)-(h): Final infected scale ($F(t_c)$) Vs. time of eight networks under the same rate of infection (β) = 0.1, tuning parameter (θ) = 0.5 for all datasets and different spreaders fraction as mentioned in each figure.	35
2.4 Final infected scale ($F(t_c)$) or fraction of recovered nodes with respect to spreaders fraction under the same infection rate (β) = 0.1 and controlling parameter (θ) = 0.5 for all datasets	36
3.1 Brief overview and working of the proposed algorithm	47
3.2 Final infected scale Vs. spraeader fractions using SIR model for a) gplus, b) facebook, c) PGP, and d) ca-Hepth dataset. The result shows the comparision of results for different values of $MDER_i$.	48
3.3 Toy Network having 25 nodes. The neighbours of node number 6 upto 2^{nd} level is represented in green color. The neighbours of node number 1 upto 2^{nd} level is outlined with red color and the neighbours of node number 4 upto 2^{nd} level is shown having edges of the color blue.	49

3.4	Infection scale at different timestamps for each dataset by taking top-100 influential nodes from each method. The results are obtained over 100 independent SIR simulations with rate of infection as β_{th} .	54
3.5	Final infected scale with respect to spreader fractions using SIR model on ten different real-life networks. The results are obtained using 100 independent simulations.	56
3.6	Number of active nodes with respect to spreader fraction using IC model on ten different real networks. The results are sampled over 100 independent simulations.	57
3.7	(a)-(j): The average shortest path distance between the selected spreaders (L_s) Vs. spreader fraction for each dataset.	59
3.8	(a)-(j) The actual running time to generate the ranking list of nodes under different methods for all the datasets used.	60
3.9	Calculating the heuristic distance between two nodes based on weights of the edges, being modeled using springs and evaluating to a single spring, following Hooke's Law of Elasticity	62
3.10	Calculating the heuristic distance between two nodes based on weights of the edges, being modeled using springs and evaluating to a single spring, following Hooke's Law of Elasticity	63
3.11	Calculating the heuristic distance between two nodes based on weights of the edges, being modeled using springs and evaluating to a single spring, following Hooke's Law of Elasticity	63
3.12	A sample Weighted Network where the edges are modelled as springs and the spring constant is analogous to the weight of the edge	67
3.13	The run of a breadth-first search with respect to A as the starting node	67
3.14	The final network with respect to node A. Notice that certain indirect neighbors became direct neighbors under the action of both series and parallel connections.	68

3.15	The infection scale with respect to the percentage of spreaders using SIR model with $\beta = 0.01$:(a) US-Airports network (b) Facebook-like weighted network (c) PowerGrid network	69
3.16	The final infection scale with respect to the time using SIR model with $\beta = 0.01$:(a) US-Airports network (b) Facebook-like weighted network (c) PowerGrid network	70
4.1	Simulation of proposed Algorithm on Zachary Karate Club network with $2-d$ node2vec embedding-(a) input Zachary Karate Club dataset, (b) 2-D embedding using node2vec, (c) Low Rank Approximation, (d) Cluster Centers With k -Means, (e) Cluster Centers after GSA, and (f) Finally, obtained communities as output.	86
4.2	Community detection results on various networks (a) Karate club with two visible communities, (b) American Football with twelve visible communities, (c) political Books with three communities, (d) Word dataset with two communities, (e) LFR-1, a synthetically generated dataset, and the graph is well separated into four communities, (f) community detection on LFR-2 dataset the division among the four communities is visible.	90
4.3	NMI results on different datasets (a) Karate club, (b)American Football, (c)Polblogs , (d)political Books, (e) Word dataset (f) LFR-1, (g)LFR-2, and (h)LFR-3.	91
4.4	Modularity results on different datasets (a) Karate club, (b)American Football, (c)Polblogs , (d)political Books, (e) Word dataset (f) LFR-1, (g)LFR-2, and (h)LFR-3.	95
4.5	Precesion value of different methods in each dataset	98
4.6	Recall value of different methods in each dataset	99

