New Generation Computing
https://doi.org/10.1007/s00354-019-00065-z

Temporal Link Prediction: A Survey

Aswathy Divakaran[1] · **Anuraj Mohan**[1]

Received: 17 November 2018 / Accepted: 31 July 2019

Abstract
The evolutionary behavior of temporal networks has gained the attention of researchers with its ubiquitous applications in a variety of real-world scenarios. Learning evolutionary behavior of networks is directly related to link prediction problem, as the addition or removal of new links or edges over time leads to the network evolution. With the rise of large-scale temporal networks such as social networks, temporal link prediction has become an interesting field of study. In this work, we provide a detailed survey of various researches carried out in the direction of temporal link prediction. We build a taxonomy of temporal link prediction methods based on various approaches used and discuss the works which come under each category. Further, we present the challenges and directions for future works.

Keywords Dynamic networks · Temporal networks · Link prediction

Introduction

Dynamic or temporal networks [15, 38, 46] are those networks where entities and relationships appear and disappear over time. Entities are represented by nodes and relationships between them are represented by links. Each link carries information on the time when it is active, along with other possible characteristics. Almost all the real-world complex phenomena can be modeled as dynamic networks. Such networks are evolving in nature since they change as a function of time. For instance, social networks, communication networks, biological networks, etc., have an underlying structure of evolving networks where nodes and edges are added and removed over time. The network structure which describes how a graph is wired helps us to predict and understand the behavior of such systems. In many cases, the edges are

✉ Aswathy Divakaran
aswathydiv36@gmail.com

Anuraj Mohan
anurajmohan@gmail.com

[1] Department of Computer Science and Engineering, NSS College of Engineering, Palakkad, India

Published online: 13 August 2019

not continuously active, but are relatively short and instantaneous. Figure 1 shows three common representations for time-varying networks which help to capture the dynamics of the network:

1. Snapshots: the network is represented as a series of static networks, one for each time step.
2. Contact sequences: if the duration of interaction is negligible, the network is represented as a triplet (i, j, t), where i and j are entities and t is the time of interaction.
3. Interval graphs: if the duration of interaction is not negligible, $T(e) = \{(t_1, t'_1), \ldots, (t_n, t'_n)\}$ over which the edge e is active.

The link structure of such networks enable to find generic solutions to various problems. Link prediction [54] is one of the central network mining tasks which infers the existence of new links or missing links in a network. In static networks, link prediction is the task of predicting hidden or missing links within a static network taken at a specific period. Moreover, the goal of link prediction in dynamic networks or otherwise temporal link prediction is to predict the links in a network that would appear in its next state of period. The time-varying nature of dynamic networks makes temporal link prediction a challenging task. Furthermore, learning the evolutionary patterns in dynamic networks is directly related to the link prediction problem. Inspired from various link prediction algorithms,

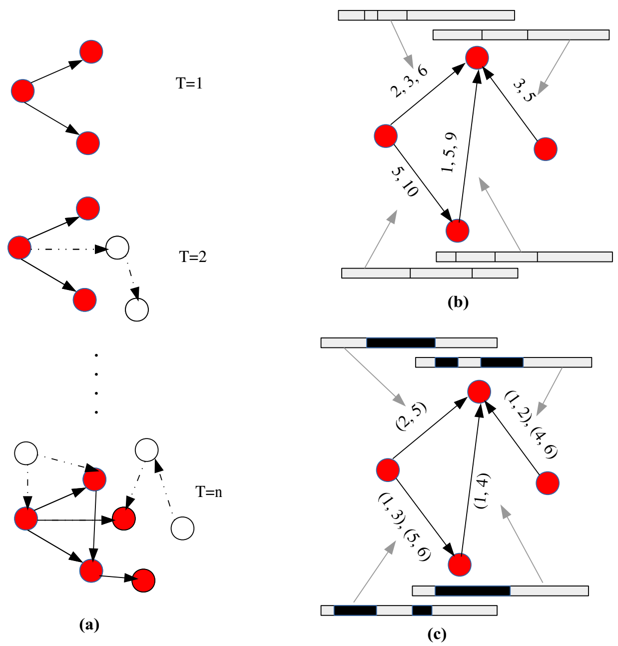

Fig. 1 Representations of dynamic network: **a** snapshots, **b** contact sequences, and **c** interval graphs

some state-of-the-art works cover the relationship between the evolution of the network and link prediction problem [97, 109]. The complex dynamic structure and non-linear varying temporal patterns make the study of dynamic networks more complicated.

With the rapid increase in real-world networks such as communication networks, co-author networks, citation networks, biological networks and email networks, temporal link prediction finds application in a variety of scenarios to analyze and solve interesting problems. Recommending new products in eBay and Amazon, friend suggestions in online social networks are some of the obvious examples. A representative survey on recommender systems [58] discussed the relationship between link prediction and personalized recommendation. In biological networks, predicting the interactions between molecules at a specific time stamp can help us better understand the temporal interaction between them. This can provide useful temporal information that indicate the stage of a specific disease such as cancer. Therefore, temporal link prediction plays an important role in disease prediction task. In addition, this task can be used to predict the academic collaborations in co-authorship and citation networks. Furthermore, temporal link prediction in terrorist communication networks help us to predict and capture the most important information related to the issue of national security.

Real-world networks exhibit different scales of dynamics ranging from small timescales (fine-grained dynamics) to the evolution of network properties over long periods (coarse-grained dynamics). Depending upon the applications, different types of granularity are more appropriate to capture meaningful behavior. At one extreme, each interval could correspond to discrete time units (e.g., day or week). In such cases, each footprint of the time-varying graph corresponds to an instant of snapshot of the network and the whole sequence becomes equivalent to the evolving graph model. But this can be seen as very coarse and causes loss of information resulting in poor predictive performance and misleading results. On the other hand, important temporal dependencies of the network can be captured at the finest granularity (e.g., at timescales of seconds or milliseconds). In such cases, contact sequence or interval graph representations are more fruitful. Although temporal information in networks are important to accurately model, predict and understand the behavior of the networks, the majority of the existing algorithms for temporal link prediction ignore such information.

Designing algorithms for temporal link prediction is non-trivial because of several reasons. First, complicated features in temporal network are difficult to characterize and extract. Complexity of the temporal networks poses difficulty in designing efficient algorithms. Furthermore, the majority of the real-world networks are sparse and sensitive to noise. Although temporal link prediction is a challenging task, an extensive amount of researchers have been devoted to solve this problem. However, the majority of the existing algorithms perform the task by taking a static network at a specific moment of time. This direction is not an efficient way for temporal link prediction, since it does not capture the dynamics of the networks. In this paper, we study various algorithms for temporal link prediction and the prevailing challenges. The contributions of this work are:

1. We build a taxonomy of temporal link prediction methods based on the various techniques used.
2. The works under temporal link prediction are categorized as techniques based on matrix factorization, probabilistic models, spectral clustering, time series, deep learning and other techniques.
3. We provide a detailed survey of works that come under each category.
4. We provide a comparison of these methods and the interesting directions for future works.

To the best of our knowledge, this is the first work that provides a comprehensive survey on various temporal link prediction techniques.

Background

In this section, we familiarize various algorithms for static link prediction and introduce a formal definition for the temporal link prediction problem.

Link Prediction in Static Networks

Although temporal link prediction is a challenging task, the majority of the algorithms perform the task by taking static networks for specific period and thereby ignore the temporal aspect of the networks. In static networks, link prediction is the task of predicting missing links in the network for a specific moment of time to capture the complete picture of the network.

Some of the works in static link prediction deployed structural similarity-based algorithms. Various local similarity indices [45, 68, 115], global similarity indices [43] and quasi-local similarity indices [56, 57] were deployed for static link prediction, where the similarity measure computed between each pair of nodes determine the existence of links between them. Several algorithms were proposed by extending some of the local and global similarity indices [23]. A vertex similarity measure called relation strength similarity (RSS) [16], which captures potential relationships in real-world networks, was utilized to discover missing links. Semi-local similarity index-based framework [63] was proposed by introducing resource allocation into local path index. Semantic and event-based features were added to improve the efficiency of link prediction [98]. A different algorithm which combines the topology and community information was introduced for predicting future links in Twitter network based on user's interest and behavior [92]. Supervised random walk model performs a PageRank-like random walk on networks and combines the network structure with node and edge features to perform missing link prediction [10]. Link prediction was treated as a matrix completion problem, and matrix factorization approach was introduced for missing link prediction [64, 77]. Moreover, link prediction problem was modeled as a linear optimization problem in [76]. Spectral clustering technique for link prediction [89] captures the global network structure by exploiting normalized graph Laplacian. A unified framework based on spectral graph transformation was introduced for link prediction in [47].

A few works based on the probabilistic model [42, 94] aim to abstract the underlying structure from the observed network. The prediction of missing links is done by using the learned model. A combination of maximum likelihood and Monte Carlo sampling algorithm [21] was introduced to infer existence likelihood of links in static networks. Markov model was deployed for website prediction by calculating the conditional probabilities [113]. An expectation maximization (EM) [44, 90] framework was used to model network completion problem, where they combined EM approach with Kronecker graph models to predict the missing links in the network. Probabilistic models such as probabilistic relational model (PRM) [28], probabilistic entity relationship model (PERM) [36] and Stochastic relational model [106] were also introduced for static link prediction. Applications of statistical relational learning [78] were introduced for link prediction, where the task is inherently relational. Based on the local topological embeddings of two nodes, vertex collocation profiles were proposed for link analysis and prediction [55]. A universal structural consistency index that is free of the prior knowledge of network organization was introduced in [59]. A general mathematical and computational framework that deals with data reliability in complex networks by incorporating the idea of stochastic block models (SBM) was introduced in [33]. Some surveys [9, 60, 95] cover link prediction in static networks. Moreover, static link prediction algorithms were deployed for various applications such as finding experts in co-authorship networks and friendship prediction in social media [7, 75].

Temporal Link Prediction: Problem Definition

We define temporal link prediction problem as "Let $G = (V, E)$ be a dynamic network, where V is the set of vertices and each edge $(u, v) \epsilon E$ represents a link between u and v. Given the snapshots of G represented as $G_1, G_2, \dots, G_t$ from time step 1 to t, how can we predict the network for a next time step G_{t+1}?" Figure 2 depicts an overview of temporal link prediction. This survey focuses on the existing techniques for temporal link prediction by considering the temporal aspect of dynamic networks. Further, we discuss concerns and challenges in existing techniques and provide future insights for temporal link prediction.

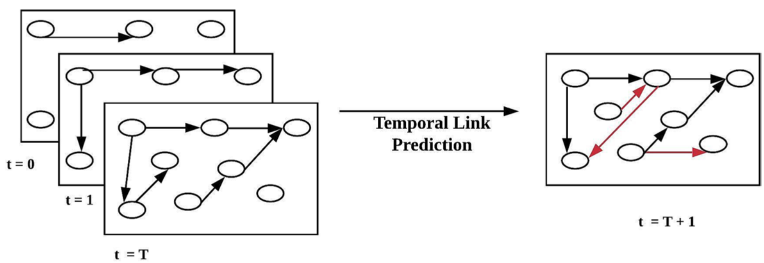

Fig. 2 An overview of temporal link prediction

Evaluation Metrics for Temporal Link Prediction

This section briefly describes the evaluation metrics used for various temporal link prediction methods described in "Temporal link prediction techniques".

1. Area under curve (AUC): AUC is a widely used evaluation metric for link prediction. This metric can be interpreted as probability that a randomly chosen missing link is given a higher score than a randomly chosen non-existent link, provided the rank of all the non-observed links. The value of this metric is bounded between 0 and 1. Among n independent comparisons, if there are n' times the missing link having a higher score and n'' times they have the same score, AUC score is calculated as:

$$\mathrm{AUC} = \frac{n' + 0.5n''}{n}. \tag{1}$$

2. Mean average precision (MAP): This metric is an extension of average precision (AP) where the average of all APs is calculated to get MAP score. It estimates the precision of every node and computes the average over all nodes. It is calculated as:

$$\begin{aligned} \mathrm{AP}(i) &= \frac{\Sigma_j \mathrm{Precision@}j(i) \cdot \Delta i(j)}{|\{\Delta i(j) = 1\}|} \\ \mathrm{MAP} &= \frac{\Sigma_{i\epsilon Q} \mathrm{AP}(i)}{|Q|}. \end{aligned} \tag{2}$$

3. Area under precision recall (AUPR): AUPR is plotted with recall against precision. It is considered as a more relevant measure for imbalanced classification problems.

$$\begin{aligned} \mathrm{Recall} &= \frac{\mathrm{True\ \ positives}}{\mathrm{True\ \ positives} + \mathrm{False\ \ negatives}} \\ \mathrm{Precision} &= \frac{\mathrm{True\ \ positives}}{\mathrm{True\ \ positives} + \mathrm{False\ \ positives}} \end{aligned}. \tag{3}$$

4. F1-score: F1-score (F-score or F-measure) is the measure of a test's accuracy. It considers both precision and recall of the test to compute the score. This metric is calculated as:

$$F1 = 2 \cdot \frac{\mathrm{Precision} \cdot \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}. \tag{4}$$

5. Root mean squared error (RMSE): This metric is defined as the square root of the mean of squared differences between predicted and actual observations. For n observations, if y_j and $y_j^{'}$ denotes the actual and predicted observations, respectively, RMSE score is given by Eq. 5.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\Sigma_{j=1}^{n}(y_j - y_j^{'})^2}. \tag{5}$$

An Overview

In this section, we provide an overview of various temporal link prediction algorithms, followed by a brief description about each method in "Temporal link prediction techniques".

