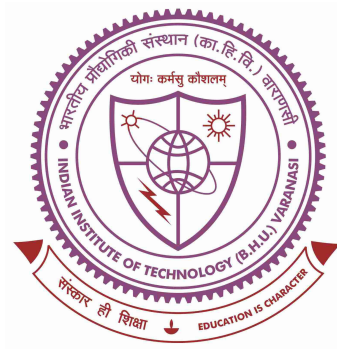


Link Prediction in Dynamic Networks using Feature and Quantum based Machine Learning Techniques



Thesis submitted in partial fulfillment

for the Award of Degree

DOCTOR OF PHILOSOPHY

By

MUKESH KUMAR

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY

(BANARAS HINDU UNIVERSITY),

VARANASI-221 005

Roll No: 19071009

2023

Certificate

It is certified that the work contained in this thesis entitled “Link Prediction in Dynamic Networks using Feature and Quantum based Machine Learning Techniques” by “MUKESH KUMAR” has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

It is further certified that the student has fulfilled all the requirements of Comprehensive Examination, Candidacy and SOTA for the award of Ph.D. Degree.

Bhaskar
30/01/2023

Dr. Bhaskar Biswas

सह आचार्य/Associate Professor

तंत्रज्ञान विज्ञान एवं अभियंत्रिकी विभाग/Department of Computer Sc. & Engg

भारतीय प्रौद्योगिकी संस्थान (बनारस हिन्दू विश्वविद्यालय)/Indian Institute of Technology

(बनारस हिन्दू विश्वविद्यालय)/(Banaras Hindu University)

Department of Computer Science and Engineering

वाराणसी / Varanasi-221 005

Indian Institute of Technology (Banaras Hindu University),

Varanasi-221 005

COPYRIGHT TRANSFER CERTIFICATE

Title of the Thesis : **Link Prediction in Dynamic Networks using Feature and Quantum based Machine Learning Techniques**

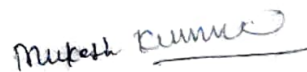
Name of the Student : **MUKESH KUMAR**

Copyright Transfer

The undersigned hereby assigns to the Indian Institute of Technology (Banaras Hindu University) Varanasi all rights under copyright that may exist in and for the above thesis submitted for the award of the "DOCTOR OF PHILOSOPHY".

Date : 30/01/2023.

Place : Varanasi



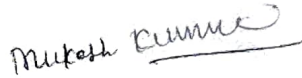
Signature of the Student

(MUKESH KUMAR)

Note: However, the author may reproduce or authorize others to reproduce material extracted verbatim from the thesis or derivative of the thesis for author's personal use provided that the source and the Institute's copyright notice are indicated.

DECLARATION BY THE CANDIDATE

I, MUKESH KUMAR, certify that the work embodied in this thesis is my own bonafide work and carried out by me under the supervision of Dr. Bhaskar Biswas from July 29, 2019 to January 29, 2023, at the Department of Computer Science and Engineering, Indian Institute of Technology (BHU), Varanasi. The matter embodied in this thesis has not been submitted for the award of any other degree/diploma. I declare that I have faithfully acknowledged and given credits to the research workers wherever their works have been cited in my work in this thesis. I further declare that I have not willfully copied any other's work, paragraphs, text, data, results, etc., reported in journals, books, magazines, reports dissertations, theses, etc., or available at websites and have not included them in this thesis and have not cited as my own work.



Date :


Signature of the Student

Place : Varanasi

(MUKESH KUMAR)

CERTIFICATE BY THE SUPERVISOR

It is certified that the above statement made by the student is correct to the best of my/our knowledge.

 30/01/2023

Signature of Supervisor

(Dr. Bhaskar Biswas)

पर्यवेक्षक/Supervisor

संगणक विज्ञान एवं अभियांत्रिकी विभाग

Department of Computer Sc. & Engg

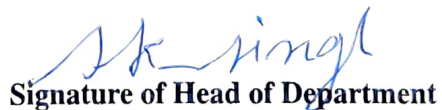
भारतीय प्रौद्योगिकी संस्थान

Indian Institute of Technology

(वाराणसी हिन्दू विश्वविद्यालय)

(Banaras Hindu University)

वाराणसी-221005



Signature of Head of Department

आचार्य व विभागाध्यक्ष

Professor & Head

संगणक विज्ञान एवं अभियांत्रिकी विभाग

Department of Computer Sc. & Engg

भारतीय प्रौद्योगिकी संस्थान

Indian Institute of Technology

(वाराणसी हिन्दू विश्वविद्यालय)

