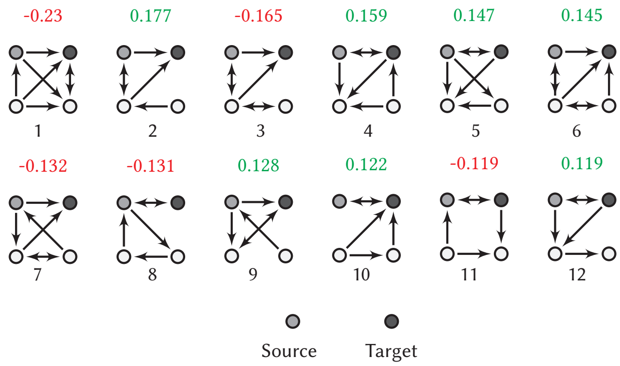

Figure 4: Top Cora SSTs with Bidirected Citations – (See Section 5.3.2) – SSTs 2, 3, and 6 indicate that if articles A and B mutually cite each other, A tends to cite whatever B cites *unless* another article C who bi-cites with B does not. SSTs 4 and 5 indicate that articles are more likely to bi-cite each other if they cite the same articles.

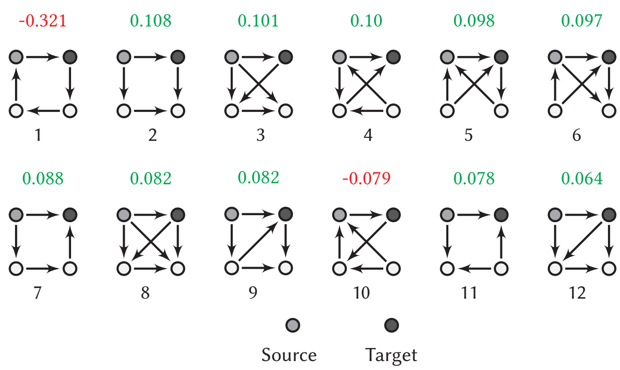

Figure 5: Top Cora ML SSTs Without Bidirected Citations – (See Section 5.3.2) – The top SST indicates that if an edge closes a 4-cycle that is considered a strong indicator that the edge is not genuine. Similarly, SST 10 suggests that a 3-cycle is unlikely, but not as unlikely as a 4-cycle. Other than SSTs 1 and 4, the top SSTs are positive indicators. SSTs 2 - 9, and 11 - 12 all include some kind of "transitivity", that nodes which cite (or are cited by) similar articles cite each other.

or deletions) using temporal graph attention layers. A Multi-Layer Perceptron decoder allows the TGN to score candidate edges with probabilities for evaluation of future link prediction.

5.4.2 Quantitative Results. Quantitative results are listed in Table 4. Once again our SST-based link predictors are among the top performers. Again, we suggest that these numbers be taken with a grain of salt because we simply used the GNNs' default hyperparameters. Chiefly, our tests demonstrate that our SSTs' elegant and interpretable results are validated by good prediction performance.

Note that we bypassed computing AUPR_3 on the Wikipedia graph due to the sheer size of the false test edge set - $O((10^5)^2)$.

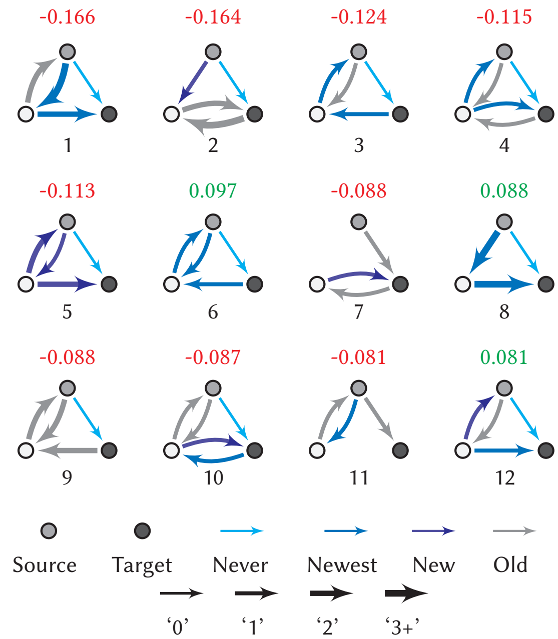

Figure 6: Top 3-Node SSTs for Wikipedia Link Additions – Main Takeaway: Wedges only close to triangles when the wedge had recent edges (*e.g.* Newest) appearing for the first or maybe second time (low frequency, *e.g.* '1'), ideally including an edge pointing to the target node of the new edge.

5.4.3 Interpretable Temporal Results. To demonstrate the interpretability of SSTs on temporal graphs, we explore the three-node SSTs on the Wikipedia edge additions graph. We find that, unlike the general assumption of triadic closure, according to our model many triangles are considered *unlikely* to close. It is only the triangles where certain connection combinations in the wedge were formed recently (indicated by our recency trait) and *for the first (or maybe second) time* (indicated by our frequency trait) that the wedge is quite likely to close into a triangle. See Figure 6. This is evidenced quantitatively by the fact that the three-node SST predictor performed much better than the Common Neighbors.