Figure 3 depicts the taxonomy of various temporal link prediction methods based on various techniques used. One of the main concerns in this task is that it deals with node and link failures. The majority of the works in this area ignore this and assume that nodes remain same at all the time steps, whereas edges appear and disappear over time. This results in the loss of information which corresponds to creating a new link due to the appearance of one of its nodes and vice versa. Table 1 shows an overview of various temporal link prediction techniques. Matrix factorization-based techniques capture the local and global features and the transitional patterns of the dynamic networks for temporal link prediction. They provide better prediction results for comparatively small real-world network datasets with linear characteristics. Furthermore, either the low-rank approximation or the feature collapse technique is time and space consuming and makes matrix factorization not scalable for large graphs. Probabilistic models have higher ability to explore evolutionary patterns in dynamic networks. Although these models perform well for small networks, they become computationally expensive for large networks. Spectral clustering-based algorithms effectively track the network evolution and perform temporal link prediction task. But they are not scalable for complex dynamic networks and incur high computational complexity. Moreover, spectral clustering-based temporal link prediction techniques perform well for datasets with wide range of cluster geometries. Time series can effectively capture the dynamics of the networks. For datasets that shows significant change in their properties for particular time periods or intervals, time series based temporal link prediction algorithms perform well. However, a majority of the time series-based techniques deal with local features and thereby ignore the global topological features of the network. Traditional time series forecasting models raise challenges in capturing the non-linear varying temporal patterns in dynamic network. DL techniques have shown promising results among various temporal link prediction algorithms because of their capability of capturing transitional patterns and automatically identifying useful representations from networks. DL-based algorithms for temporal link prediction performs well for datasets with complex non-linear properties. For example, embedding techniques ignored the painstaking feature engineering by mapping the networks into a low-dimensional latent space. RBM-based models efficiently capture the temporal and

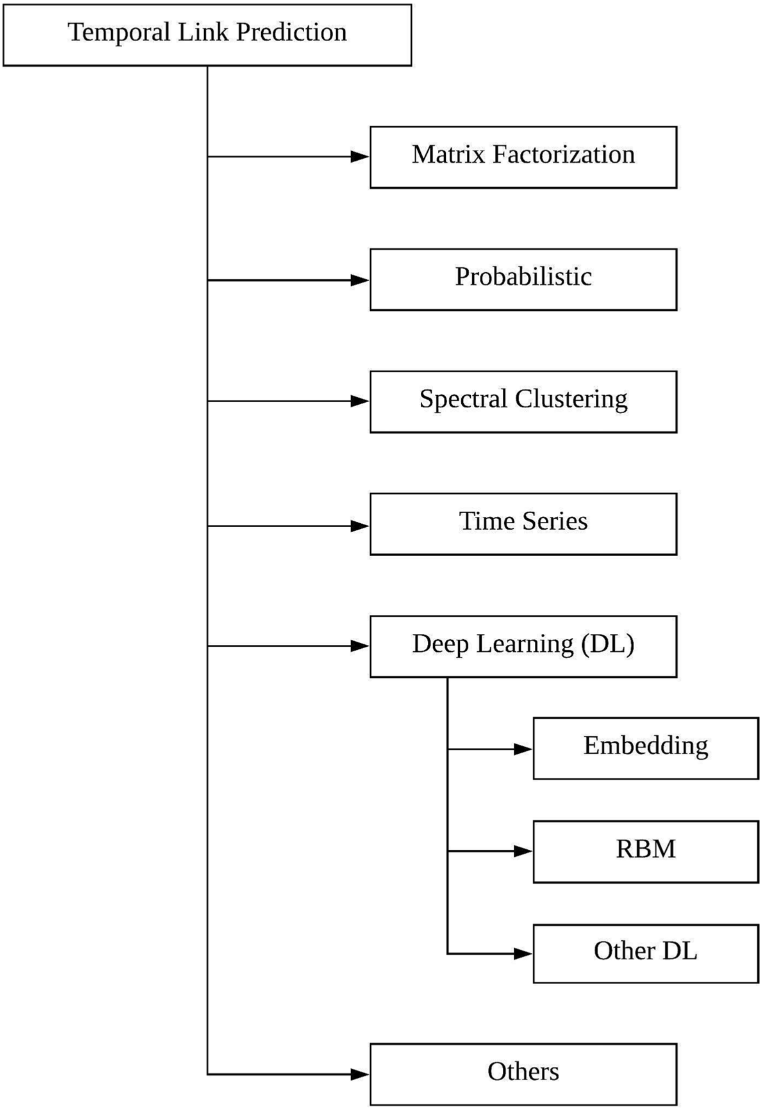

Fig. 3 Taxonomy of temporal link prediction techniques

Table 1 Comparison of temporal link prediction techniques

Category	Method used	Pros and cons
Matrix factorization	Fast and scalable link propagation [82]	Significantly less time and space complexity than traditional methods, but cannot handle temporal aspect of dynamic networks effectively
	Tensor decomposition [24]	Tensor decomposition fails to effectively capture network structure and dynamics through high-dimensional tensors which incurs high computational cost
	NMF [30]	Effectively incorporates content & structure information but tri-matrix factorization incurs high computational complexity
	DeepEye [6]	Effectively learns latent features from dynamic and topological structure of the network, but not acceptable for large, dense networks, time complexity for temporal link prediction is at least $O(n^2)$
	SNMF-FC [61]	Extracts features at each time point effectively with communicability matrices, time complexity of $O(rn^2lT)$, not acceptable for large sparse networks
	GrNMF [62]	Eliminates collapsing of networks or features and improved accuracy, but not acceptable for large-scale networks with millions of nodes
	AM-NMF [49]	Effectively performs temporal link prediction by incorporating linear characteristics of the networks, non-linear characteristics of the dynamic networks are not considered
Probabilistic	etERGM [70]	Performs effectively for small dynamic networks, but this approach is not directly scalable for large network size of YouTube or Facebook
	Non-parametric time series prediction [86, 87]	Effective prediction for large-scale dynamic networks, but incurs high computational complexity
	Time series-random walk [4]	Temporal and global information provides better accuracy for temporal link prediction, but it requires a huge amount of computation time and memory space for large-scale uncertain dynamic networks
	PBSPM [96]	Fast and robust method which gets rid of high computational complexity of temporal link prediction, largely depends on the popularity of nodes
	Stochastic Markov prediction model [22]	Scalable prediction over much larger datasets, but Markov models are not efficient to capture the non-linear varying temporal patterns
	Multilayer model [104]	Outperforms other probabilistic approaches for temporal link prediction, but not accepted for undirected dynamic networks
	TCOP [48]	Stable measure for predicting links even after a long time gap with missing history but not acceptable for large scale dynamic networks
	NP-GLM [84]	Predicts the link's apparition time effectively via a non-parametric approach, but the maximum likelihood approach is time consuming

Table 1 (continued)

Category	Method used	Pros and cons
Spectral clustering	Spectral regression [81]	Fails to capture non-linear varying temporal patterns of the dynamic network
	Tracking network evolution [99]	Performs temporal link prediction efficiently for small networks by tracking the evolution of networks, but incurs high computational complexity for large networks
Time series	TBNS [102]	Fails to capture the evolving nature of the networks, exponential smoothing cannot handle varying temporal patterns efficiently, hence forecast results are not accurate
	Interaction Prediction [83]	Performs temporal link prediction in a better way but incurs high time complexity, MA models are not efficient to capture the non-linear temporal variations in dynamic networks
	Indices Integration [3]	Performs temporal link prediction in O(n) time, but fails to efficiently capture the temporal behavior of the networks
	ITM [39]	Integration of node importance improved prediction accuracy
	TS-VLP [5]	Reduced time complexity to $O(n^2)$, but not scalable for large networks, SRW index is not efficient to completely capture the network properties
	Temporal link prediction based on ARMA, ARIMA model [34]	Fails to capture the non-linear temporal variations in dynamic networks
	Multivariate temporal link prediction based on VAR model [72]	Achieved better accuracy for temporal link prediction, but failed to incorporate global topological features
	Multivariate temporal link prediction based on NARX network [73]	Quasi-local similarity measures extract more information about underlying social structure and hence improved prediction accuracy
	Supervised temporal link prediction model [74]	Fails to capture the non-linear temporal variations in dynamic networks using ARIMA models, forecast results are not accurate in all cases
	CALAs [65]	This model considers only the local properties of the networks, fails to preserve the global properties of the networks

Table 1 (continued)

Category	Method used	Pros and cons
Embedding based	Triadic closure [111]	Preserves structural information and evolutionary patterns of the network, but fails to capture the temporal aspects of the networks
	CTDNE [69]	Incorporates temporal dependencies of the network into account and performs temporal link prediction efficiently
	SemiGraph [69]	Does not give importance to temporal dependencies in networks, first and second order proximities are not preserved efficiently
	DynLink2Vec [80]	Efficiently encodes neighborhood information and past history needed for temporal link prediction in the metric embeddings
	DynGEM [32]	Achieved consistent 2-3X speed up across a variety of different networks, but stability of the embeddings remains a concern
	DDNE [51]	Captures the non-linear transformations of nodes by preserving all the network properties and therefore improved prediction accuracy, but incurs high computational complexity
RBM based	RBM model [108]	Not acceptable for large sparse networks, incurs high computational cost, shallow learning process limits the ability of prediction
	ctRBM [52]	Reduced computational cost compared to [108], but incurs high space complexity, shallow learning process limits the ability of prediction
Other DL approaches	LIST [107]	Performs streaming link prediction by preserving both local and global properties of the network
	GRATFEL [79]	GTEs effectively captures features for temporal link prediction than other topology-based methods, but time complexity for each iteration depends on the number of possible edges
	SLWE [18]	Improves the prediction accuracy, but fails to capture the non-linear varying temporal patterns of the network effectively
Other techniques	TTM [41]	Good performance in case of sparse dynamic networks analyzed in short timescales, but does not consider important characteristics of the network
	Cross temporal link prediction [71]	Improved accuracy of temporal link prediction with the time-dependent feature projection, but it depends solely on locality preserving
	Proximity measure based on temporal events [88]	Efficient proximity measure for temporal link prediction in small networks, but fails in case of large complex networks

Table 1 (continued)

Category	Method used	Pros and cons
	CMA-ES [11]	Combines various topological measures and node-specific measures, but not acceptable for large and dense dynamic networks, does not consider persistence or decay of links over time
	MDS [27]	Ensures reliable prediction, but fails to capture the temporal or dynamic aspect of the networks effectively
	Feature engineering [67]	Performs prediction task by considering topological features, but fails to capture the global properties of networks
	Common neighbor based [103]	Fails to capture global properties and transitional patterns of the networks
	Temporal latent space model [114]	Scalable approach for temporal link prediction, but temporal smoothness assumptions are not satisfied in all circumstances
	Exended neighbor-based methods [14]	Prediction task is based on local features only, but fails to capture global properties of the networks
	DLP-ILS, DLP-IRA [110]	Low time complexity, but prediction is solely on the basis of node's adjacency relationship
	Weighting criteria based [66]	Incorporates temporal and topological information, but fails to capture the complex dynamic nature of networks
	SLIDE [50]	Cost-effective technique which stores all historical data for attributed temporal link prediction
	Evolutionary community based [20]	Computationally inexpensive and hence applicable to any size of networks, but not scalable for complex non-linear varying temporal networks
	Tensorial and bipartite model [91]	Takes into account inter-layer correlations of networks to make better predictions, but incurs high complexity

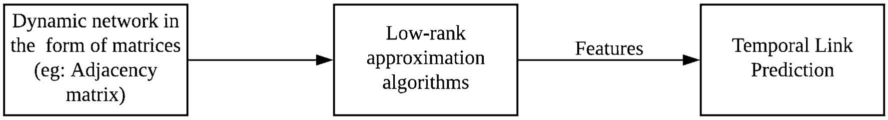

Fig. 4 Matrix factorization-based temporal link prediction

neighborhood-based features to perform the prediction task efficiently. However, they incur high computational complexity for large complex dynamic networks.

Temporal Link Prediction Techniques

In this section, we categorize various temporal link prediction techniques and propose a taxonomy based on the methods used. Given a series of snapshots of the network from period 1 to T, the goal of temporal link prediction is to predict the links in a network that would appear in its future time $T+1$. The difference between various temporal link prediction algorithms lies in how they capture the temporal or dynamic nature of the networks and also how they define the network property to be preserved. In the next section, we introduce the insights of each temporal link prediction techniques as well as how they quantify the network properties and perform link prediction task.

Matrix Factorization

Matrix factorization, also called matrix decompositions, are the methods that factorize a matrix into its constituent factors thereby making more complex operations easier. It is an effective tool for large-scale data processing and analysis. Eigen decomposition (ED) [1] and singular value decomposition (SVD) [31] were some of the low rank approximation algorithms commonly used in solving the static link prediction problem. Some of these techniques were applied for temporal link prediction.

Matrix factorization-based temporal link prediction represents the network property in the form of a matrix (e.g., adjacency matrices) and factorize this matrix to form the features for performing the link prediction task. These algorithms demonstrate how matrix factorization technique describe network properties efficiently. Pioneer studies in temporal link prediction usually solve the task in this way. The difference between various algorithms in this area lies in how they capture the dynamics of the network by preserving all the network properties. Figure 4 shows the general strategy followed in almost all the matrix factorization-based temporal link prediction techniques. The steps are as follows:

1. Dynamic networks are represented in the form of matrices (for eg: adjacency matrices).

2. Various low-rank approximation algorithms are deployed to characterize and extract underlying features of the dynamic network
3. Several parameters are introduced to capture the temporal information and network properties effectively.
4. Temporal link prediction is performed based on the latent features obtained.

Raymond et al. [82] introduced matrix factorization techniques such as Cholesky decomposition and eigen decomposition (ED) for link prediction in both static and dynamic graphs. This framework concentrated on link prediction in static graphs and was extended for link prediction in dynamic graphs. Here, the aim was to predict the existence of links for an arbitrary node pair $(x_i, y_j) \epsilon X \times Y$ incorporating the idea of various link propagation techniques. A matrix F^* is given as input whose each entry is '1' if there exists a link between a pair of nodes and '0' otherwise. Moreover, they assumed that the similarity matrices W_x and W_y among the nodes X and Y, respectively, are also known. Since the entities and relationships in a dynamic network are relatively short and instantaneous, matrix F^* can be represented as $F^*_{new} = F^* + \Delta F^*$, where ΔF^* denotes the change in F^*. Incomplete Cholesky decomposition gives the low rank approximation G_x and G_y of similarity matrices W_x and W_y obtained by Kronecker sum and product Laplacian. They computed eigenvectors of G_x and G_y followed by ED. A matrix F is obtained as output, each of whose elements $[F]_{i,j}$ is the link strength between x_i and y_j. Incremental updates in the link prediction scores are obtained based on F^*_{new} given by Eq. 6,

$$D' * \left(\bar{V}_X^T \Delta F^* \bar{V}_Y\right) \tag{6}$$

where D' is the diagonal matrix, and $\bar{V}_X$ and $\bar{V}_Y$ are the eigenvectors of G_x and G_y, respectively. Their experimental results on web graph dataset using AUC metric gives a score of 0.954. Truncated SVD (TSVD) and tensor factorization-based technique [24] represented time evolving link data as a third-order tensor $Z(i, j, t)$ given by Eq. 7 and found that the well-known Katz method [43] for link prediction can be extended to bipartite graphs. TSVD approximation was performed in a scalable way by first collapsing the temporal data into a single matrix. They proved that matrix-based approaches used were not able to fully leverage the temporal aspect and presented CANDECOMP/PARAFAC (CP) [26] tensor decomposition technique that fully leverage the temporal aspect of the data. The k-component CP decomposition of a 3-way tensor Z of size $M \times N \times T$ is given by Eq. 8,

$$Z(i,j,t) = \begin{cases} 1, & \text{if node } i \text{ links to node } j \text{ at time } t \\ 0, & \text{Otherwise.} \end{cases} \tag{7}$$

where $\circ$ denotes the outer product, $\lambda_k \epsilon R_+$, $a_k \epsilon R^M$, $b_k \epsilon R^N$ and $c_k \epsilon R^T$ for $k = 1, 2, \ldots, k$, λ_k contains scalar weight of the kth component and each summand $\lambda_k \circ a_k \circ b_k \circ c_k$ is called a component. The components extracted by CP model is used to assign scores to each pair (i, j) based on their likelihood of linking in future. For each pair of nodes (i, j), similarity score is defined using k-component CP model.

$$Z \approx \sum_{k=1}^{K} \lambda_k a_k \circ b_k \circ c_k. \tag{8}$$

Holt–Winter forecasting method which is suitable for time series data with periodic patterns is used to predict the similarity scores for future period $T+1$. Even though tensor factorization provides a better way to capture the temporal patterns, it incurs high computational cost. Their experimental result shows that CP tensor factorization-based temporal link prediction gives an AUC score of 0.928 on prediction with DBLP dataset.