(Banaras Hindu University)

वाराणसी-221005

Acknowledgements

This thesis is the end of my journey in obtaining my Ph.D. I have not traveled in a vacuum in this journey. This thesis has been kept on track and been seen through to completion with the support and encouragement of numerous people including my well wishers, my friends, colleagues and various institutions. At the end of my thesis, I would like to thank all those people who made this thesis possible and an unforgettable experience for me. At the end of my thesis, it is a pleasant task to express my thanks to all those who contributed in many ways to the success of this study and made it an unforgettable experience for me.

Firstly, I would like to express my sincere gratitude to my advisor **Dr. Bhaskar Biswas** for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

I specially thank the members of my Research Progress Evaluation Committee, **Dr. Sanjay kumar Singh** and **Dr. Rajeev** for their invaluable suggestions regarding the thesis, their insightful comments and encouragement, but also for the hard question which incited me to widen my research from various perspectives.

I would like to express my sincere thanks for **Prof. Sanjay Kumar Singh**, Head, Computer Science and Engineering Department, for his kindness and valuable support in carrying out the research. I express my sincere thanks to the faculty members, **Prof. K. K. Shukla**, **Prof. A. K. Tripathi**, **Dr. Ravi Shankar Singh**, **Dr. Vinayak Srivastava**, **Dr. R. N. Choudary**, **Dr. H. P. Gupta**, **Dr. T. Dutta**, **Dr. A. K. Singh** and staffs of the department.

I am also thankful to all my friends for the fun moments in my PhD student life. I wish to express my deep sense of thanks to my senior **Dr. Ajay Kumar** and **Mr. Shivansh Mishra** for being a great friends and the best advisers I could ever have. Their advice, encouragement and critics were always the source of inspiration. This dissertation would not have been possible without their invaluable suggestions and persistent help.

Most importantly, my deepest gratitude is for my family for their constant support, inspiration, guidance, and sacrifices. My parents were constant source of motivation and

inspiration. Their affection and guidance was instrumental in me choosing Engineering and eventually continuing on to my Ph.D. I would like to thank my mother, who was also my childhood teacher and is always there to stands by my side.

Last but not the least, I would like to special thanks is to my wife **Miss Anjali Kumari**, my daughters **Srishti Yadav and Trisha** and my son **Shivansh Yadav** for supporting me unconditionally throughout my study. I sincerely thank all of them who contributed in for helping me to see the light at the end of every scary tunnel during my Ph.D.

- MUKESH KUMAR

PREFACE

Link prediction is a task that involves network structure evaluation and identifying missing and future links in social networks. In dynamic networks, each edge is associated with a temporal component, and link prediction in such networks is the task of identifying edge existence probability based on previously observed network dynamics. This research is crucial to comprehending network evolution and its effects on individual node behavior with respect to time. Link prediction algorithms can be applied to extract missing data, identify erroneous interactions, evaluate the factors governing network evolution, and more. Node pair similarity-based link prediction methods use different structural or topological information to calculate possible edges' chance. In the snapshot-based paradigm for link prediction on dynamic graphs, feature sets created from combining different similarity-based methods are used to track the temporal changes in edges. These features are then used to make predictions for non-existing edges. Machine learning algorithms are used on these feature sets to build prediction models. Compared to the individual similarity-based approaches, machine learning techniques have been used to improve prediction accuracy as they are more helpful in modeling changes in patterns over time. This thesis aims to enhance link prediction on dynamic networks using different feature sets and machine learning algorithms.

The following four fundamental aspects of link prediction are explored in this thesis since they are used to build the feature sets for machine learning classifiers:

- 1) incorporating local, global, and quasi-local similarity indices for link prediction, 2) merging snapshot-based features with snapshot-impartial path-based and cost-based features, 3) feature selection to identify best feature sets based on similarity indices, 4) exploring community detection enhanced link prediction features, and 5) studying the effect of quantum kernel-inspired transformation on link prediction in dynamic networks.

One of the well-known categories of link prediction methods is the similarity-based method, which uses a node pair similarity score for edge estimation. The three widely used categories of similarity-based indices are Local (L), Global (G), and Quasi-local (Q) indices. All three types of these methods can be combined to create a rich feature set

that would be more useful for link prediction than its components individually. The proposed LGQ framework explores the relative effect of combining these categories of link prediction features on different machine learning algorithms.

