ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

Link prediction heuristics for temporal graph benchmark

Manuel Dileo and Matteo Zignani

Department of Computer Science, University of Milan, Italy
{manuel.dileo, matteo.zignani}@unimi.it

Abstract. Link prediction is one of the most well-known and studied problems in graph machine learning, successfully applied in different settings, such as predicting network evolution in online social networks, protein-to-protein interactions, or completing links in knowledge graphs. In recent years, we have witnessed several solutions based on deep learning methods for solving this task in the context of temporal networks. However, despite their effectiveness on static graphs, traditional heuristic-based approaches from network science research have never been considered potential benchmarks' baselines. For this reason, in this work, we tested four of the most well-known and simple heuristics for link prediction on the most adopted temporal graph benchmark (TGB). Our results show that simple link prediction heuristics can reach comparable results with state-of-the-art deep learning techniques and, thanks to their interpretability, give insights into the network being studied. We believe considering heuristic-based baselines will push the temporal graph learning community toward better models for link prediction.

1 Introduction and Background

Link prediction is one of the most well-known and studied problems in graph machine learning, successfully applied in different settings, such as predicting network evolution in online social networks [1], drug-drug interactions [2], or forecasting financial markets [3]. In recent years, in the context of link prediction on temporal networks, several solutions based on deep learning (DL) methods have been proposed for solving the task. Most of these techniques are covered in two recent surveys on the topic [4, 5]. In this research field, Temporal Graph Benchmark [6] - TGB - represents a widely adopted collection of benchmarks for link prediction on temporal networks. It makes available five datasets capturing different domains, spanning time, granularities, and sizes. Overall, the TGB leaderboard[1] shows that DL models exhibit high variability in the performance over the different datasets. Huang *et al.* [7] analyzes the temporal edge re-occurrence in the TGB datasets, highlighting that DL models may vary their performance due to the level of re-occurrence/novelty of the edges in the datasets, depending on how much DL models are based on memorization (e.g. TGN [8], DyRep [9]), or inductive reasoning (e.g. CAWN [10], GraphMixer [11]). However, due to the challenging interpretability of DL models, the underlying rationales behind this behavior remain largely opaque. In network science [12],

[1] `https://tgb.complexdatalab.com/docs/leader_linkprop/`, august 2024.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

before the advent of machine learning, heuristic-based methods were proposed to solve the link prediction problem [13]. On static networks, heuristics demonstrate their effectiveness even against feature learning methods such as Graph Neural Networks on OGB, the main graph machine learning benchmark[2]. However, traditional heuristic-based approaches from network science research have never been considered potential benchmarks' baselines for link prediction on temporal networks. This is an important issue to solve as, in general, there is a lack of proper baselines for link prediction in the temporal graph learning community [14]. Guided by these observations, we tested four of the most well-known and simple heuristics for link prediction on TGB. Our results show that simple link prediction heuristics can reach comparable results with state-of-the-art deep learning techniques. Moreover, thanks to heuristic easy interpretability, we give some insights into the networks being studied that go beyond the simple edge re-occurrence patterns. Overall, the main objective of this work is not to propose simple heuristics as new state-of-the-art solutions for machine learning on temporal networks, but to emphasize the importance of considering network science research when dealing with temporal graph learning (TGL), to push the community toward better models for link prediction.

2 Methodology

Given a graph $\mathcal{G} = (V, E)$ and a candidate pair (u, v) *s.t.* $u, v \in V$, the link prediction problem consists of finding a score function that, based on the information contained in $\mathcal{G}$, maps (u, v) in a value. The higher the value, the higher the probability that the link (u, v) exists. We selected four well-known heuristics for link prediction and developed new implementations using sparse matrices[3]. Specifically, we choose *Common Neighbors* (CN), the *Adamic Adar Index* (AA), the *Preferential Attachment* (PA)[13], and the *Resource Allocation Index* (RA) [15]. They can be defined as follows:

$$\mathrm{CN}(u, v) = |\Gamma(u) \cap \Gamma(v)| \tag{1}$$

$$\mathrm{RA}(u, v) = \sum_{w \in \Gamma(u) \cap \Gamma(v)} \frac{1}{|\Gamma(w)|} \tag{2}$$

$$\mathrm{AA}(u, v) = \sum_{w \in \Gamma(u) \cap \Gamma(v)} \frac{1}{\log |\Gamma(w)|} \tag{3}$$

$$\mathrm{PA}(u, v) = |\Gamma(u)||\Gamma(v)| \tag{4}$$

where $\Gamma(u)$ denotes the set of neighbours of u. *Common neighbors* is the simplest idea for giving a link prediction score, following the natural intuition that if two nodes interact with many nodes in common, they are likely to interact. AA and RA refine the simple counting of common neighbors by assigning more weights to