The majority of the works on matrix factorization-based temporal link prediction approaches deploy non-negative matrix factorization (NMF) technique. NMF approximation is a group of algorithms where a matrix S is factorized into two matrices V and H with all the three matrices having no negative elements and the product of the two matrices approximately equal to the original matrix. NMF-based temporal link prediction was first introduced in [30]. The temporal link prediction by integrating both content and structure information was proposed in this model. Here, the temporal graph series $G = \{G_1, G_2, \ldots, G_T\}$ is represented by a series of adjacency matrices $\{A_t, t = 1, 2, \ldots, T\}$, $A_t \epsilon R^{N\times N}$.

$$\min_{U,S,V} J = \| A - USU^T \|_F^2 + \alpha \| C - UV^T \|_F^2 + \lambda \mathrm{tr}(U^T LU). \tag{9}$$

Given a series of adjacency matrices and content matrix C, the model predicts the occurrence probabilities of links in a future time $T+1$ by integrating latent NMF tri-factorization and graph regularization techniques. The objective function for this model is given by Eq. 9, where the factor matrices $U \epsilon R_+^{N\times K}$, $S \epsilon R_+^{K\times K}$ and $V \epsilon R_+^{M\times K}$ are non-negative and $\| . \|_F$ denotes Frobenius norm. α and λ are the parameters that control the influence of content information and graph regularization respectively. The first term in Eq. 9 corresponds to temporal link matrix tri-factorization, second term incorporates the content and semantic attributes of the nodes and the third term integrates graph proximity information into the joint model. Here, tr(.) denotes the trace of a matrix. The probability of a link existing between a pair of nodes can be obtained by the inner product of U and S, which gives the score matrix for predicting future link occurrence . Their experimental result shows that this method gives an average AUC score above 70% for hep-th dataset. In contrast to matrix tri-factorization, a different approach for NMF-based temporal link prediction was introduced in [6], where they modeled NMF as a problem of optimizing the Frobenius norm given by Eq. 10 such that X is the adjacency matrix and the factor matrices $U \geq 0$ and $V \geq 0$.

$$\min_{U,V} \| X - UV^T \|_F^2 \tag{10}$$

They aimed at finding a consensus latent space matrix U^* and a coefficient matrix V^*, so that for every consensus latent space matrix $U^{(t)}$ and coefficient matrix $V^{(t)}$, the differences between V^* and $V^{(t)}$, U^* and $U^{(t)}$ will be minimized. Similarities between row vectors in V^* are computed using correlation coefficient and cosine similarity metrics. Temporal link prediction is performed based on the similarity score obtained. This framework presented an iterative model for NMF that learns latent features from temporal and topological structure of dynamic networks. Their results show that this method gives an AUC score above 80% for enron dataset. Later, [61] introduced communicability in networks [25] into account and solved NMF-based temporal link prediction in a different way. They first proved the equivalence of NMF and ED taking into account graph communicability. Communicability in networks can be defined as weighted sum of all the walks connecting a pair of vertices. Based on the proven equivalence, symmetric NMF- based feature collapsing (SNMF-FC) was introduced. SNMF is an extension of NMF in which a matrix X is approximately factorized into a matrix $Z_{n\times r}$ such that $X \approx ZZ'$, where $Z \geq 0$. SNMF algorithm factorizes communicability matrix corresponding to each snapshot into $n \times r$ feature matrix Z_t and solves the optimization by squared error. Once the feature matrices $Z_1, Z_2, \dots, Z_T$ are obtained, they are collapsed into a feature matrix F where a parameter is introduced to balance the relative importance of feature matrix Z_t. From the collapsed feature matrix, temporal links are predicted as Eq. 11,

$$W_{T+1} = \sum_{l=1}^{m} W_{T+1}^{[l]}/m \tag{11}$$

where m is the number of multiple runs. SNMF-FC requires total time complexity of $O(rn^2lT)$, where r is the number of features and l is the number of iterations. Their results on DBLP dataset give an AUC score of 0.65. To further improve the accuracy of the SNMF-FC algorithm, graph regularized NMF (GrNMF) [62] eliminated the feature collapsing technique and provided a better way to characterize the topological information in temporal networks. Firstly, NMF factorizes the network at time T into a base matrix ($B_{n\times k}$) and a feature matrix ($F_{k\times n}$). On the second concern, graph regularizationprovides an efficient strategy. Optimization function computes update rules, where at each iteration, either matrix B or matrix F is optimized while keeping the other fixed.

$$W_{T+1} = BF. \tag{12}$$

Temporal link prediction in GrNMF is given by Eq. 12. Their result shows that this method that leverages the power of NMF and graph regularization technique improved the accuracy of prediction without increasing the time complexity. They have proved that DBLP dataset gives an AUC score of 0.8 for temporal link prediction using this framework. Recently, NMF-based network embedding [49] was introduced for temporal link prediction. The goal of this technique was to predict the adjacency matrix for a future time $T + 1$, given a series of adjacency matrix for each snapshot network. The standard NMF process is given by Eq. 13, where $A_t \epsilon R^{N\times N}$ is the adjacency matrix of the snapshot at time t, $Y_t \epsilon R^{N\times K}$ and $X_t \epsilon R^{N\times K}$ are the coefficient matrix and base matrix, respectively, and k is the dimension of the hidden space. Here, they introduced adaptive multiple NMF (AM-NMF) algorithm which learns a unified embedding representation of the given network snapshots by simply combining multiple NMF components of several successive time slices. For the temporal link prediction task, the objective function for AM-NMF is modeled as a linear combination of multiple NMF components with respect to $T + 1$ successive time slices.

$$\text{argmin}_{X_t, Y_t \geq 0} \parallel A_t - Y_t X_t^T \parallel_F^2 \tag{13}$$

Each component's relative contribution can be automatically adjusted by the corresponding adaptive parameter. This technique effectively performs temporal link prediction by incorporating linear characteristics of the dynamic networks. Their result shows that this framework gives an AUC score of 97.73% for the KAIST dataset (Table 2).

Probabilistic Approaches

Probabilistic models deploy maximum likelihood approaches or probability distributions instead of fixed values, thereby enabling variation and uncertainty to be quantified. Probability distribution describes the range of possible values and shows which values within the corresponding range are most likely to occur. This should provide a better basis for temporal link prediction, because the full range of possible outcomes can be taken into account. A few works in temporal link prediction focused on building probabilistic models. Figure 5 shows the general idea behind probabilistic temporal link prediction models.

The general steps followed by probabilistic models for temporal link prediction are as follows:

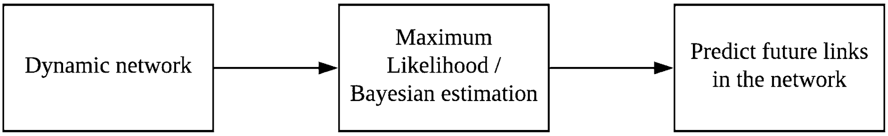

Fig. 5 Probabilistic models for temporal link prediction

Table 2 Comparison of matrix factorization-based temporal link prediction techniques

Method	Objective function	Description
Fast and scalable link propagation [82]	Equation 6	Fast and scalable semi-supervised algorithm using both link and node information
Tensor decomposition [24]	Equation 8	Incorporated temporal information into CP tensor-based link prediction analysis
NMF [30]	Equation 9	A joint model based on matrix tri-factorization which integrates global network structure, contents of nodes and graph proximity information
DeepEye [6]	Equation 10	Demonstrates how latent NMF features can express network dynamics efficiently
SNMF-FC [61]	Equation 11	NMF-based temporal link prediction using graph communicability and feature collapsing
GrNMF [62]	Equation 12	Introduced graph regularization and eliminated feature collapsing for temporal link prediction
AM-NMF [49]	Equation 13	Under the framework of NMF, embeds the dynamic network into a low-dimensional hidden space where the characteristics of the network are preserved

1. Dynamic networks are represented as a sequence of snapshots at regular intervals of time.
2. Either maximum likelihood approach or probability distributions like joint probability distribution, Bernaulli distribution, etc., are used to infer the likelihood of links in a future time.

Extended temporal exponential random graph model (etERGM) [70] predicts node attributes and links from temporal networks solely based on the historical data. This model builds two conditional exponential random graph models instead of training a single joint probability prediction model. The model takes Markovian assumption that each network matrix is conditionally independent of all other prior observations, given its immediate prior observed matrix. Given the social network link observations $N^1, N^2, \dots, N^T$, and the node attribute observations $x_1, x_2, \dots, x_T$, where x_t is a k-length vector, they aimed at making accurate predictions of the N^{T+1} and x_{T+1} in the future step. For node prediction, the log linear model $P(x_t|x_{t-1}, N^t, \gamma)$ in Eq. 14 is used. It describes the transition of attributes from time $t-1$ to time t, conditioning on the network structure N^t at time t. Here, Z is a normalization constant, $\Pr(x^{t-1})$ is Gaussian multivariate regularization prior estimated from training data, and γ is a vector corresponding to sufficient statistic vectors ψ which encodes the dependencies between actor links and attributes. They used a sampling procedure to learn γ, where Metropolis–Hastings algorithm [17] is used to sample from $P(x_t|x_{t-1}, N^t, \gamma)$ distribution. Their results show that the model gives an accuracy score of 85.5% for the delinquency dataset.

$$P\left(x^t|x^{t-1}, N^t, \gamma\right) = \frac{1}{Z(x^{t-1}, N^t, \gamma)} \exp\left\{\gamma' \psi\left(x^t, x^{t-1}, N^t\right)\right\} \Pr\left(x^{t-1}\right). \tag{14}$$

Some of the probabilistic temporal link prediction algorithms follow non-parametric approaches [86, 87], where a framework for non-parametric time series prediction is proposed which models the evolution of a sequence x_t over time. Based on the features and local neighborhood around the endpoints, the model predict links under the assumption that links are formed independent of each other. This allows for different types of neighborhoods in a graph, each with its own dynamics. The model is given by $Y_{t+1}(i,j)G \sim \text{Bernoulli}(g(\psi_t(i,j))), \psi_t(i,j) = \{s_t(i,j), d_t(i)\}$, where, $0 < g(.) < 1$ is a function of two sets of features: neighborhood-specific features ($d_t(i)$) and pair-specific features ($s_t(i,j)$). Their estimate of the function $g(.)$ for two pairs of nodes (i,j) and (i',j') in future time T is given by Eq. 15,

$$g_T(\psi_T(i,j)) = \frac{\sum_{i',j',t'} \Gamma(\psi_T(i,j), \psi_{T'}(i',j')).Y_{t'+1}(i',j')}{\sum_{i',j',t'} \Gamma(\psi_T(i,j), \psi_{T'}(i',j'))} \tag{15}$$

where the kernel function $\Gamma(\psi_T(i,j), \psi_{T'}(i',j'))$ is further divided into pair-specific and neighborhood-specific parts. Their evaluation results show that this framework gives an AUC score of 0.89 on hep-th dataset. A different algorithm was proposed for temporal uncertain networks [4], which takes into account time series-based probabilistic random walk approach for temporal link prediction.

They described temporal uncertain networks as a sequence of snapshots at regular intervals of time $G_t = (V, E_t, P_t, A_t)$, where $t = t_0, t_{0+1}, \dots, t_{0+T-1}$, P_t is the probability matrix and A_t is the adjacency matrix of G_t. Here, the goal of time series link prediction problem in temporal uncertain networks is to predict the occurrence probabilities of edges at time t_{0+T}. For a pair of nodes (u, v) in different snapshots, the algorithm first computes probabilistic random walk transformation matrices $W_{t_0}(u, v), W_{t_{0+1}}(u, v), \dots, W_{t_{0+T-1}}(u, v)$ and then combines this sequence into one transformation matrix W as in Eq. 16, where γ is a damping factor which gives more importance to recent snapshots. Based on the transformation matrix, SimRank index which gives the occurrence probability score of the edge is applied to calculate the similarity score between the pair of nodes.

$$W = \sum_{t=t_0}^{t_0+T-1} \gamma^{t_0+T-1} W_t. \tag{16}$$

Wang et al. introduced popularity-based structural perturbation method (PBSPM) [96] and its fast algorithm to characterize the likelihood of an edge from both existing connectivity structure and current popularity of its two end points. This framework was proposed under the hypothesis that observed network is determined by some latent attractors (e.g., similar hobbies, ages, gender, location) that independently influence the structural properties. They represented an attractor as $x_k = [x_{k,1}, x_{k,2}, \dots, x_{k,n}]^T$, where $x_{k,i}$ represents attractiveness of node i for the latent attractor x_k. Probability P_{ij} is defined as the weighted influence of each attractor, where a parameter is used to balance the relative influence of each attractor x_k.

$$A' = \sum_{k=1}^{n} \left(\lambda_k + \Delta\lambda_k\right) x'_k x'^T_k \tag{17}$$

They argued that the ability for a node i to attract new edges is determined by both latent attractors and its current popularity and introduced an advanced attractiveness $x'_{k,i}$ which is a combination of the static attractiveness and popularity. Temporal link prediction was performed with the advanced attractiveness and is given by Eq. 17, where A' is the adjacency matrix for a future time and λ is the eigenvalue. They evaluated the performance of the framework using precision metric on the haggle dataset and their results show that the method gives a precision score of 0.3475 on the haggle dataset. Das et al. presented a stochastic Markov model [22] that takes into account time-varying graphs as its basis for temporal link prediction. In this framework, they visualized network evolution as a process in which the network goes through a sequence of states. Each state corresponds to a graph that varies with respect to time. The network evolution is interpreted as a transition from one network state to another. The temporal analysis in this model considers fine-grained and coarse-grained timescales along with link and cluster evolution, respectively.

$$\mathrm{Corr}(t_\mathrm{b}, t_\mathrm{a}) = \frac{|L(t_a) \cap L(t_\mathrm{b})|}{|L(t_\mathrm{a}) \cup L(t_\mathrm{b})|} \tag{18}$$