Further exploring the role of snapshot-impartial paths, Path Weight Aggregation Feature (PWAFF) for Link Prediction in Dynamic Networks is proposed. Different topological aspects of the networks (Local, Global, and Quasi-local), as well as Clustering Coefficient based features, are considered for feature generation, in addition to the suggested Path Weight-Based Aggregation Feature (PWAFF). Some of these features also take advantage of longer paths between nodes for a more thorough similarity estimation.

A feature-based solution that considers individual snapshots and the overall network throughout the full-time span to answer the link prediction problem is explored. A novel feature called Cost-based feature for link prediction (CFLP) for estimating edge behavior throughout the entire network is proposed, which uses a reward and penalty structure to summarize node activity across the whole network (snapshot-impartial), is presented. The feature set also uses similarity indices, classified into four major categories: local similarity, global similarity, quasi-local similarity, and clustering coefficient-based similarity, to measure edge activity change in individual snapshots. Different feature selection methods are used to correctly quantify the relative effect of features among themselves and the overall link prediction problem.

The highly connected groups of nodes within the network are called communities, which have an important role in understanding and uncovering various functional properties of the system. A community information-based feature estimation and link prediction (COMMLP) method applied to dynamic graphs in a per snapshot feature estimation-based setting is proposed to take advantage of the group behavior of nodes. First, a link prediction framework is presented to predict missing links using parameterized influence regions of nodes and their contribution to community partitions. Then, a unique feature set is generated using local, global, and quasi-local similarity-based and community information-based features. This feature set is further optimized using scoring-based feature selection methods to select only the most relevant features.

Finally, the solution to the issue of link prediction in dynamic networks using supervised learning and the Projected Quantum Kernel (PQK) is addressed. It is conceivable to

expect that quantum computers may outperform classical computers in machine learning tasks since quantum systems display aberrant behavior that conventional systems are thought to be incapable of creating. The QML approaches use the huge dimensionality of quantum Hilbert space to get an optimized solution by modeling the feature space of a classification problem with a quantum state. The projected quantum kernel-based link prediction (PQKLP) approach is presented to address the link prediction problem using both local and global information. Using PQK, features are transformed from the popularly used snapshot-based feature set form into quantum space such that the effectiveness of machine learning-based classification can be improved.

Contents

Certificate	iii
Declaration by the Candidate	v
Copyright Transfer Certificate	vii
Preface	xi
Acknowledgements	xv
Contents	xvii
List of Figures	xxiii
List of Tables	xxvii
Abbreviations	xxxii
Symbols	xxxiii
1 Introduction	1
1.1 Link prediction	3
1.2 Challenges in Link Prediction	4
1.3 Motivation of the thesis	5
1.4 Contribution of the thesis	8
1.4.1 Feature fusion based link prediction in dynamic network	8
1.4.2 Path Weight Aggregation Feature for link prediction	9
1.4.3 A new cost based feature for link prediction	10
1.4.4 Community Enhanced Link Prediction	10
1.4.5 Projected Quantum Kernel based Link prediction	11

1.5	Organization of the thesis	11
2	Literature Review	13
2.1	Similarity-based methods	13
2.1.1	Local similarity based indices	15
2.1.2	Global similarity indices	17
2.1.3	Quasi-local indices	18
2.1.4	Clustering based Features	21
2.2	Dynamic Networks	23
2.2.1	Dynamic Datasets	25
2.2.2	Link Prediction Framework in dynamic networks	27
2.3	Machine learning classifiers in link prediction	30
2.3.1	Machine learning classifier	31
2.3.2	Ensemble learning	33
2.3.3	Performance evaluation metrics	34
3	Features Fusion based Link Prediction in Dynamic Networks	37
3.1	Introduction	38
3.2	Proposed work	41
3.2.1	Proposed Feature generation model Algorithm	42
3.3	Result Analysis	43
3.3.1	Performance of LGQ model and its variations with Neural Network (NN)	43
3.3.2	Performance of LGQ and its variations with XGBoost	47
3.3.3	Comparison of LGQ model with state-of-the-art methods	48
3.4	Conclusion	51
4	Path Weight Aggregation Feature for Link Prediction in Dynamic Networks (PWAFF)	53
4.1	Introduction	54
4.2	Proposed work	57
4.2.1	Path Weight Aggregation Feature (<i>PWAFF</i>) for link prediction	57
4.2.2	Analysis for selecting \sqrt{N}	59
4.2.3	Proposed PWAFF Feature Generation Algorithm	59
4.3	Result Analysis	63
4.3.1	Performance of PWAFF model and its machine learning variations-XGBoost (XGB)	67
4.3.2	Performance of PWAFF model and its machine learning variations-Random Forest Classifier (RFC)	68
4.3.3	Performance of PWAFF model and its machine learning variations-Linear Discriminant Analysis (LDA)	68