[2]`https://ogb.stanford.edu/docs/leader_linkprop/`, august 2024

[3]`https://docs.scipy.org/doc/scipy/reference/sparse.html`, august 2024.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

the lower connected neighbors, using two different scaling factors. PA is based on the homonymous mechanism to generate evolving scale-free networks [12] where the probability that a new link is connected to the node u is proportional to $|\Gamma(u)|$.

Link prediction heuristics can also be applied for future link prediction on temporal networks. Specifically, temporal networks can be modeled as timestamped edge streams consisting of triplets of source, destination, and timestamp $\mathcal{G} = \{(u, v, t_0), ..., (u, v, t_T)\}$ where the timestamps are ordered ($t_0 \leq t_1 \leq ... \leq t_T$) [4, 6]. Denoting with $\mathcal{G}_t$ the augmented graph of all the edges observed in the stream up to time t, the future link prediction problem consists of predicting edges based on $\mathcal{G}_t$ in a future timestamp $t^+ > t$. Hence, given a triple (u, v, t^+), its score can be computed as $H(u, v)$ using $\mathcal{G}_t$, where H is a heuristic. Timestamps on edges can also be leveraged to allow link prediction heuristics to consider only the most recent interactions. Specifically, given $\mathcal{G}_t$, a candidate triple (u, v, t^+) and a timestamp $t^- < t$, we can define the set of temporal neighbours of u as:

$$\Gamma_{[t^-,t]}(u) = \{(s, d, t') \mid (s = u \vee d = u) \wedge t^- \leq t' \leq t\} \tag{5}$$

and use the temporal neighborhood to compute the heuristic scores.

3 Evaluation and Discussion

Data. We evaluate link prediction heuristics on TGB [6], the most well-known and used collection of benchmarks for machine learning on temporal networks. Specifically, we choose the following link prediction datasets [4]:

- TGBL-WIKI: it stores the co-editing network on Wikipedia pages over one month. The network is a bipartite interaction network where editors and wiki pages are nodes, while one edge represents a given user who edits a page at a specific timestamp.
- TGBL-REVIEW: This dataset is an Amazon product review network from 1997 to 2018 where users rate different products. Therefore, the network is a bipartite network where both users and products are nodes and each edge represents a particular review from a user to a product at a given time.
- TGBL-COIN: This is a cryptocurrency transaction dataset based on the Stablecoin ERC20 transactions dataset. Each node is an address and each edge represents the transfer of funds from one address to another at a time. The network starts on April 1st, 2022, and ends on November 1st, 2022, and contains transaction data of 5 stablecoins and 1 wrapped token.
- TGBL-COMMENT: This dataset is a directed reply network of Reddit where users reply to each other's threads. Each node is a user and each

[4] TGBL-FLIGHT was excluded as it was not available through TGB Dataloaders.

ESANN 2024 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium) and online event, 9-11 October 2024, i6doc.com publ., ISBN 978-2-87587-090-2. Available from http://www.i6doc.com/en/.

interaction is a reply from one user to another. The network started in 2005 and ended in 2010.

To sum up, the four datasets range from different domains, spanning times, granularities, and sizes [6].

Experimental setting. We adopt the experimental setting presented in TGB [6]. All datasets are split chronologically into the training, validation, and test sets, respectively containing 70%, 15%, and 15% of all edges. Following the streaming setting [8], information from the test set is only employed for updating the memory module in TGL methods. However, no back-propagation or model update is possible based on the test set information. For the heuristics, no information contained in the test set is used to compute the scores. For each dataset, we consider all the deep learning (DL) models in the TGB leaderboard and made available by the TGB team. For TGBL-WIKI and TGBL-REVIEW, we consider the *Preferential attachment* heuristic only, as they are bipartite networks and nodes do not share neighbors. We denote with the subscript rec the heuristics that take into account only the recent edges. Specifically, the size of the time window for recent edges is set to the duration of the test split. The code to reproduce the results is available in a Github Repository [5].