The observation time interval [0, T] is divided into coarse-grained time intervals, called time slots $T_1, T_2, \ldots, T_i, T_{i+1}, \ldots, T_k$ and each of the time slots T_i is further divided into fine-grained time interval called time stamps which reflects the smallest interaction time unit. The correlated link evolution is determined by Eq. 18, where $L(t_\mathrm{a})$ and $L(t_\mathrm{b})$ corresponds to the corresponding sets of links over time stamps t_a and t_b, respectively. They considered network evolution as a process where the network goes through a sequence of states in a Markovian model. The algorithm for computing link prediction probability initializes a Markovian transition matrix. For subsequent intervals, it selects the source node and calculates destination node selection probabilities. Interaction probabilities are calculated by considering whether the source and destination node pairs are in same or different clusters. The interaction takes place across the selected start and destination nodes with interaction probability above certain threshold limit and the iteration is continued for each successive interval till the end of training. The parameters for the model are obtained from correlated evolution of links (local correlation) and semi-global correlation (over clusters). Their evaluation results show that the model gives an AUC score of 0.973 on DBLP dataset. A multilayer model for temporal link prediction [104] examines the layers evolution (layers birth/death process and lifetime) throughout the network. They modeled the evolution of each node's membership in different layers by an infinite factorial hidden Markov model [29] considering feature cascade and thereby formulated link generation process for inter-layer and intra-layer links. The probability that an active layer α at time step τ_{t-1} remains active until the next time step τ_t described by exponential distribution is given by Eq. 19, where γ is introduced to control the rate of layers birth/death.

$$P(w^{\alpha\tau_t} | w^{\alpha\tau_{t-1}} = 1) = e^{-\gamma}. \tag{19}$$

This model exhibited more predictive power for cascade link prediction. Their experimental results show that the framework gives an F1 score of 0.55 for the DBLP dataset. Temporal probabilistic measure called temporal co-occurrence probability (TCOP), which is an extended form of co-occurrence probability (COP) was introduced for link prediction by incorporating the temporal information in collaborative networks [48]. For a pair of nodes (i, j), TCOP algorithm first computes central neighborhood set (CNS) of i and j and then constructs Markov random field (MRF) with the nodes in CNS(i,j). MRF construction involves computation of clique potentials, which is an assignment of weights to all subsets of C, called factors F of C. Temporal weight for a factor F is defined as Eq. 20,

$$\text{Temporal-weight}(F) = \frac{w(F) \cdot \beta^{r(F)}}{|t_{\max}(F) - t_{\min}(F)| + 1} \tag{20}$$

where $\beta < 1$ is a damping factor. Finally, the algorithm infers the joint probability between nodes i and j. Their results proved that TCOP is a stable measure for predicting links even after a long time gap with missing history and gives an AUROC score of 0.797 for the DBLP dataset. Probabilistic non-parametric approach called non-parametric generalized linear model (NP-GLM) [84] infers the hidden underlying probability distribution of the time of link creation between a pair of nodes, given its features. Here, they proposed an inference method to answer queries such as the most probable time by which a link will appear between two nodes and the probability of link creation between two nodes during a specific period. Specifically, given the feature vector x_l for a missing link l at time t_0, NP-GLM predicts t_l, which indicates how long after t_0, the link l will appear in the network. Here, the probabilistic approach for this problem models a conditional distribution $f_T(t_l|x_l)$ given by Eq. 21.

$$\pi_{i=1}^{N} f_T(t_i|x_i)^{y_i} P\left(T \geq t_i|x_i\right)^{1-y_i}. \tag{21}$$

The strength of this framework is that it can learn the underlying distribution of data as well as the amount of contribution of each extracted feature for the link advent time in the network. Their experimental results show that this framework gives an MAE score of 10.59 for the Weibo dataset (Table 3).

Spectral Clustering

A few works on temporal link prediction rests on spectral graph theory, which is the study of properties of a graph in relationship to the eigenvalues and eigenvectors of adjacency matrix or Laplacian matrix associated with the graph. These algorithms deploy low-rank approximation approach that works well for large-scale graphs for which full matrix decomposition algorithms does not work well. Figure 6 depicts the basic idea of spectral clustering-based temporal link prediction, where the spectra of

Table 3 Comparison of probabilistic models for temporal link prediction

Method	Objective function	Description
etERGM [70]	Equation 14	Model that facilitates simultaneous prediction of links and attribute values in temporal networks solely based on historical data
Non-parametric time series prediction [86, 87]	Equation 15	Non parametric model for link prediction which uses graph-based features of node pairs as well as local neighborhoods for temporal link prediction
Time series random walk [4]	Equation 16	Temporal link prediction problem is formalized by designing a time series-based random walk in temporal uncertain networks
PBSPM [96]	Equation 17	Introduced a hypothesis that the ability of each node to attract links depends not only on its structural importance, but also on its current popularity
Stochastic Markov prediction model [22]	Equation 18	A Markov prediction model which considers multiple time scales in leveraging temporal analysis for link prediction
Multilayer model [104]	Equation 19	Probabilistic multilayer model which examines evolution of layers throughout the network evolution
TCOP [48]	Equation 20	Time score is incorporated into graphical model framework, yielding a new measure called TCOP
NPGLM [84]	Equation 21	A probabilistic non-parametric approach which infers hidden underlying probability distribution of link advent time, given its features

graphs which are the collection of eigenvalues of corresponding Laplacian matrix and low-rank approximation of the graph Laplacian together provide the meaningful features for temporal link prediction (Table 4).

The steps below describe the basic principle used in spectral clustering-based temporal link prediction.

1. Dynamic network is represented as a sequence of snapshots in the form of Laplacian matrices.
2. The spectra of graphs and low-rank approximation of the graph Laplacian form essential features for link prediction.
3. Several models such as time series forecasting models and linear system models are used to predict the feature vector for a future time.
4. Link prediction is performed based on the predicted feature vector.

Fang et al. [81] introduced a framework that rests on spectral graph theory and low-rank approximation for the graph Laplacian matrix. The idea behind this framework is to use ARMA model to predict links based on time series of graph spectra. This spectral regression model with low-rank approximation takes the normalized Laplacian matrices corresponding to each snapshot of the dynamic network and performs the temporal link prediction task in three steps which includes feature extraction, edge set prediction and graph reconstruction. Low-rank approximation of an $N \times N$ graph Laplacian matrix at time t (L_t) is given by Eq. 22,

$$H_t^k = \sum_{i=1}^{k} \lambda_t^i x_t^i (x_t^i)^T \tag{22}$$

where $K \leq N$, λ_t^i and x_t^i are the orthonormal eigenvectors and eigenvalues of L_t. To preserve the neighborhood structure and contribute more to the graph reconstruction,

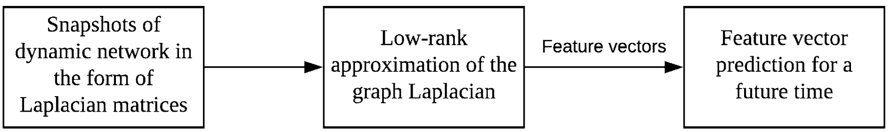

Fig. 6 Spectral clustering-based temporal link prediction

Table 4 Comparison of spectral clustering-based temporal link prediction

Method	Objective function	Description
Spectral regression [81]	Equation 22	Temporal regression model for link prediction that rests on spectral graph theory and low-rank approximation
Tracking network evolution [99]	Equation 23	A framework to track, model and predict dynamic network structures using FIR filters

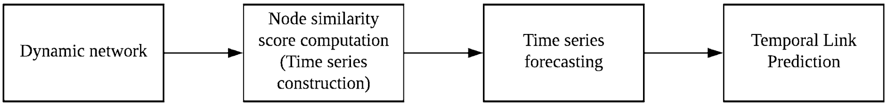

Fig. 7 Time series-based temporal link prediction

eigenvectors corresponding to k largest eigenvalues of the graph Laplacian were chosen. The behavior of their components over time were modeled using auto regressive moving average (ARMA) model [13]. They observed that the evolution of dynamic network and the change in the eigenvalues are related which suggested that the spectrum of Laplacian at time T (λ_T) can be used for the low-rank approximation at time $T+1$ (λ_{T+1}). Eigenvalues λ_T corresponding to the eigenvectors X_T are used to estimate λ_{T+1}. Using the low-rank approximation in Eq. 22 and predicted eigenvectors X_{T+1}, H^K_{T+1} is constructed. Link prediction result is obtained by comparing H^K_t and H^K_{T+1}. Wu et al. [99] presented another framework that deployed spectral graph theory and low rank approximation for temporal link prediction. In contrast to the ARMA model, this framework utilized finite impulse response (FIR) filter that is a linear system model to track the evolution of latent feature vectors of each node.

$$z'_{i,k}(T) = \alpha_{i,k,1} z_{i,k}(T-1) + \cdots + \alpha_{i,k,J} z_{i,k}(T-J) + b_{i,k}. \tag{23}$$

The objective function for feature vector prediction using FIR filter is given by Eq. 23, where $\alpha_{i,k,j}$, for some $j = 0, 1, \ldots, J$ is the impulse response and $b_{i,k}$ is the bias. The best estimate for $z'_{i,k}$ can be obtained by solving ridge regression. For a pair of nodes (u, v), cosine similarity score is calculated from the predicted feature vectors. If the score is beyond a predefined threshold, there exists a link between them.

Time Series

Time series-based temporal link prediction deploys various time series forecasting models for predicting links in the network for a future time period. Time series score is constructed by computing various similarity measures between each node pairs in the network. Time series forecasting is the practice of predicting future time series scores based on the previously observed scores. Figure 7 shows the general strategy deployed in time series-based temporal link prediction. The frequently changing structure of the network data makes time series a promising option for temporal link prediction. Various time series forecasting models were also introduced to predict the future similarity scores between node pairs.

General steps followed by time series-based temporal link prediction techniques are as follows:

1. Dynamic networks are represented as a sequence of snapshots for different time periods.
2. For each pair of nodes in the network, time series is constructed based on various node similarity measures.

3. Time series forecasting models are deployed to predict the time series score for a future time period.
4. Link prediction is performed based on the predicted score.

Tensor-based node similarity (TBNS) [102] for link prediction eliminates the calculation of three-dimensional tensor and improves efficiency. This algorithm consists of two independent steps. First, data are stored into a tensor Z which is a set of slices which represents the relational data of T time periods. They mapped the three-dimensional tensor into a two-dimensional matrix which contains trends within times of each edge using the exponential smoothing method given by Eq. 24,

$$Z_{t+1} = \alpha^* Z_t + \alpha^* (1-\alpha)^* Z_{t-1} + \cdots + (1-\alpha)^{t*} Z_1 \tag{24}$$

where α is the smooth factor whose value range is $0 < \alpha < 1$, Z_{t+1} stands for two-dimensional matrix with weight that is transformed by exponential smoothing method, and the value of $Z_{t+1}(i,j)$ is the intimacy between objects i and j at time $T+1$. Higher the weight in the matrix, closer is the relationship between the two objects. They applied common neighbors, which quantize the intimacy between two objects to compute the degree of similarity for each node. The values in the matrices now show the intimacy between nodes at time $T+1$. Greater the value, the higher is the similarity or greater is the chance that they will have connection at $T+1$. Their experimental result shows that this method gives an AUC score of 0.938 on Facebook wall posts dataset. A different approach for interaction prediction in dynamic networks [83] was introduced by exploiting the community discovery. The idea behind this framework is to make use of time-stamped network observations and community knowledge, besides classical features for learning a model which is able to forecast new interactions between nodes. This method is built upon a supervised learning strategy, where for each temporal snapshot $t \epsilon T$, a partition $C_t = \{C_{t,0}, C_{t,1}, \dots, C_{t,k}\}$ is computed using community discovery algorithms such as Louvain, Infohiermap and Daemon [12]. For each t and C, $G_{C_t} = (V_t, C, E_t, C)$ is defined as the subgraph induced on G_t by nodes in C_t, such that $V_{t,C} \subseteq V_t$ and $E_{t,C} \subseteq E_t$. Interaction between communities C_t of G_t are considered for each time t, and set of measures f is computed for each pair of nodes $(u, v)\ \epsilon\ C_t$. The values $f^{u,v}$ describe topological, structural and community features of node pair (u, v). For each pair of nodes (u, v), a time series $S_{f^{u,v}}$ is constructed using a sequence of measures $f_0^{u,v}, f_1^{u,v}, \dots, f_t^{u,v}$. They used a well-known forecasting technique, moving average (MA) model to predict the future expected value $f_{T+1}^{u,v}$. The objective function for MA model is given by Eq. 25, where μ denotes the mean of the series, ϵ_i indicates white noise errors and θ is the parameter of the model. The forecasted values are given as features for a classifier which predicts the future intra-community interactions.

$$z(t) = \mu + w_t + \sum_{i=1}^{q} \theta_i w_{t-i} \tag{25}$$

Their results show that this method gives an AUC score of 0.907 on DBLP dataset. Ahamed et al. introduced two new approaches for temporal link prediction [3]. One approach is based on the reduced static graph by using a modified reduced adjacency matrix to reflect the frequency of each link. Another approach integrates similarity indices of the nodes to exploit both temporal and topological information. Indices integration algorithm for temporal link prediction was based on five different similarity indices such as common neighbor (CN), Adamic Adar (AA) [2], Jaccard Coefficient (JC) [40], preferential attachment (PA) and Katz. The five different index matrices $S_1, S_2, \dots, S_T$ for graphs $G_1, G_2, \dots, G_T$ were integrated to construct $S_{1,T}$ given by Eq. 26,

$$S_{1,T} = \sum_{t=1}^{T} \gamma^{T-(t-1)} S_t \tag{26}$$

which is a matrix consisting of the indices of the probabilities for future links where the damping factor assigns more importance to recent topological information. Their results show that the method gives an AUC score of 0.8914 for CN similarity index on Enron dataset. Integrated time series model (ITM) [39] combines three different types of information which includes the community information, node centrality information and time series information. For different sequences of snapshot networks with window size T, a modified static graph representation is generated. They defined two different matrices N_{com} and N_{cen} which represents community and centrality informations respectively and used these matrices to construct probability matrix $S_{t_0,T}$ as given by Eq. 27, where β is the community matrix integration parameter. The resulting matrix consists of occurrence probabilities for future links and carries time, inner link and community information.

$$S_{t_0,T} = \beta * N_{\text{com}} + N_{\text{cen}} \tag{27}$$

Time series-vertex link prediction (TS-VLP) algorithm [5] takes into account sampling technique for similarity computation that reduces the computation time. The algorithm first combines snapshots of temporal network into a weighted graph using a damping factor to assigngreater importance to more recent snapshots. Subgraphs are generated by random walk staring from each node.

$$S_{xy}^{\text{SRW}} = q_x T'(L, x, y) + q_y T'(L, y, x) \tag{28}$$

Superposed random walk (SRW) index is computed within the small subgraph centered at each node as given by Eq. 28, where q_x and q_y are the initial configuration functions. $T'(L, x, y)$ is calculated by counting the number of paths including node y in the set of sampled paths starting from node x and $T'(L, y, x)$ is calculated by counting the number of paths including node x in the set of sampled paths starting from node y. The error of the estimated similarities is restricted within a given threshold by choosing a proper number of sampled paths. They proved that the algorithm effectively integrates the local and global topology of the networks and reduces the computation time to $O(n^2)$.