4.3.4	Performance comparison of PWAF model machine learning variations	69
4.3.5	Comparison of PWAF variations with state-of-the-art methods	71
4.4	Conclusion	73
5	A New Cost Based Feature for Link Prediction in Dynamic Networks (CFLP)	75
5.1	Introduction	76
5.2	Proposed Method	78
5.2.1	A New Cost Based Feature for Link Prediction(CFLP) algorithm	78
5.2.2	Algorithm for training and testing of CFLP model	79
5.2.3	Feature score used in CFLP	83
5.3	Result Analysis	85
5.3.1	Performance comparison of CFLP model with different similarity methods on Neural Network (NN)	85
5.3.2	Performance comparison of CFLP model with different similarity methods on Logistic Regression (LR)	86
5.3.3	Performance comparison of CFLP model with different similarity methods with Extreme Gradient boosting XGBoost (XGB)	89
5.3.4	Performance comparison of CFLP model with different similarity methods with Random Forest Classifier (RFC)	91
5.3.5	Performance comparison of CFLP model with different similarity methods Linear Discriminant Analysis (LDA)	91
5.3.6	Comparison of CFLP with state-of-the-art methods after optimization	94
5.4	Conclusion	97
6	Community Enhanced Link Prediction in Dynamic Networks	99
6.1	Introduction	99
6.1.1	Community Detection	102
6.1.2	Community-based Feature Generation Methods	103
6.1.3	Classification Models	106
6.2	Proposed work	108
6.2.1	Topological Feature Generation	109
6.2.2	Community-based Feature Generation	109
6.2.3	Feature Set Engineering and Link Prediction	112
6.2.4	Feature Reduction	113
6.2.5	Algorithm Description	114
6.2.6	Demonstration with Example	116
6.3	Result Analysis	118
6.3.1	Performance comparison of <i>COMMLP – FULL</i> with individual feature based methods using Neural Network model	119

6.3.2	Performance comparison of <i>COMMLP – FULL</i> with individual feature based methods using XGBoost model	121
6.3.3	Performance comparison of <i>COMMLP – FULL</i> with individual feature based methods using Linear Discriminant Analysis model	124
6.3.4	Performance comparison of <i>COMMLP – FULL</i> with individual feature based methods using Random Forest Classifier model	124
6.3.5	Comparison of Individual and Community Information based Link Prediction Features	126
6.3.6	Comparison of <i>COMMLP-DYN</i> with State-of-the-Art Methods after Optimization	129
6.4	Conclusion	132
7	Projected Quantum Kernel based Link Prediction in Dynamic Networks (PQKLP)	133
7.1	Introduction	134
7.1.1	Quantum Computation	137
7.1.2	Projected Quantum Kernels (PQK)	139
7.2	Proposed Method	145
7.2.1	Algorithms used in the proposed framework	145
7.2.2	Feature generation algorithm for proposed PQKLP approach	148
7.3	Result Analysis	149
7.3.1	Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Neural Network (NN)	149
7.3.2	Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Logistic Regression (LR)	151
7.3.3	Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Linear Discriminant Analysis (LDA)	153
7.3.4	Performance comparison and analysis of PQKLP model with its component individual link prediction methods using XGBoost (XGB)	153
7.3.5	Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Random Forest Classifier (RFC)	154
7.3.6	Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Gaussian Naive Baiyes classifier (GNB)	156
7.3.7	Performance comparison and analysis of classical machine learning model with proposed PQKLP-based machine learning model with same feature set	157

7.3.8	Performance comparison and analysis of Projected Quantum Kernel based Link Prediction (PQKLP) model with various state-of-the art algorithms	159
7.4	Conclusion	160
8	Conclusion and future directions	163
8.1	Summary of Contributions	163
8.2	Scope for Future Work	166
	Bibliography	169
A	List of Publications	193