Results. We report the performance of the DL models and heuristics methods for the four datasets in terms of Mean Reciprocal Rank (MRR) in Table 1 and 2. Results for DL models are taken from the TGB leaderboard. On TGBL-WIKI and TGBL-COIN, *Preferential Attachment* reaches the second best and the top performance overall, respectively. PA performances are achieved using zero learnable parameters and only a few minutes of computation, in contrast with the ones of DL models, achieved using millions of parameters and hours of computation. These results highlight that a simple mechanism such as *Preferential Attachment* is crucial in creating new interactions between nodes in these datasets. This pattern goes beyond the simple repetition/novelty of edges: the test set of TGBL-WIKI exhibits a lot of already-seen edges; in contrast, the test set of TGBL-COMMENT is characterized mainly by new interactions [6]. Hence, the memorization capability is not sufficient to discriminate TGL models. On TGBL-REVIEW, the performance of PA is almost equal to zero. The test set of this dataset is almost entirely composed of new edges. In this case, the drastically low result, in conjunction with the high level of new unseen edges, indicates that the new reviews are made mainly by new users on new products. Therefore, in this case, DL models may focus on giving representation for timestamped edges [11] instead of generating node-level embeddings ([8], [9]), avoiding giving too much importance to past node influence and their motif [10]. Lastly, on TGBL-COMMENT, only the heuristics based on CNs achieve comparable results with DL models. This result may be connected to results on the strength of weak ties [16] in online social networks. Therefore, TGL models may need to

[5] `https://github.com/manuel-dileo/lp-heuristics`, august 2024.

distinguish link structural roles [17] over time. In general, we believe the TGL community should focus on temporal network analysis of their benchmark to push towards better models for future link prediction.

Method	TGBL-WIKI	TGBL-REVIEW
CAWN [10]	**0.711 ± 0.006**	0.193 ± 0.001
TGN [8]	0.396 ± 0.060	0.349 ± 0.020
TCL [18]	0.207 ± 0.025	0.193 ± 0.009
TGAT [19]	0.141 ± 0.007	0.355 ± 0.012
GraphMixer [11]	0.118 ± 0.002	**0.521 ± 0.015**
DyRep [9]	0.050 ± 0.017	0.220 ± 0.030
PA	0.463 ± 0.000	0.019 ± 0.000
PA_{rec}	0.488 ± 0.000	0.020 ± 0.000

Table 1: Performance of deep learning models and heuristic techniques for temporal link prediction on test set of TGBL-WIKI and TGBL-Review, in terms of MRR.

Method	TGBL-COIN	TGBL-COMMENT
TGN	0.586 ± 0.037	**0.379 ± 0.021**
DyRep	0.452 ± 0.046	0.289 ± 0.033
PA	0.481 ± 0.000	0.09 ± 0.000
PA_{rec}	**0.584 ± 0.000**	0.124 ± 0.000
CN	0.408 ± 0.000	0.131 ± 0.000
CN_{rec}	0.322 ± 0.000	0.242 ± 0.000
AA	0.314 ± 0.000	0.130 ± 0.000
AA_{rec}	0.324 ± 0.000	0.245 ± 0.000
RA	0.327 ± 0.000	0.126 ± 0.000
RA_{rec}	0.334 ± 0.000	0.245 ± 0.000

Table 2: Performance of deep learning models and heuristic techniques for temporal link prediction on the test sets of TGBL-COIN and TGBL-COMMENT, in terms of MRR.

Future works In future works, we will extend the set of experiments by including other heuristics and datasets to give new insights into TGL models. Then, based on the observations, we will develop a new model that tries to combine key heuristic ideas with neural link prediction scores.

References

[1] Manuel Dileo, Matteo Zignani, and Sabrina Gaito. Temporal graph learning for dynamic link prediction with text in online social networks. *Machine Learning*, Nov 2023.