Univariate time series-based temporal link prediction [34] takes into account the neighborhood-based similarity measures for nodes in the network. This framework takes adjacency and occurrence matrices corresponding to each snapshot network as input and performs temporal link prediction in three steps: node similarity score computation, node similarity score prediction and link prediction. For each pair of nodes (u, v), different neighborhood-based similarity scores are computed for each time period.

$$\mathrm{score}(u, v, t) = \sum_{i=1}^{p} \phi_i \mathrm{score}(u, v, t - i) + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \tag{29}$$

They used ARMA and ARIMA models [19] for time series forecasting, which takes the similarity scores computed in the previous phase as input and predicts the time series score for a future time $T + 1$ given by Eq. 29, where the coefficients ϕ and θ are calculated using maximum likelihood estimation, and ϵ_t is a constant. Forecasted future scores are used to predict how likely two nodes are to connect in future. Their evaluation result shows that this method gives an AUC score above 0.9 on the hep-th dataset for time series-based temporal link prediction using common neighbor similarity measure. Unlike univariate time series models, multivariate time series link prediction models [72, 73] integrate temporal evolution of the network, node similarities and node connectivity information. The use of multivariate time series models can integrate covariance structures and link occurrence information simultaneously for accurate prediction of both new links and repeated links. In [72], they constructed the multivariate time series for each node pair (u, v) in different snapshot network using five different local neighborhood-based similarity measures such as CN, AA, JC, PA and resource allocation (RA) [112]. Vector autoregression (VAR) model was used for multivariate time series forecasting which enables to represent time series information over a combination of node similarity and node connectivities. The objective function for VAR model prediction is given by Eq. 30,

$$Y_t' = C + \pi^1 Y_{t-1} + \cdots + \pi^p Y_{t-p} + \epsilon_t \tag{30}$$

where C is a vector of dimension n of intercepts, $\pi^1, \ldots, \pi^p$ are the coefficient matrices and ϵ_t is a vector of dimension n of errors. Temporal link prediction is performed based on the predicted time series score. Their result shows that this framework gives an AUC score of 0.82 on the unweighted DBLP dataset. In contrast to the neighborhood-based similarity measures [72], quasi-local and local topology measures such as local random walk (LRW), SRW and local path (LP) index [73] were used to construct the multivariate time series for each pair of nodes (u, v). This model incorporated NARX network [100] architecture in series–parallel mode for efficient time series forecasting. NARX networks better implement the correlation among similarity measures and link occurrence information simultaneously as given by Eq. 31,

$$y'_{\mathrm{sp}}(t + 1) = f\left[y(t), \ldots, y(t - d_y); u(t), \ldots, u(t - d_u)\right] \tag{31}$$

where the function $f(.)$ denotes the non-linear mapping function, and d_u and d_y denote the number of past exogenous variables and number of delayed target variables, respectively. Link prediction is performed based on the predicted time series score. Their result shows that this framework yields an AUC score of 0.92 with DBLP dataset for SRW index.

Supervised temporal link prediction using time series of similarity measures [74] considers the dynamic topology of social networks. For each pair of nodes (u, v) in different snapshots of the network, time series is constructed using various similarity measures such as CN, AA, JC, SI, HPI and PA. ARIMA model was deployed for time series forecasting, where they predicted the future similarity for each node pair. Each pair of nodes can be represented as a vector of features which consists of a set of forecasts of all similarity measures. The predicted similarity scores are selected as features which are employed by supervised classifiers such as random forest [53] and support vector machine (SVM) [8] to apply link prediction task. Their result shows that this framework gives an AUC score above 0.8 on the DBLP dataset using Random Forest and SVM classifiers. Another work introduced a different approach for temporal link prediction based on temporal similarity metrics and continuous action set learning automata (CALA) [65]. CALA is a reinforcement-based optimization tool which tries to learn the optimal behavior from the environmental feedbacks. Here, each pair of nodes (u, v) in different snapshot networks is represented as a feature vector where seven different similarity indices such as CN, JC, Katz, Salton [85], PA, AA and LP are used as features. The goal of this work was to model link prediction problem as a noisy optimization problem given by Eq. 32,

$$\text{maximize}\left(\sum_{t=1}^{T} \gamma^{T-t+1} |A_t(X_t W - \theta_t)|^2\right) + \delta \left| \left(\wedge_{T+1}(X_{T+1} W - \theta_{T+1}\right) \right|_2 \tag{32}$$

where X_t is the similarity matrix at time t, A_t is an indicator matrix, γ is a time decay factor, θ_t is a threshold for each time t and δ is a regularization term. They used a team of CALAs to solve the noisy optimization problem. To determine the importance of different time periods and similarity metrics on the prediction result, they defined a coefficient for each different period and similarity metrics and used a CALA for each coefficient. Each CALA tries to learn the true value of the corresponding coefficient. Final link prediction is obtained from a combination of different similarity metrics in different time periods based on the obtained coefficients. Their results show that time series-based temporal link prediction using CALAs gives an AUC score above 0.9 for the haggle dataset (Table 5).

Deep Learning

Deep learning (DL) also called deep structured learning has shown its outstanding performance in various fields such as financial services and health care. DL techniques for temporal link prediction can be divided into three categories: embedding based, restricted Boltzmann machine (RBM) based and other DL techniques.

Table 5 Comparison of time series-based temporal link prediction

Method	Objective function	Description
TBNS [102]	Equation 24	Tensor-based prediction method which exploits exponential smoothing and common neighbor similarity metric for each node
Interaction prediction [83]	Equation 25	Supervised learning strategy which exploits time series forecasting model and community discovery for interaction prediction
Indices Integration [3]	Equation 26	Exploits both temporal and topological information by integrating similarity indices at all the time steps
ITM [39]	Equation 27	Performs temporal link prediction by integrating temporal information, community structure and node centrality in the networks
TS-VLP [5]	Equation 28	Fast similarity-based link prediction that can achieve high-quality results via sampling-based random walks
Temporal link prediction based on ARMA, ARIMA model [34]	Equation 29	Link prediction based on time series forecasting of neighborhood-based node similarity scores using ARMA and ARIMA models
Multivariate temporal link prediction based on VAR model [72]	Equation 30	Proposed multivariate time series forecasting-based temporal link prediction model which integrates temporal evolution, node similarities and connectivity information
Multivariate temporal link prediction based on NARX network [73]	Equation 31	Temporal link prediction model based on NARX network which incorporated quasi-local node similarity scores
Supervised temporal link prediction model [74]	Equation 29	Supervised temporal link prediction technique where the forecasted time series scores are employed as features for a classifier
CALAs [65]	Equation 32	Temporal link prediction method based on temporal similarity metrics and CALAs

Embedding-Based Temporal Link Prediction

Network representation learning (NRL) [101] or otherwise graph embedding techniques eliminated the need for painstaking feature engineering. The goal of this approach is to represent a graph in a low-dimensional vector space by preserving all the network properties. Different algorithms for graph embedding differ in the way they preserve all the network properties. Figure 8 depicts how embedding-based temporal link prediction is performed. Embedding-based temporal link prediction aims at predicting the links of a network in a future time from the low-dimensional embedding vectors.

The basic steps followed by various embedding-based temporal link prediction algorithms are as follows:

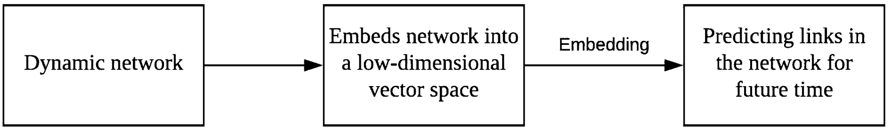

Fig. 8 Embedding-based temporal link prediction

1. To capture the temporal aspect, dynamic networks are represented as a sequence of snapshots or contact sequences or interval graphs.
2. Deploy various DL-based techniques to embed the network into a low-dimensional vector space by preserving all the network properties.
3. Embeddings are predicted for a future time period which gives the structure of the network for that time.
4. Temporal link prediction is performed based on the predicted embeddings.

Dynamic network embedding by modeling the triadic closure process [111] analyzes how an open triad evolves into a closed triad. This model was introduced under two assumptions: social homophily and temporal smoothness. Social homophily suggests that highly connected vertices should be embedded closely in the latent representation space and the goal of temporal smoothness is to minimize the Euclidean distance between the embedding vectors in adjacent time steps. The objective function for the framework is given as Eq. 33,

$$\mathrm{argmin}_{\{u_t^i\},\theta} \sum_{t=1}^{T} L_{\mathrm{sh}}^t + \beta_0 L_{\mathrm{tr}}^t + \beta_1 L_{\mathrm{smooth}}^t \tag{33}$$

where u_t^i is the embedding vector of v_i at time t, β_0 and β_1 are the hyper parameters, θ is the social strategy parameter, and $L_{\mathrm{sh}}^t, L_{\mathrm{tr}}^t, L_{\mathrm{smooth}}^t$ are the loss-based functions corresponding to social homophily, triadic closure and temporal smoothness, respectively. The goal of this objective function is to learn a representation for dynamic networks by preserving the structural information and evolution patterns. In particular, a uniform framework that quantifies the probability of an open triad developing into a closed triad and learns the embedding vectors of each vertex at different time steps jointly was introduced here. Their experimental result shows that the framework gives a precision score of 0.968 for mobile dataset. Nguyen et al. [69] introduced another framework for learning time respecting embeddings from continuous time dynamic networks. In contrast to the discrete snapshots of the network at regular intervals of time, this method takes the finest timescale at the level of seconds and milliseconds as input. At the finest granularity, the graph is represented as $G = (V, E_T, T)$, where V is the set of vertices, E_T is the set of temporal edges between vertices and T is a function that maps each edge to a corresponding time stamp. Moreover, this framework incorporates temporal dependencies into account and learns more meaningful and accurate temporal embeddings by searching over the space of temporal random walks that obey time. A temporal random walk is similar to a random walk in static graphs with an additional constraint that walk respects time. The algorithm for learning time-preserving embedding generalizes SkipGram architecture. Given a temporal walk S_t, the model takes temporal aspect of the dynamic networks into account and learns the embedding by optimizing the equation in Eq. 34,

$$\max_f \log \Pr(W_T = \{v_{i-w}, \dots, v_{i+w}\} \backslash v_i | f(v_i)) \tag{34}$$

where $f : V \rightarrow R^D$ is the node embedding function, w is the context window size for optimization and W_T is an arbitrary temporal context window $W_T \subseteq S_t$. The model is under the assumption of conditional independence between the nodes of a temporal context window when observed with respect to the source node v_i. Their result shows that this method gives an AUC score of 0.671 for the Enron dataset. Hisano et al. presented another semi-supervised graph embedding approach called Semi-Graph [37] for dynamic link prediction. This is performed by defining the loss function as a weighted sum of the supervised loss from the past dynamics and the unsupervised loss of predicting the neighborhood context in the current network. The learned embedding reflects information from both the temporal and cross-sectional network structures. For a set of links (j, k), the loss function is given by Eq. 35, where Ne(j) is the set of all edges that did not form links in the past formation networks, v_{fj} is a complex vector representation for node j, v' denotes conjugate of v, Re() is a function that keeps only the real part of a complex function and W_f is a diagonal complex-valued matrix defining the scaling of bias. The loss function for this framework is expressed for both link formation and dissolution processes.

$$\sum_{j,k\epsilon(j,k)} \log P(k|j) = \sum_{j,k\epsilon(j,k)} \left(\mathrm{Re}\left({v'_{fj}}^T W_f v_{fk} \right) - \log \sum_{k'\epsilon Ne(j)} \exp\left(\mathrm{Re}\left({v'_{fj}}^T W_f v_{fk'} \right) \right) \right). \tag{35}$$

DyLink2Vec [80] provides an effective feature representation framework for link prediction in dynamic networks. For each pair of nodes (u, v) in a snapshot network, the model learns a metric embedding $\alpha^{uv} \epsilon R^d$, where d represents the dimension of the embedding metric. In contrast to the language modeling techniques for learning network representations, this framework utilizes a compression reconstruction framework which preserves higher-order neighborhood and links history patterns of node pairs in its latent space representation. They modeled the metric embedding task as an optimal coding problem where the objective is to minimize the reconstruction error and solved the optimization task by solving the gradient descent method. The optimization problem is solved by two operations: compression and reconstruction using the back-propagation approach. Metric embedding is obtained from the objective function given by Eq. 36,

$$\alpha^{uv} = f\left(W^{(c)} e^{uv} + b^{(c)} \right) \tag{36}$$

where $W^{(c)}$ and $b^{(c)}$ are the weight matrix and bias of compression operation, respectively. Once the embedding metric for each node pair is obtained, a supervised classification model is trained using the feature representations to perform the link prediction task efficiently. Their evaluation result shows that this framework gives a

PRAUC score of 0.50 on the Enron dataset. Goyal et al. introduced a framework called DynGEM [32], where they deployed the idea of deep autoencoders to learn embeddings effectively. Deep autoencoder leverages the recent advancements in DL to generate highly non-linear stable embeddings of dynamic graphs. This framework deployed a dynamically expanding deep autoencoder to map the input data into a highly non-linear latent space and capture trends in connectivity of graph snapshots at any time period. This is a semi-supervised model which optimizes the objective function corresponding to the first- and second-order proximities. For any pair of nodes v_i and v_j, from snapshot G_T, the model takes their neighborhoods as input and passes them through the autoencoder to get d-dimensional embedding vectors at the output of the encoder. Decoder reconstructs the neighborhoods from the embeddings. The model minimizes a weighted combination of three objectives to learn the parameters as given by Eq. 37,

$$L_{\mathrm{net}} = L_{\mathrm{glob}} + \alpha L_{\mathrm{loc}} + v_1 L_1 + v_2 L_2 \tag{37}$$

where L_{loc} and L_{glob} correspond to the first- and second-order proximities of the network, α, v_1 and v_2 are the hyper parameters and L_1, L_2 are the regularizers. Their experimental result shows that this framework gives a MAP score of 0.084 for the Enron dataset. Deep dynamic network embedding (DDNE) [51] is another autoencoder-based framework that utilized the historical information obtained from the network snapshots at past time stamps to learn latent representations of the future network. Moreover, this framework utilized gated recurrent unit (GRU) that makes each recurrent unit to adaptively capture the dependencies of different timescales. The objective function for embedding is designed to incorporate both network internal and dynamic transition structures as given by Eq. 38,

$$Y_i^{(m)} = \sigma\left(W^{(m)} y_i^{m-1} + b^{(m)}\right) \tag{38}$$

where $m = 2, \ldots, M$ are the hidden layers. Since GRU is known to learn problems with long range temporal dependencies and fast convergence, it makes the encoder efficient to capture the dynamic patterns by mapping the input to a highly non-linear latent space. Interaction proximity was introduced to exploit the pairwise closeness between vertices which tries to map two frequently connected vertices into similar latent space. The window is shifted one step toward the future to obtain a fixed observation which contains the previous $N - 1$ snapshots and the current snapshot. The known observations fixed as history are fed into the deep model and a forward inference was performed to get embeddings corresponding to future time $T + 1$. The decoder reconstructs the graph corresponding to time $T + 1$ from which the temporal link prediction task is performed. Their result shows that this model gives an AUC score of 0.96 for the Enron dataset (Table 6).