List of Figures

1.1	The Link Prediction (LP) finds missing links (i.e., 1 – 2, 2 – 7, and 5 – 6) in this observed network.	4
2.1	Taxonomy of Link Prediction Approaches [1]	14
2.2	Structure of a dynamic network with different snapshots	25
2.3	Flow of link prediction framework in dynamic networks	29
3.1	Dynamic Networks with different snapshots. In the snapshot-based paradigm for link prediction on dynamic graphs, feature sets are used to track the behavioral changes in edges. These features are then used to make predictions for non-existing edges.	39
3.2	Comparison of LGQ approach with state-of-the-art	50
4.1	Path-based approaches to link prediction [2]	55
4.2	Graph of $F(\frac{W}{\sqrt{N}})$ Vs $\frac{W}{\sqrt{N}}$ graph	58
4.3	Analysis for selecting \sqrt{N}	58
4.4	Representations of link prediction model using machine learning techniques in dynamic networks	60
4.5	PWAF model overview	61

4.6	Performance Comparison among PWAF machine-learning variations	70
4.6	Performance Comparison among PWAF machine-learning variations (contd..)	71
5.1	Proposed <i>CFLP</i> Model Overview	81
6.1	Structure of snapshots of a dynamic network at different time intervals (green nodes are nodes common with previous snapshot and red nodes are the nodes being added in current snapshot)	104
6.2	The Proposed Model Framework. The Proposed Model Framework. First, 4 dynamic network snapshots T_1, T_2, T_3, T_4 are used to generate 12 topological features based on local, global, and quasi-local information. Also, 8 community-based features are generated based on the three-degree theory. Then we have utilized these 20 features (12 topological and 8 community-specific) for training machine learning models NN, XGB, LDA, and RFC for the COMMLP-FULL prediction model. We have also presented another version of the proposed solution <i>COMMLP – DYN</i> , by applying feature reduction methods TREECL, KBREG & KBMIR.	107
6.3	Example of Common Influence Region of nodes $X&Y$ (represented as orange color nodes 1,3,5) when influence is assumed to be spread to 2-hop region from the central node - $CIR(X, Y)$	111
6.4	Example Graph for Feature Calculation	116
6.5	ROC curve based comparison of <i>COMMLP</i> (<i>COMMLP – DYN</i>) with other state-of-the-art feature based methods on Random Forest Classifier-based machine learning prediction	127
7.1	Temporal Networks	137
7.2	Projected Quantum Kernel [3]	140

7.3 Projected Quantum Kernel based Link Prediction Model Overview 146

List of Tables

2.1	Comparison of similarity-based approaches [1]	23
2.2	Link Prediction in Dynamic Networks	25
2.3	Dataset Information along with corresponding snapshot information and time intervals (AVG DEG - Average Degree, AVG SP - Average Shortest Path Length, CLUSTER - Average Clustering Coefficient, HETERO - Heterogeneity, ASSOC - Associativity)	28
3.1	Performance of LGQ model and its variations with Neural Network (NN)	46
3.2	performance of LGQ and its variations with XGBoost (XGB)	48
3.3	Comparison of LGQ model and its variation with state-of-the-art methods	49
4.1	Performance of PWAF model and its machine learning variations- XGBoost (XGB)	64
4.2	Performance of PWAF model and its machine learning variations- Random Forest Classifier (RFC)	65
4.3	Performance of PWAF model and its machine learning variations - Linear Discriminant Analysis (LDA)	66
4.4	Performance comparison of PWAF machine learning variation with state-of-the-art methods	72