[2] Ke Han, Peigang Cao, Yu Wang, Fang Xie, Jiaqi Ma, Mengyao Yu, Jianchun Wang, Yaoqun Xu, Yu Zhang, and Jie Wan. A review of approaches for predicting Drug-Drug interactions based on machine learning. *Front Pharmacol*, 12:814858, January 2022.

[3] Manuel Dileo and Matteo Zignani. Discrete-time graph neural networks for transaction prediction in web3 social platforms. *Machine Learning*, Jun 2024.

[4] Seyed Mehran Kazemi, Rishab Goel, Kshitij Jain, Ivan Kobyzev, Akshay Sethi, Peter Forsyth, and Pascal Poupart. Representation learning for dynamic graphs: A survey. *J. Mach. Learn. Res.*, 21:70:1–70:73, 2020.

[5] Antonio Longa, Veronica Lachi, Gabriele Santin, Monica Bianchini, Bruno Lepri, Pietro Lio, Franco Scarselli, and Andrea Passerini. Graph neural networks for temporal graphs: State of the art, open challenges, and opportunities, 2023.

[6] Shenyang Huang, Farimah Poursafaei, Jacob Danovitch, Matthias Fey, Weihua Hu, Emanuele Rossi, Jure Leskovec, Michael M. Bronstein, Guillaume Rabusseau, and Reihaneh Rabbany. Temporal graph benchmark for machine learning on temporal graphs. In *NeurIPS*, 2023.

[7] Razieh Shirzadkhani, Shenyang Huang, Elahe Kooshafar, Reihaneh Rabbany, and Farimah Poursafaei. Temporal graph analysis with TGX. In *WSDM*, pages 1086–1089. ACM, 2024.

[8] Emanuele Rossi, Ben Chamberlain, Fabrizio Frasca, Davide Eynard, Federico Monti, and Michael Bronstein. Temporal graph networks for deep learning on dynamic graphs, 2020.

[9] Rakshit Trivedi, Mehrdad Farajtabar, Prasenjeet Biswal, and Hongyuan Zha. Dyrep: Learning representations over dynamic graphs. In *ICLR (Poster)*. OpenReview.net, 2019.

[10] Yanbang Wang, Yen-Yu Chang, Yunyu Liu, Jure Leskovec, and Pan Li. Inductive representation learning in temporal networks via causal anonymous walks. In *ICLR*. OpenReview.net, 2021.

[11] Weilin Cong, Si Zhang, Jian Kang, Baichuan Yuan, Hao Wu, Xin Zhou, Hanghang Tong, and Mehrdad Mahdavi. Do we really need complicated model architectures for temporal networks? In *The Eleventh International Conference on Learning Representations*, 2023.

[12] Albert-Laszlo Barabasi and Marton Posfai. *Network science*. Cambridge University Press, Cambridge, 2016.

[13] David Liben-Nowell and Jon Kleinberg. The link prediction problem for social networks. In *Proceedings of the Twelfth International Conference on Information and Knowledge Management*, CIKM '03, pages 556–559, New York, NY, USA, 2003. Association for Computing Machinery.

[14] Julia Gastinger, Christian Meilicke, Federico Errica, Timo Sztyler, Anett Schuelke, and Heiner Stuckenschmidt. History repeats itself: A baseline for temporal knowledge graph forecasting. 2024.

[15] Tao Zhou, Linyuan Lü, and Yi-Cheng Zhang. Predicting missing links via local information. *The European Physical Journal B*, 71(4):623–630, Oct 2009.

[16] Mark S. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, 1973.

[17] Muhan Zhang, Pan Li, Yinglong Xia, Kai Wang, and Long Jin. Revisiting graph neural networks for link prediction, 2021.

[18] Lu Wang, Xiaofu Chang, Shuang Li, Yunfei Chu, Hui Li, Wei Zhang, Xiaofeng He, Le Song, Jingren Zhou, and Hongxia Yang. TCL: transformer-based dynamic graph modelling via contrastive learning. *CoRR*, abs/2105.07944, 2021.

[19] Da Xu, Chuanwei Ruan, Evren Körpeoglu, Sushant Kumar, and Kannan Achan. Inductive representation learning on temporal graphs. In *ICLR*. OpenReview.net, 2020.