Restricted Boltzmann Machine (RBM) Based

Restricted Boltzmann machine is a special case of Markov random field which consists of two layers: a visible layer (*V*) and a hidden layer (*H*). Very few works in temporal link prediction have concentrated on modeling an RBM which can learn the probability distribution over $(V, H)\epsilon 0, 1^{N_V} \times 0, 1^{N_H}$, where N_V and N_H are the dimensions of V and H, respectively. Figure 9 depicts the basic idea of RBM-based temporal link prediction. The joint probability distribution in RBM is given by Eq. 39,

$$\Pr(V, H) = \frac{\exp\left(V^T WH + a^T V + b^T H\right)}{\sum_{V,H} \exp\left(V^T WH + a^T V + b^T H\right)}, \tag{39}$$

where $w\epsilon R^{N_V \times N_H}$ is the weight matrix between V and H, and a and b are biases for V and H, respectively. Unlike Boltzman machine (BM), there are no connections

Table 6 Comparison of embedding-based temporal link prediction

Method	Objective function	Description
Triadic closure [111]	Equation 33	Representation of learning approach that preserves both structural information and evolution patterns of the network by modeling triadic closure
CTDNE [69]	Equation 34	Time-dependent network representation that captures important temporal dependencies of continuous-time dynamic networks
SemiGraph [69]	Equation 35	Discrete semi-supervised graph embedding approach that reflects both temporal and cross-sectional network structures
DynLink2Vec [80]	Equation 36	Learns feature embedding of node pairs by considering both network topology and link history for temporal link prediction
DynGEM [32]	Equation 37	Recent advancements in deep autoencoders to learn graph embeddings
DDNE [51]	Equation 38	Proposed a deep architecture which can leverage historical linkage to make an embedding for new links

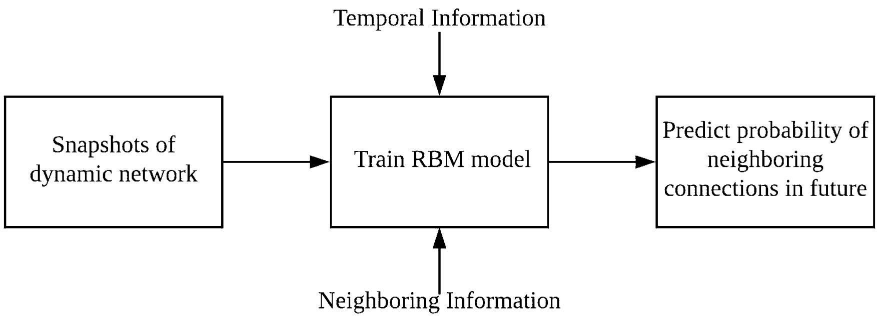

Fig. 9 RBM-based temporal link prediction

Table 7 Comparison of RBM-based temporal link prediction

Method	Objective function	Description
RBM model [108]	Equation 39	Incorporates historical linkage and neighboring information into the RBM model by adding temporal and neighboring connections between hidden and visible layers
ctRBM [52]	Equation 39	Incorporates historical linkage and neighboring information into the RBM model and neighbor influence clustering was introduced to improve the efficiency

between nodes in each layer in RBM. The goal is to minimize the distance between V and V', where V' represents the estimation of the model. Figure 9 depicts the general strategy followed by RBM-based temporal link prediction techniques. The general steps followed by RBM-based temporal link prediction models are as follows:

1. Dynamic networks are represented as a sequence of snapshots at regular interval of time.
2. RBM model which incorporates temporal and neighboring information is trained based on a sequence of observation of the dynamic network structures.
3. The probability of neighbor connections are calculated to perform temporal link prediction.

Traditional RBMs can only model static frames of data. Yu et al. [108] modeled RBM by adding temporal and neighboring connections between hidden and visible layers. They suggested that the temporal and neighboring connections can be added on each node in a dynamic network. Moreover, they introduced a new generative model in which the weights for temporal connections can model global non-linear temporal structure and weights for neighboring connection signify local structures. This model was under the assumption that a node's linkage behavior will affect the future linkage status and predicted the probability of neighboring connections in future time. Given the historical linkage information, the neighbor influence of node i is defined as in Eq. 40,

$$P_i^t = 1 - \left(1 - \lambda_i\right)^{\sum_{j=1}^{N} inf(i,j,t)} \tag{40}$$

where $\inf(i, j, t)$ is the influence of node j on i at time t, and λ_i is the infection rate of node i. Training approach was based on the contrastive divergence algorithm. Their experimental result shows that the RBM model gives the least scores for predicted errors than state-of-the-art methods. A similar framework called conditional temporal restricted Boltzmann machine (ctRBM) [52] inherits the advantages from RBM family and predicts the links based on individual transition variance as well as influence introduced by local neighbors. The model makes prediction based on the current time window and the expectation of local neighbor prediction. They proposed a Neighbor Influence Clustering algorithm without reducing the model capacity for further improving the computational cost. A ctRBM denoted as M_p is trained for each node p based on contrastive divergence algorithm and, finally, a collection of ctRBMs are obtained for all nodes denoted as M. The final prediction was obtained by collecting the results of each M_p. Although M makes a high space complexity, it is easily deployed on a large distributed platform. Their evaluation result shows that the framework gives an AUC score of 0.788 for the exposure dynamic network dataset (Table 7).

Other DL Approaches

Some other DL methods were also deployed for temporal link prediction. Link prediction model with spatial and temporal consistency (LIST) [107] predicts the links in a sequence of networks over time by incorporating network propagation and temporal matrix factorization techniques simultaneously. This is guaranteed by the joint optimization of the network propagation loss and the temporal network reconstruction error. The prediction result is given by Eq. 41,

$$A(T+1) = \left(\sum_{i=0}^{d} W^{(i)}(T+1)^i\right)\left(\sum_{i=0}^{d} W^{(i)}(T+1)^i\right)^T \tag{41}$$

where $A(T+1)$ is the adjacency matrix at time $T+1$, $\{W^{(i)}\}_{i=1}^{d}$ are the factor matrices and d is the order. LIST utilizes a user-defined sliding window to learn parameters and therefore supports streaming link prediction as well. This model characterizes the network dynamics as a function of time, which integrates the spatial topology of the network at each time stamp and the temporal network evolution. This model has the advantage of generality in addressing various temporal applications like temporal network compression and expanding community detection. Their result shows that this framework gives an RMSE score of 0.0009 on the DBLP dataset.

GRATFEL [79] method was introduced for temporal link prediction, where they designed a novel graphlet transition-based feature representation of the node pair instances. The unsupervised feature learning task is modeled as an optimal coding function h constituted by compression and reconstruction operations. The objective is to minimize the reconstruction error and the optimization task is solved by gradient descent method as given by Eq. 42, where e is the graphlet transition feature vector, $W^{(c)}$ and $b^{(c)}$ represent the weight matrix and bias for compression, respectively, $W^{(r)}$ and $b^{(r)}$ represent the weight matrix and bias for reconstruction respectively, β is the output of reconstruction, λ is the regularization parameter, and E represents the union of training and prediction datasets and $\| \, . \, \|_F$ denotes the Frobenius norm of the matrix.

$$J(W,b) = \frac{1}{2m}\sum_{e \epsilon E}\left(\frac{1}{2}\parallel \beta - e \parallel^2\right) + \frac{\lambda}{2}\left(\parallel W^{(c)} \parallel_F^2 + \parallel W^{(r)} \parallel_F^2\right). \tag{42}$$

Based on the feature representations for time snapshots $[1, t-1]$, the ground truth y' is constructed from G_t. Links in G_{t+1} is predicted using a supervised classification model. Their evaluation results show that GRAFTEL-based temporal link prediction

Table 8 Comparison of other DL techniques for temporal link prediction

Method	Objective function	Description
LIST [107]	Equation 41	Characterizes network dynamics as a function of time which integrates spatial topology of the network at each time stamp and temporal network evolution
GRATFEL [79]	Equation 42	Temporal link prediction is performed using the extracted feature representations of graphlet transition events
SLWE [18]	Equation 43	Builds feature vector for deep neural network by the application of weak estimators in addition to the traditional similarity metrics

gives an AUC score of 0.948 on the Enron dataset. In addition to the utilization of traditional similarity metrics, [18] introduced a novel application of weak estimators that inexpensively build an effective feature vector for a deep neural network. Weak estimators have been widely used in a variety of machine learning algorithms to improve the model accuracy, owing to their capacity to estimate the changing probabilities in dynamic systems. They introduced stochastic learning weak estimators (SLWEs) to characterize the changing likelihood of the existence of a link over time. Their evaluation results show that the framework gives an AUC score of 0.806 on MathOverflow dataset. The goal of SLWE was to estimate class probabilities in a dynamic system by applying stochastic learning principles as new instances are observed. The probability that a link is present at time $t + 1$ is given by Eq. 43,

$$p(t + 1) = \lambda \cdot p(t) + x(t)(1 - \lambda), \tag{43}$$

where λ is a learning coefficient and $x(t)$ denotes the existence of a link at time t. Finally, they built and trained a DL network for link prediction where the training process is treated as a supervised classification problem (Table 8).

Other Techniques

Subgraph evolution-based temporal link prediction [41] defined triad transition matrix (TTM) which contains the probabilities of transitions between triads found in the network. The idea behind TTM is to use the data from the history of the network to derive the probabilities of transitions between triads of chosen network nodes. Oyama et al. [71] introduced a new link prediction problem called cross-temporal link prediction problem, where links among the nodes in different time periods were observed. They introduced a dimension reduction approach to this problem, where the data objects in different time periods are mapped into a common low-dimensional space and links are identified based on the distance between the data objects. The majority of the link prediction algorithms focus on the application of proximity measures. A limitation usually observed in all those works were that only the current state of the network is used to compute the proximity scores without considering the temporal information. A new proximity measure for link prediction was introduced in [88] based on the concept of temporal events. An event-based score was computed and updated along time by rewarding the temporal events observed between the pair of nodes under the analysis of their neighborhoods. Link prediction was performed by ranking the node pair scores. Based on the concept of temporal events, the evolutionary approach in [11] suggested that incorporation of temporal features and node attributes can improve the link prediction task. This model provided an approach for predicting future links by adopting the covariance matrix adaptation evolution strategy (CMA-ES) [35] to optimize weights which are used in a linear combination of 16 neighborhoods and node similarity indices.

Graph embedding framework for temporal link prediction was introduced in [27]. The framework consists of four major modules: graph embedding, manifold alignment, trajectory prediction and graph reconstruction. For each graph snapshot G_t, multi-dimensional scaling (MDS) [105] or other graph embedding algorithms are applied to map each graph into a Euclidean space. Manifold alignment [93] aligns each G_t in the embedded space with different choices of alignment algorithms. Trajectory prediction step in the algorithm performs trajectory estimation and prediction for each graph vertex in the embedded space using the ARMA model. The graph embedding X_{T+1} is optimally predicted after this step from which the graph G_{T+1} is reconstructed using pairwise distances between the embeddings. Another work concentrated on feature engineering for supervised temporal link prediction [67], which is a two-step strategy with a novel yet simple feature construction approach using a combination of domain and topological attributes of the graph in the first phase. Based on the time-feature matrix, unconstrained edge selection is performed to identify the potential candidates for prediction by any generic two-class learner. Another approach utilized three metrics, time-varied weight, change degree of common neighbor and the intimacy between common neighbors [103]. Moreover, they redefined the common neighbors by finding them within two hops. The goal of scalable temporal latent space model [114] is to predict links over time based on a sequence of graph snapshots. This model was under the assumption that each user lies in an unobserved latent space and interactions are more likely to occur between similar users in the latent space. They presented a global optimization algorithm which effectively infers the temporal latent space. Later, [14] introduced a directional link prediction measure by extending neighbor-based measures as directional pattern. They also considered weight and time information of links which are effective to improve the link prediction accuracy. The two dynamic algorithms DLP-ILS and DLP-IRA proposed in [110] assumed that two nodes with the shorter distance in latent space are more likely to generate an edge in future. The latter adopted dynamic resource allocation for common neighbor metrics to reduce calculation time. These algorithms have significantly improved accuracy of link prediction than baseline methods. A new weighting criteria, where unsupervised link prediction is performed by combining the topological, temporal and contextual data aspects simultaneously, was introduced in [66]. This general weighting model can be used to configure different weighting criteria based on the three aspects of dynamic networks. Here, the idea is to combine profile similarity between nodes with frequency and time of their interactions so that connected nodes that interacted frequently and share similar profiles have higher link strength. The problem of streaming link prediction for dynamic attributed networks was studied in [50]. They introduced a framework SLIDE that maintains and updates a sketching matrix which contains all the observed data and utilized the sketching matrix to infer new links between the nodes. Considering the evolutionary aspect of communities in the network, [20] build a dynamic similarity metric or dynamic features to measure the similarity or proximity between node pairs. A supervised link prediction model was built, where these features were used as instance features. Two stochastic block models, namely tensorial model and bipartite model, was introduced in [91], where the former uses nodes as its fundamental unit and the latter focuses on links.

Dataset Description

The majority of the works in the area of temporal link prediction utilize real-world networks from various domains such as co-author networks, citation networks, social networks, human contact information networks and others. Such datasets cover relationships between the entities for a certain period of time. To capture their dynamic behavior, snapshots of these datasets are taken at regular intervals of time. In this section, we briefly describe various datasets used for temporal link prediction.

Co-author and Citation Networks

Co-author networks are the most widely used category of real-world dataset that consists of authors and their collaborations. Each node in the network represents an author, and the edge between two authors represents a common publication. Citation networks contain documents that reference each other. Each node in such network represents a paper and an edge represents a citation. Some of the frequently used co-author and citation network datasets for temporal link prediction are described below.