5.1	Feature Score of all similarity measures using Mutual Information Classifier (MIC) and F-Regression method (FREG). In this table R-EMAIL, COLLEGE, LKML and MATH are shortforms for RADOSLAW-EMAIL, COLLEGEMSG, LKML-REPLY and MATHOVERFLOW respectively.)	83
5.2	Performance comparison and analysis of <i>CFLP</i> model with various individual similarity indices with Neural Network (NN)	87
5.3	Performance comparison and analysis of <i>CFLP</i> model with various individual similarity indices with Logistic Regression (<i>LR</i>)	88
5.4	Performance comparison and analysis of <i>CFLP</i> model with various individual similarity indices with XGBoost (<i>XGB</i>)	90
5.5	Performance comparison and analysis of <i>CFLP</i> model with various individual similarity indices with Random Forest Classifier (<i>RFC</i>)	92
5.6	Performance comparison and analysis of <i>CFLP</i> model with various individual similarity indices with Linear Discriminant Analysis (<i>LDA</i>)	93
5.7	Comparison of <i>CFLP</i> machine learning variation with state-of-the-art methods after optimization	95
6.1	Scoring of different link prediction methods in comparison to our proposed feature in example graph (<i>CIR(X,Y)</i> is Common Influence Region between <i>X&Y</i> , <i>CL – COM(X)&CL – COM(Y)</i> are community labels of <i>X&Y</i> and <i>COMMLP – DYN</i> is the feature calculated using this information)	118
6.2	Performance comparison of <i>COMMLP – FULL</i> with individual feature based link prediction algorithms using Neural Network model on five datasets, three <i>Ratio</i> values and three performance metrics	120
6.3	Performance comparison of <i>COMMLP – FULL</i> with individual feature based link prediction algorithms using XGBoost model on five datasets, three <i>Ratio</i> values and three performance metrics	122

6.4	Performance comparison of <i>COMMLP – FULL</i> with individual feature based link prediction algorithms using Linear Discriminant Analysis model on five datasets, three <i>Ratio</i> values and three performance metrics .	123
6.5	Performance comparison of <i>COMMLP – FULL</i> with individual feature based link prediction algorithms using Random Forest Classifier model on five datasets, three <i>Ratio</i> values and three performance metrics	125
6.6	Features Scores of all individual features using TREECL, KBMIR and KBREG methods across all datasets	128
6.7	Comparison of <i>COMMLP (COMMLP – DYN)</i> with state-of-the-art methods after optimization using truncated feature set on six datasets, 3 <i>Ratio</i> values and three performance metrics	130
7.1	Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Neural Network (NN)	150
7.2	Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Logistic Regression (LR)	151
7.3	Performance comparison and analysis of <i>PQKLP</i> model with its component individual link prediction methods using Linear Discriminant Analysis (<i>LDA</i>)	152
7.4	Performance comparison and analysis of PQKLP model with its component individual link prediction methods using XGBoost (XGB) . . .	154
7.5	Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Random Forest Classifier (RFC)	155
7.6	Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Gaussian Naive Bayes (GNB)	156

7.7	Performance Comparison of link prediction in different machine learning models with proposed projected Quantum machine learning model	157
7.8	Performance comparison and analysis of Projected Quantum Kernel based Link Prediction (PQKLP) model with various state-of-the art algorithms .	160

Abbreviations

L	Local Similarity Indices
G	Global Similarity Indices
Q	Quasi-local Similarity Indices
LG	Local and Global Similarity Indices
LQ	Local and Quasi-Local Similarity Indices
GQ	Global and Quasi-Local Similarity Indices
LGQ	Local, Global and Quasi-Local Similarity Indices
NN	Neural Network
LR	Logistic Regression
XGB	XGBoost
LDA	Linear Discriminant Analysis
RFC	Random Forest Classifier
GNB	Gaussian Naive Bays
CCLP2	Level-2 Node Clustering Coefficient-based Link Prediction
AUROC	Area Under the Receiver Operating Characteristic Curve
AP	Average Preccision
AUPR	Area Under the Precision-Recall Curve
L3	Path of length 3
CCLP	Clustering Coefficient-based Link Prediction
PWAF	Path Weight Aggregation Feature
CFLP	Cost Based Feature for Link Prediction
PQKLP	Projected Quantum Kernel based Link Prediction

Symbols

$G(V, E)$	A social network with vertex set V and edge set E
A	Adjacency matrix of a network
n	The number of nodes in the network ($ V $)
m	The number of Edges in the network ($ E $)
$\Gamma(z)$	The neighbors set of node z
k_z	Degree of the node z
$t(z)$	Number of triangle passing through the node z
λ_1	Maximum eigen value of a matrix
β	damping factor
$\langle K \rangle$	Average degree of the a network
$\langle D \rangle$	Average path length of a network
$\langle C \rangle$	Average clustering coefficient of a network
r	Coefficient of assortativity of a network
H	Degree of heterogeneity of a network
ρ	Network density
D	Diagonal matrix
L	Laplacian matrix
$C(z)$	Clustering coefficient of node z
$CC(z)$	Level-2 clustering coefficient a node z
$S(x, y)$	Similarity score between the node x and the node y