4.7	Toy network to explain the working of the proposed Algorithm: (a)Input graph (b) Graph after finding communities (c) Computing local ranking of nodes in each community and then doing normalization (d) Computing gateway rank of each node, (e) Computing community ranking based on number of outgoing connection to other communities, and (f) Toy network after executing the proposed algorithm where top five spreaders are colored.	109
4.8	(a)-(h): Infection scale ($F(t)$) Vs. time using the same rate of infection (β) = 0.08 for all networks. The number of initial spreaders is taken as 2% for PGP, Cond-Mat, LFR-0.2 datasets, and 5% for Hep-Th, Gr-Qc, Facebook, Gnutella08, and Guntella09 datasets. The results are obtained over 100 independent simulations of the SIR model.	112
4.9	(a)-(h): (a)-(h): Number of infected nodes at the end of spreading process with respect to different number of spreaders using the SIR model with infection rate (β) = 0.08 for all the datasets. The results are obtained from 100 independent simulations of SIR model.	115
4.10	(a)-(h): Number of infected nodes at the end of spreading process with respect to different number of spreaders using the IC model with P_{uv} = 0.12 for all the datasets. The results are obtained from 100 independent simulations of IC model.	116
4.11	(a)-(h): Average shortest distance between spreaders with respect to spreaders fraction using the SIR model and the same infection rate (β) = 0.08 for all datasets. The results are obtained from 100 independent simulations of SIR model.	119
5.1	Toy network for comparison of degree, h -index and extended h -index centrality	125

5.2	(a)-(g) Final infected scale vs. spreader fraction values for different datasets- Football, Email, Jazz, C. Elegans, Facebook, Wiki Vote, Tech Router, and Gnutella P2P 08. The results are generated over 100 independent simulations of the IC model with given activation probability (p) value in the result of each dataset.	139
5.3	(a)-(g) Infected Scale $F(t)$ vs. Time values for different datasets - Football, Email, Jazz, C. Elegans, Facebook Pages, Wiki Vote, Tech Routers, and Gnutella P2P. The results are generated over 100 independent simulations of the IC model with given activation probability (p) value in the result of each dataset	140
5.4	(a)-(h) Average shortest path length (L_s) Vs. Spreader Fraction values for different datasets - Football, Email, Jazz, C. Elegans, Facebook Pages, Wiki Vote, Tech Routers, and Gnutella P2P 08.	141
6.1	Graph-based Long Short Term Memory architecture	151
6.2	Model architecture of the proposed work for influence maximisation using graph-based transfer learning	153
6.3	(a)-(h) Final infected scale vs seed set size for test networks. The infection probability β is set to 0.1 under the SIR model.	162
6.4	(a)-(h) Final infected scale vs. seed set size value obtained by various methods on different datasets. The results are obtained using 100 independent simulations of the SIR model with infection probability (β) as 0.1.	165
6.5	(a)-(h) Final infected scale Vs. infection probability values obtained by different methods for various chosen algorithms. The results are obtained using 100 independent simulations of the SIR model with infection rate (β) as 0.1. The number of initial seed nodes is taken as two percent of the total nodes in each network.	166

6.6	(a)-(h) Kendall's Tau value vs infection probability for different chosen datasets obtained for different algorithms.	170
6.7	(a)-(h) The absolute running time (in ms.) for finding the top 100 influential nodes using various algorithms of IM on different datasets.	172

List of Symbols

Symbol	Meaning
\mathbb{N}	Set of <i>natural numbers</i>
$\forall x$	For <i>all</i> x
$x \in X$	x is a <i>member</i> of X
$A \subseteq X$	A is a <i>subset</i> of X
$ X $	The <i>cardinality</i> of a set X
$A \cap B$	The <i>intersection</i> of A and B
$A \cup B$	The <i>union</i> of A and B
$A \setminus B$	The <i>set difference</i> of A and B
$A \times B$	The <i>Cartesian product</i> of A and B
$E(G)$	The set of <i>edges</i> of G
$V(G)$	The set of <i>vertices</i> of G
$d_G(v)$	The <i>degree</i> of the vertex v in G
Δ or $\Delta(G)$	The <i>maximum degree</i> in G
$DC(v)$	The <i>degree centrality</i> of the vertex v in G
$CC(v)$	The <i>closeness centrality</i> of the vertex v in G
$BC(v)$	The <i>betweenness centrality</i> of the vertex v in G
k -shell(v)	The <i>k-shell centrality</i> of the vertex v in G

δ or $\delta(G)$	The <i>minimum degree</i> in G
$N_G(v)$	<i>neighborhood</i> of v in G
$\sigma_{G,M}(S)$	<i>expected influence spread</i> or numbers of users influenced by the users in set S where G is the complex network and M is information diffusion model
$N_G[v]$	<i>closed neighborhood</i> of v in G
P	The class of <i>deterministic</i> polynomial-time solvable problems
NP	The class of <i>nondeterministic</i> polynomial-time solvable problems