1. DBLP: The most widely used dataset for temporal link prediction. This is a bibiliographic database for computer science which consists of 1,314,050 nodes and 18,986,618 edges.
2. Condensed matter (Cond-mat): This is an authorship network extracted from arXiv archive which consists of 38,741 nodes and 58,595 links for a period from January 1995 to September 1999.
3. General relativity and quantum cosmology (GR-QC): This is a collaboration network from the e-print arXiv that covers scientific collaboration between authors who submitted papers to the GR-QC category. Nodes represent authors and edges between them represents collaboration. The dataset consists of papers for a period from January 1993 to April 2003 with 542 nodes and 14,496 edges.
4. High energy physics-phenomonology (Hep-Ph): This is a collaboration graph of authors of scientific papers from Hep-Ph section of arXiv archive. It consists of 28,093 nodes and 4,596,803 edges. The data covers papers in the period from January 1993 to April 2003. Citation network of Hep-ph is a collaboration graph

of publications in the arXiv's Hep-Ph section that consists of 34,546 vertices and 421,578 edges.
5. High energy physics-theory (Hep-Th): This is a collaboration graph consisting authors of scientific papers from Hep-Th section of the arXiv archive. The dataset consists of 22,908 nodes and 26,73,133 links.
6. Academic: This is a co-author network extracted from the academic network of AMiner. It consists of 51,060 nodes that represent authors and 7,94,552 edges which represent co-author relationships.
7. Astro physics (Astro-Ph): This co-author dataset contains authors of scientific papers from arXiv astro-physics section. It consists of 18,771 nodes and 198,050 edges.
8. Hieph-collab: This dataset represents co-author network for a period of 1992 to 2003. It consists of 8381 nodes representing authors and 40,736 edges representing co-author relationship.
9. Citeseer: This is a citation network extracted from Citeseer Digital Library. The dataset consists of 384,413 vertices and 17,51,463 edges.
10. Hieph-cite: This is a citation network based on High Energy Physics publications submitted to arXiv. It consists of 8249 nodes and 3,35,028 edges.

Communication Networks

Communication network consists of nodes that represents person and edges that indicates individual communication between two persons. The most widely used communication network datasets for temporal link prediction are described below.

1. Facebook: This is a directed network that contains wall posts of users on Facebook. The nodes of the network are Facebook users, and each directed edge represents one post, linking the users writing a post to the users whose wall the post is written on. The network consists of 46,952 vertices and 876,993 edges.
2. Enron email: This dataset consists of emails between the employees in Enron Inc. from January 1999 to July 2002. Each node in the network represents a user, and a link represents email communication between them. The network consists of 150 users and 2609 edges.
3. Digg: This is a reply network of the social news website Digg. Each node represents a user of the website and each directed edge indicates that a user replied to another user. It consists of 30,398 nodes and 87,627 edges.
4. Nodobo: This network contains cell phone calls between high school students. Here, nodes represent cell phone users and edges indicates the cell phone communication between users.The dataset consists of 395 nodes and 453 edges.
5. Radoslaw (Manufacturing email): This network represents the email communication between employees in a mid-sized manufacturing company. Nodes in the network represent employees and edges between them are individual emails. The dataset covers a period from January 2010 to October 2010 and consists of 167 vertices and 82,927 edges.

6. Eu-Core: This is an email dataset that contains over 3,00,000 emails sent among faculty at European University, with an edge representing an email sent from one user to another at a specific time. This network consists of 986 nodes and 3,32,334 edges.
7. CollegeMsg: This dataset represents the student interactions in a social network operating at the University of California, Irvine, over half a year. Each node in the network represents a student and edge represents a private message sent from one student to another at a specific time. It consists of 1899 nodes and 59,835 edges.
8. Mobile: This is a mobile network provided by China Telecom and consists of more than 2 million call logs between 3,40,751 users. The dataset considers users as nodes and there exists an edge between two users if they have called each other.
9. MathOverflow: This network describes interactions between users of the popular mathematics help forum mathoverflow.net recorded over the span of more than 6 years. It consists of 24,818 nodes and 5,06,550 edges.

Social Networks

Social networks represent relationships between individual persons in online social networking platforms. Nodes in such network represent individual person and link represents relationships between them.

1. Twitter: This is an online social network where each node is a Twitter user and an edge from user A to user B indicates that user A has mentioned user B in a tweet using '@username' syntax. It consists of 29,19,613 nodes and 1,28,87,063 edges.
2. Flickr: This social network consists of nodes that represents Flickr users and links denote their friendship relationships. The dataset consists of 2,302,925 vertices and 33,140,017 edges.
3. Epinions: This dataset represents trust and distrust network of Epinions, which is an online product rating site. Each node in the network represents individual users and link represents trust or distrust. The dataset consists of 131,828 vertices and 841,372 edges.
4. Youtube: This dataset represents social network of Youtube users and their friendship connections. The network consists of 3,223,589 vertices and 9,375,374 edges.

Human Contact Network

Human contact network dataset describes some kind of actual contact between two persons such as talking with each other, spending time together or face-to-face contacts. Each node in those networks indicates a person and an edge between two persons indicates a contact between them. A few such networks deployed for temporal link prediction are described below.

1. Reality mining: This network contains human contact data among 100 students of Mansachusetts Institute of Technology (MIT). Each node represents a person and link indicates that the corresponding persons had a physical contact. The dataset consists of 96 nodes and 10,86,404 edges.
2. Haggle: This network describes human contact information where contacts between people are measured by some wireless devices. Nodes represent users and links between them indicates a contact. The dataset consists of 274 nodes and 28,244 edges.
3. Infectious SocioPatterns: This is a temporal network which describes face-to-face behavior of people during the exhibition Infectious: stay away in 2009. Nodes represent exhibition visitors and edges represent face-to-face contacts that were active for at least 20 s. The dataset consists of 410 vertices and 17,298 edges.
4. Hypertext 2009: This network represents face-to-face contacts of the attendees of the ACM Hypertext 2009 Conference. Nodes in the network represent conference visitors and edge represents face-to-face contacts that were active for at least 20 s. It consists of 113 vertices and 20,818 edges for a period of June 2009 to July 2009.

Future Directions

In this section, we summarize future directions in the field of temporal link prediction. Matrix factorization techniques incorporated the linear characteristics of the dynamic networks. To improve the prediction accuracy, further improvements can be done to consider the non-linear varying complex patterns of the temporal networks. The majority of the probabilistic methods deploy maximum likelihood approaches. In contrast to the time-consuming maximum likelihood approaches, probabilistic multilayer model [104] can be extended to incorporate undirected dynamic networks. Rather than using single linear system models like FIR filters, it would be better to extend the spectral clustering-based techniques to deploy distributed filters to capture the latent features for different users. Time series can effectively capture the evolving patterns of the networks. However, computing a time series score which preserves all the network properties remains a challenging task. Moreover, traditional time series forecasting techniques fail to capture the non-linear varying temporal patterns of the networks. Therefore, an efficient time series forecasting method which predicts the future score by incorporating the non-linear varying temporal patterns of the networks can improve the accuracy of temporal link prediction. Embedding techniques have shown their outstanding performance in mapping the networks into a low-dimensional space by preserving all the network properties. Time series score constructed from the embeddings would efficiently capture the evolving nature of the networks. Moreover, deploying efficient time series forecasting models on this time series would improve the accuracy of temporal link prediction.

Temporal link prediction techniques discussed in this paper are under the assumption that the nodes in the network remain same over all the time periods. However, in real-world heterogeneous networks, both nodes and edges are added or removed over time. How to extend temporal link prediction techniques to incorporate heterogeneous networks remains an open question. Moreover, a vast majority of the the real-world networks are associated with a rich set of node attributes and their attribute values also vary. Extending temporal link prediction to incorporate attributed networks will be a promising task for the future. Further improvements can be done in existing systems to reduce the computational complexity and thereby improve the performance.

Conclusions

In this work, we present a comprehensive survey of the literature in the field of temporal link prediction. We familiarize the temporal aspect of the networks and provide a formal definition for temporal link prediction problem. Based on various existing techniques, we propose a taxonomy for temporal link prediction methods, which consists of seven major categories: matrix factorization, probabilistic models, spectral clustering, time series, DL, and other techniques. We briefly discuss the approaches that are presented by the existing techniques to capture the temporal aspect of the networks and to perform temporal link prediction efficiently. Finally, we present the prevailing challenges in the existing techniques and point out the promising future directions in the field of temporal link prediction.

Acknowledgements The infrastructure used for conducting this study is funded by FIST which is sanctioned by DST to NSS College of Engineering, Palakkad. We would like to express our gratitude to the Department of Computer Science and Engineering, NSS College of Engineering, Palakkad, for providing the required infrastructure.

References

1. Abdi, H.: The eigen-decomposition: Eigenvalues and eigenvectors. In: Encyclopedia of Measurement and Statistics, pp. 304–308 (2007)
2. Adamic, L.A., Adar, E.: Friends and neighbors on the web. Soc. Netw. **25**(3), 211–230 (2003)
3. Ahmed, N.M., Chen, L.: New approaches for link prediction in temporal social networks. Comput. Model. New Technol. **18**, 87–94 (2014)
4. Ahmed, N.M., Chen, L.: An efficient algorithm for link prediction in temporal uncertain social networks. Inf. Sci. **331**, 120–136 (2016)
5. Ahmed, N.M., Chen, L., Wang, Y., Li, B., Li, Y., Liu, W.: Sampling-based algorithm for link prediction in temporal networks. Inf. Sci. **374**, 1–14 (2016)
6. Ahmed, N.M., Chen, L., Wang, Y., Li, B., Li, Y., Liu, W.: Deepeye: link prediction in dynamic networks based on non-negative matrix factorization. Big Data Min. Anal. **1**(1), 19–33 (2018)
7. Aiello, L.M., Barrat, A., Schifanella, R., Cattuto, C., Markines, B., Menczer, F.: Friendship prediction and homophily in social media. ACM Trans. Web (TWEB) **6**(2), 9 (2012)
8. Al Hasan, M., Chaoji, V., Salem, S., Zaki, M.: Link prediction using supervised learning. In: SDM06: Workshop on Link Analysis, Counter-terrorism and Security (2006)
9. Al Hasan, M., Zaki, M.J.: A survey of link prediction in social networks. In: Social Network Data Analytics, pp. 243–275. Springer, Boston, MA (2011)
10. Backstrom, L., Leskovec, J.: Supervised random walks: predicting and recommending links in social networks. In: Proceedings of the 4th ACM International Conference on Web Search and Data Mining, pp. 635–644. ACM (2011)

11. Bliss, C.A., Frank, M.R., Danforth, C.M., Dodds, P.S.: An evolutionary algorithm approach to link prediction in dynamic social networks. J. Comput. Sci. **5**(5), 750–764 (2014)
12. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. **2008**(10), P10008 (2008)
13. Brockwell, P.J., Davis, R.A., Calder, M.V.: Introduction to Time Series and Forecasting, vol. 2. Springer, Cham (2002)
14. Bütün, E., Kaya, M., Alhajj, R.: Extension of neighbor-based link prediction methods for directed, weighted and temporal social networks. Inf. Sci. **463**, 152–165 (2018)
15. Casteigts, A., Flocchini, P., Quattrociocchi, W., Santoro, N.: Time-varying graphs and dynamic networks. Int. J. Parallel Emerg. Distrib. Syst. **27**(5), 387–408 (2012)
16. Chen, H.H., Gou, L., Zhang, X.L., Giles, C.L.: Discovering missing links in networks using vertex similarity measures. In: Proceedings of the 27th Annual ACM Symposium on Applied Computing, pp. 138–143 (2012)
17. Chib, S., Greenberg, E.: Understanding the metropolis-hastings algorithm. Am. Statist. **49**(4), 327–335 (1995)
18. Chiu, C., Zhan, J.: Deep learning for link prediction in dynamic networks using weak estimators. In: IEEE Access, pp. 35937–35945 (2018)
19. Cholette, P.A.: Prior information and ARIMA forecasting. J. Forecast. **1**(4), 375–383 (1982)
20. Choudhury, N., Uddin, S.: Evolutionary community mining for link prediction in dynamic networks. In: International Conference on Complex Networks and their Applications, pp. 127–138. Springer (2017)
21. Clauset, A., Moore, C., Newman, M.E.: Hierarchical structure and the prediction of missing links in networks. Nature **453**(7191), 98 (2008)
22. Das, S., Das, S.K.: A probabilistic link prediction model in time-varying social networks. In: 2017 IEEE International Conference on Communications (ICC), pp. 1–6. IEEE (2017)
23. Dong, L., Li, Y., Yin, H., Le, H., Rui, M.: The algorithm of link prediction on social network. Math. Probl. Eng. **2013** (2013)
24. Dunlavy, D.M., Kolda, T.G., Acar, E.: Temporal link prediction using matrix and tensor factorizations. ACM Trans. Knowl. Discov. Data (TKDD) **5**(2), 10 (2011)
25. Estrada, E., Hatano, N.: Communicability in complex networks. Phys. Rev. E **77**(3), 036111 (2008)
26. Faber, N.K.M., Bro, R., Hopke, P.K.: Recent developments in CANDECOMP/PARAFAC algorithms: a critical review. Chemom. Intell. Lab. Syst. **65**(1), 119–137 (2003)
27. Fang, C., Kohram, M., Meng, X., Ralescu, A.: Graph embedding framework for link prediction and vertex behavior modeling in temporal social networks. In: Proceedings of the SIGKDD Workshop on Social Network Mining and Analysis (2011)
28. Friedman, N., Getoor, L., Koller, D., Pfeffer, A.: Learning probabilistic relational models. In: IJCAI, vol. 99, pp. 1300–1309 (1999)
29. Gael, J.V., Teh, Y.W., Ghahramani, Z.: The infinite factorial hidden Markov model. In: Advances in Neural Information Processing Systems, pp. 1697–1704 (2009)
30. Gao, S., Denoyer, L., Gallinari, P.: Temporal link prediction by integrating content and structure information. In: Proceedings of the 20th ACM International Conference on Information and Knowledge Management, pp. 1169–1174. ACM (2011)
31. Golub, G.H., Reinsch, C.: Singular value decomposition and least squares solutions. In: Linear Algebra, pp. 134–151. Springer, Berlin, Heidelberg (1971)
32. Goyal, P., Kamra, N., He, X., Liu, Y.: Dyngem: deep embedding method for dynamic graphs (2018). arXiv:1805.11273
33. Guimerà, R., Sales-Pardo, M.: Missing and spurious interactions and the reconstruction of complex networks. Proc. Natl. Acad. Sci. **106**(52), 22073–22078 (2009)
34. Güneş, İ., Gündüz-Öğüdücü, Ş., Çataltepe, Z.: Link prediction using time series of neighborhood-based node similarity scores. Data Min. Knowl. Discov. **30**(1), 147–180 (2016)
35. Hansen, N., Ostermeier, A.: Completely derandomized self-adaptation in evolution strategies. Evolut. Comput. **9**(2), 159–195 (2001)
36. Heckerman, D., Meek, C., Koller, D.: Probabilistic entity-relationship models, PRMs, and plate models. In: Introduction to Statistical Relational Learning, pp. 201–238 (2007)
37. Hisano, R.: Semi-supervised graph embedding approach to dynamic link prediction. In: International Workshop on Complex Networks, pp. 109–121. Springer, Cham (2018)
38. Holme, P., Saramäki, J.: Temporal networks. Phys. Rep. **519**(3), 97–125 (2012)

39. Ibrahim, N.M.A., Chen, L.: Link prediction in dynamic social networks by integrating different types of information. Appl. Intell. **42**(4), 738–750 (2015)
40. Jaccard, P.: Étude comparative de la distribution florale dans une portion des Alpes et des jura. Bull. Soc. Vaudoise Sci. Nat. **37**, 547–579 (1901)
41. Juszczyszyn, K., Musial, K., Budka, M.: Link prediction based on subgraph evolution in dynamic social networks. In: 3rd IEEE International Conference on Privacy, Security, Risk and Trust and Third IEEE International Conference on Social Computing, pp. 27–34 (2011)
42. Kashima, H., Abe, N.: A parameterized probabilistic model of network evolution for supervised link prediction. In: 6th International Conference on Data Mining (ICDM'06), pp. 340–349. IEEE (2006)
43. Katz, L.: A new status index derived from sociometric analysis. Psychometrika **18**(1), 39–43 (1953)
44. Kim, M., Leskovec, J.: The network completion problem: inferring missing nodes and edges in networks. In: Proceedings of the 2011 SIAM International Conference on Data Mining, pp. 47–58. Society for Industrial and Applied Mathematics (2011)
45. Kossinets, G.: Effects of missing data in social networks. Soc. Netw. **28**(3), 247–268 (2006)
46. Kostakos, V.: Temporal graphs. Phys. A Stat. Mech. Appl. **388**(6), 1007–1023 (2009)
47. Kunegis, J., Lommatzsch, A.: Learning spectral graph transformations for link prediction. In: Proceedings of the 26th Annual International Conference on Machine Learning, pp. 561–568. ACM (2009)
48. Lakshmi, T.J., Bhavani, S.D.: Temporal probabilistic measure for link prediction in collaborative networks. Appl. Intell. **47**(1), 83–95 (2017)
49. Lei, K., Qin, M., Bai, B., Zhang, G.: Adaptive multiple non-negative matrix factorization for temporal link prediction in dynamic networks. In: Proceedings of the 2018 Workshop on Network Meets AI & ML, pp. 28–34. ACM (2018)
50. Li, J., Cheng, K., Wu, L., Liu, H.: Streaming link prediction on dynamic attributed networks. In: Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, pp. 369–377. ACM (2018)
51. Li, T., Zhang, J., Philip, S.Y., Zhang, Y., Yan, Y.: Deep dynamic network embedding for link prediction. IEEE Access **6**, 29219–29230 (2018)
52. Li, X., Du, N., Li, H., Li, K., Gao, J., Zhang, A.: A deep learning approach to link prediction in dynamic networks. In: Proceedings of the 2014 SIAM International Conference on Data Mining, pp. 289–297. SIAM (2014)
53. Liaw, A., Wiener, M.: Classification and regression by randomforest. R News **2**(3), 18–22 (2002)
54. Liben-Nowell, D., Kleinberg, J.: The link-prediction problem for social networks. J. Am. Soc. Inf. Sci. Technol. **58**(7), 1019–1031 (2007)
55. Lichtenwalter, R.N., Chawla, N.V.: Vertex collocation profiles: subgraph counting for link analysis and prediction. In: Proceedings of the 21st International Conference on World Wide Web, pp. 1019–1028. ACM (2012)
56. Liu, W., Lü, L.: Link prediction based on local random walk. EPL (Europhysics Letters) **89**(5), 58007 (2010)
57. Lü, L., Jin, C.H., Zhou, T.: Similarity index based on local paths for link prediction of complex networks. Phys. Rev. E **80**(4), 046122 (2009)
58. Lü, L., Medo, M., Yeung, C.H., Zhang, Y.C., Zhang, Z.K., Zhou, T.: Recommender systems. Phys. Rep. **519**(1), 1–49 (2012)
59. Lü, L., Pan, L., Zhou, T., Zhang, Y.C., Stanley, H.E.: Toward link predictability of complex networks. Proc. Natl. Acad. Sci. **112**(8), 2325–2330 (2015)
60. Lü, L., Zhou, T.: Link prediction in complex networks: a survey. Phys. A Stat. Mech. Appl. **390**(6), 1150–1170 (2011)
61. Ma, X., Sun, P., Qin, G.: Nonnegative matrix factorization algorithms for link prediction in temporal networks using graph communicability. Pattern Recognit. **71**, 361–374 (2017)
62. Ma, X., Sun, P., Wang, Y.: Graph regularized nonnegative matrix factorization for temporal link prediction in dynamic networks. Phys. A Stat. Mech. Appl. **496**, 121–136 (2018)
63. Meng, B., Ke, H., Yi, T.: Link prediction based on a semi-local similarity index. Chin. Phys. B **20**(12), 128902 (2011)
64. Menon, A.K., Elkan, C.: Link Prediction via Matrix Factorization. In: Gunopulos, D., Hofmann, T., Malerba, D., Vazirgiannis, M. (eds.) Machine Learning and Knowledge Discovery in Databases.

ECML PKDD 2011. Lecture Notes in Computer Science, vol. 6912, pp. 437–452. Springer, Berlin, Heidelberg (2011)
65. Moradabadi, B., Meybodi, M.R.: Link prediction based on temporal similarity metrics using continuous action set learning automata. Phys. A Stat. Mech. Appl. **460**, 361–373 (2016)
66. Muniz, C.P., Goldschmidt, R., Choren, R.: Combining contextual, temporal and topological information for unsupervised link prediction in social networks. Knowl. Based Syst. **156**, 129–137 (2018)
67. Narasimhan, J., Holder, L.: Feature engineering for supervised link prediction on dynamic social networks. In: Proceedings of the International Conference on Data Mining (DMIN), p. 1. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp) (2014)
68. Newman, M.E.: Clustering and preferential attachment in growing networks. Phys. Rev. E **64**(2), 025102 (2001)
69. Nguyen, G.H., Lee, J.B., Rossi, R.A., Ahmed, N.K., Koh, E., Kim, S.: Continuous-time dynamic network embeddings. In: Companion Proceedings of the Web Conference 2018, pp. 969–976. International World Wide Web Conferences Steering Committee (2018)
70. Ouzienko, V., Guo, Y., Obradovic, Z.: Prediction of attributes and links in temporal social networks. In: ECAI, pp. 1121–1122 (2010)
71. Oyama, S., Hayashi, K., Kashima, H.: Cross-temporal link prediction. In: 2011 IEEE 11th International Conference on Data Mining (ICDM), pp. 1188–1193. IEEE (2011)
72. Özcan, A., Öğüdücü, Ş.G.: Multivariate temporal link prediction in evolving social networks. In: IEEE/ACIS 14th International Conference on Computer and Information Science (ICIS), pp. 185–190. IEEE (2015)
73. Özcan, A., Öğüdücü, Ş.G.: Temporal link prediction using time series of quasi-local node similarity measures. In: 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), pp. 381–386. IEEE (2016)
74. Özcan, A., Öğüdücü, Ş.G.: Supervised temporal link prediction using time series of similarity measures. In: 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN), pp. 519–521. IEEE (2017)
75. Pavlov, M., Ichise, R.: Finding experts by link prediction in co-authorship networks. FEWS **290**, 42–55 (2007)
76. Pech, R., Hao, D., Lee, Y.L., Yuan, Y., Zhou, T.: Link prediction via linear optimization. Phys. A Stat. Mech. Appl. **528**, 121319 (2019)
77. Pech, R., Hao, D., Pan, L., Cheng, H., Zhou, T.: Link prediction via matrix completion. EPL (Europhysics Letters) **117**(3), 38002 (2017)
78. Popescul, A., Ungar, L.H.: Statistical relational learning for link prediction. In: IJCAI Workshop on Learning Statistical Models from Relational Data, vol. 2003. Citeseer (2003)
79. Rahman, M., Hasan, M.A.: Link prediction in dynamic networks using graphlet. In: Frasconi, P., Landwehr, N., Manco, G., Vreeken, J. (eds.) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2016. Lecture Notes in Computer Science, vol. 9851, pp. 394–409. Springer, Cham (2016)
80. Rahman, M., Saha, T.K., Hasan, M.A., Xu, K.S., Reddy, C.K.: Dylink2vec: effective feature representation for link prediction in dynamic networks (2018). arXiv:1804.05755
81. Ralescu, A., Kohram, M., et al.: Spectral regression with low-rank approximation for dynamic graph link prediction. IEEE Intell. Syst. **26**(4), 48–53 (2011)
82. Raymond, R., Kashima, H.: Fast and scalable algorithms for semi-supervised link prediction on static and dynamic graphs. In: Balcázar, J.L., Bonchi, F., Gionis, A., Sebag, M. (eds.) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2010. Lecture Notes in Computer Science, vol. 6323, pp. 131–147. Springer, Berlin, Heidelberg (2010)
83. Rossetti, G., Guidotti, R., Pennacchioli, D., Pedreschi, D., Giannotti, F.: Interaction prediction in dynamic networks exploiting community discovery. In: 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 553–558. IEEE (2015)
84. Sajadmanesh, S., Zhang, J., Rabiee, H.R.: NPGLM: a non-parametric method for temporal link prediction (2017). arXiv:1706.06783
85. Salton, G., McGill, M.J.: Introduction to Modern Information Retrieval. McGraw-Hill (1986)
86. Sarkar, P., Chakrabarti, D., Jordan, M.: Nonparametric link prediction in dynamic networks (2012). arXiv:1206.6394

87. Sarkar, P., Chakrabarti, D., Jordan, M.: Nonparametric link prediction in large scale dynamic networks. Electron. J. Stat. **8**(2), 2022–2065 (2014)
88. Soares, P.R., Prudêncio, R.B.: Proximity measures for link prediction based on temporal events. Expert Syst. Appl. **40**(16), 6652–6660 (2013)
89. Symeonidis, P., Mantas, N.: Spectral clustering for link prediction in social networks with positive and negative links. Soc. Netw. Anal. Min. **3**(4), 1433–1447 (2013)
90. Tang, J., Wu, S., Sun, J., Su, H.: Cross-domain collaboration recommendation. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1285–1293. ACM (2012)
91. Tarrés-Deulofeu, M., Godoy-Lorite, A., Guimerà, R., Sales-Pardo, M.: Tensorial and bipartite block models for link prediction in layered networks and temporal networks. Phys. Rev. E **99**(3), 032307 (2019)
92. Valverde-Rebaza, J., de Andrade Lopes, A.: Exploiting behaviors of communities of twitter users for link prediction. Soc. Netw. Anal. Min. **3**(4), 1063–1074 (2013)
93. Wang, C., Mahadevan, S.: Manifold alignment using procrustes analysis. In: Proceedings of the 25th International Conference on Machine Learning, pp. 1120–1127. ACM (2008)
94. Wang, C., Satuluri, V., Parthasarathy, S.: Local probabilistic models for link prediction. In: 7th IEEE International Conference on Data Mining (ICDM), pp. 322–331. IEEE (2007)
95. Wang, P., Xu, B., Wu, Y., Zhou, X.: Link prediction in social networks: the state-of-the-art. Sci. China Inf. Sci. **58**(1), 1–38 (2015)
96. Wang, T., He, X.S., Zhou, M.Y., Fu, Z.Q.: Link prediction in evolving networks based on popularity of nodes. Sci. Rep. **7**(1), 7147 (2017)
97. Wang, W.Q., Zhang, Q.M., Zhou, T.: Evaluating network models: a likelihood analysis. EPL (Europhysics Letters) **98**(2), 28004 (2012)
98. Wohlfarth, T., Ichise, R.: Semantic and Event-Based Approach for Link Prediction. In: Yamaguchi, T. (ed.) Practical Aspects of Knowledge Management. PAKM 2008. Lecture Notes in Computer Science, vol. 5345, pp. 50–61. Springer, Berlin, Heidelberg (2008)
99. Wu, T., Chang, C.S., Liao, W.: Tracking network evolution and their applications in structural network analysis. IEEE Trans. Knowl. Data Eng (2018)
100. Xie, H., Tang, H., Liao, Y.H.: Time series prediction based on NARX neural networks: an advanced approach. In: 2009 International Conference on Machine Learning and Cybernetics, vol. 3, pp. 1275–1279. IEEE (2009)
101. Yang, C., Liu, Z., Zhao, D., Sun, M., Chang, E.: Network representation learning with rich text information. In: 24th International Joint Conference on Artificial Intelligence, pp. 2111–2117 (2015)
102. Yang, X., Tian, Z., Cui, H., Zhang, Z.: Link prediction on evolving network using tensor-based node similarity. In: 2012 IEEE 2nd International Conference on Cloud Computing and Intelligent Systems (CCIS), vol. 1, pp. 154–158. IEEE (2012)
103. Yao, L., Wang, L., Pan, L., Yao, K.: Link prediction based on common-neighbors for dynamic social network. Proc. Comput. Sci. **83**, 82–89 (2016)
104. Yasami, Y., Safaei, F.: A novel multilayer model for missing link prediction and future link forecasting in dynamic complex networks. Phys. A Stat. Mech. Appl. **492**, 2166–2197 (2018)
105. Young, F.W., Hamer, R.M.: Theory and Applications of Multidimensional Scaling. Eribaum Associates, Hillsdale (1994)
106. Yu, K., Chu, W., Yu, S., Tresp, V., Xu, Z.: Stochastic relational models for discriminative link prediction. In: Advances in Neural Information Processing Systems, pp. 1553–1560 (2007)
107. Yu, W., Cheng, W., Aggarwal, C.C., Chen, H., Wang, W.: Link prediction with spatial and temporal consistency in dynamic networks. In: Proceedings of the International Joint Conference on Artificial Intelligence, pp. 3343–3349 (2017)
108. Yu, X., Chu, T.: Dynamic link prediction using restricted Boltzmann machine. In: Chinese Automation Congress (CAC), pp. 4089–4092. IEEE (2017)
109. Zhang, Q.M., Xu, X.K., Zhu, Y.X., Zhou, T.: Measuring multiple evolution mechanisms of complex networks. Sci. Rep. **5**, 10350 (2015)
110. Zhang, Z., Wen, J., Sun, L., Deng, Q., Su, S., Yao, P.: Efficient incremental dynamic link prediction algorithms in social network. Knowl. Based Syst. **132**, 226–235 (2017)
111. Zhou, L., Yang, Y., Ren, X., Wu, F., Zhuang, Y.: Dynamic network embedding by modeling triadic closure process. In: Thirty-Second AAAI Conference on Artificial Intelligence (2018)

112. Zhou, T., Lü, L., Zhang, Y.C.: Predicting missing links via local information. Eur. Phys. J. B **71**(4), 623–630 (2009)
113. Zhu, J., Hong, J., Hughes, J.G.: Using markov chains for link prediction in adaptive web sites. In: Bustard, D., Liu, W., Sterritt, R. (eds.) Soft-Ware 2002: Computing in an Imperfect World. Lecture Notes in Computer Science, vol. 2311, pp. 60–73. Springer, Berlin, Heidelberg (2002)
114. Zhu, L., Guo, D., Yin, J., Ver Steeg, G., Galstyan, A.: Scalable temporal latent space inference for link prediction in dynamic social networks. IEEE Trans. Knowl. Data Eng. **28**(10), 2765–2777 (2016)
115. Zhu, Y.X., Lü, L., Zhang, Q.M., Zhou, T.: Uncovering missing links with cold ends. Phys. A Stat. Mech. Appl. **391**(22), 5769–5778 (2012)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.