6 CONCLUSION

We defined an elegant generalization of Triadic Closure, the Subgraph-to-Subgraph Transition (SST). This generalization allowed us to use a simple classifier, the Linear SVM, to create interpretable link prediction models which performed comparatively with state of the art graph neural networks. We expect that the Subgraph-to-Subgraph Transition will become a standard tool in modeling graphs and that future research will produce new and creative ways to use and efficiently count SSTs.

ACKNOWLEDGEMENTS

This research is supported by a grant from the US National Science Foundation (#1652492).

REFERENCES

[1] Ghadeer AbuOda, Gianmarco De Francisci Morales, and Ashraf Aboulnaga. 2019. Link prediction via higher-order motif features. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 412–429.

[2] Lada A Adamic and Eytan Adar. 2003. Friends and neighbors on the web. *Social networks* 25, 3 (2003), 211–230.

[3] David Avis and Komei Fukuda. 1996. Reverse search for enumeration. *Discrete Applied Mathematics* 65, 1-3 (1996), 21–46.

[4] Albert-László Barabási and Réka Albert. 1999. Emergence of scaling in random networks. *science* 286, 5439 (1999), 509–512.

[5] Christoph Berkholz, Paul Bonsma, and Martin Grohe. 2017. Tight lower and upper bounds for the complexity of canonical colour refinement. *Theory of Computing Systems* 60, 4 (2017), 581–614.

[6] Ginestra Bianconi, Richard K Darst, Jacopo Iacovacci, and Santo Fortunato. 2014. Triadic closure as a basic generating mechanism of communities in complex networks. *Physical Review E* 90, 4 (2014), 042806.

[7] Jesse Davis and Mark Goadrich. 2006. The relationship between Precision-Recall and ROC curves. In *Proceedings of the 23rd international conference on Machine learning*. 233–240.

[8] Mark S Granovetter. 1977. The strength of weak ties. In *Social networks*. Elsevier, 347–367.

[9] Martin Grohe, Kristian Kersting, Martin Mladenov, and Pascal Schweitzer. 2017. Color refinement and its applications. *Van den Broeck, G.; Kersting, K.; Natarajan, S* (2017), 30.

[10] Madhav Jha, C Seshadhri, and Ali Pinar. 2015. Path sampling: A fast and provable method for estimating 4-vertex subgraph counts. In *Proceedings of the 24th International Conference on World Wide Web*. 495–505.

[11] Krzysztof Juszczyszyn, Katarzyna Musial, and Marcin Budka. 2011. Link prediction based on subgraph evolution in dynamic social networks. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*. IEEE, 27–34.

[12] Hemant Kasat, Sanket Markan, Manish Gupta, and Vikram Pudi. 2019. Temporal Link Prediction in Dynamic Networks. In *Proceedings of the Mining and Learning on Graphs workshop*.

[13] Thomas N Kipf and Max Welling. 2016. Variational Graph Auto-Encoders. *NIPS Workshop on Bayesian Deep Learning* (2016).

[14] Peter Klimek and Stefan Thurner. 2013. Triadic closure dynamics drives scaling laws in social multiplex networks. *New Journal of Physics* 15, 6 (2013), 063008.

[15] AA Leman. 1970. On automorphisms of certain classes of graphs. *Avtomatika i Telemehanika* 2 (1970), 75–82.

[16] David Liben-Nowell and Jon Kleinberg. 2007. The link-prediction problem for social networks. *Journal of the American society for information science and technology* 58, 7 (2007), 1019–1031.

[17] Penghang Liu, Valerio Guarrasi, and A Erdem Sarıyüce. 2020. Temporal Network Motifs: Models, Limitations, Evaluation. *arXiv preprint arXiv:2005.11817* (2020).

[18] Eugene M Luks. 1982. Isomorphism of graphs of bounded valence can be tested in polynomial time. *Journal of computer and system sciences* 25, 1 (1982), 42–65.

[19] Brendan D McKay and Adolfo Piperno. 2014. Practical graph isomorphism, II. *Journal of Symbolic Computation* 60 (2014), 94–112.

[20] Ashwin Paranjape, Austin R Benson, and Jure Leskovec. 2017. Motifs in temporal networks. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. 601–610.

[21] Ali Pinar, C Seshadhri, and Vaidyanathan Vishal. 2017. Escape: Efficiently counting all 5-vertex subgraphs. In *Proceedings of the 26th International Conference on World Wide Web*. 1431–1440.

[22] Nataša Pržulj. 2007. Biological network comparison using graphlet degree distribution. *Bioinformatics* 23, 2 (2007), e177–e183.

[23] Emanuele Rossi, Ben Chamberlain, Fabrizio Frasca, Davide Eynard, Federico Monti, and Michael Bronstein. 2020. Temporal Graph Networks for Deep Learning on Dynamic Graphs. In *ICML 2020 Workshop on Graph Representation Learning*.

[24] Ryan A Rossi, Anup Rao, Sungchul Kim, Eunyee Koh, and Nesreen Ahmed. 2020. From Closing Triangles to Closing Higher-Order Motifs. In *Companion Proceedings of the Web Conference 2020*. 42–43.

[25] Guillaume Salha, Romain Hennequin, and Michalis Vazirgiannis. 2020. Simple and Effective Graph Autoencoders with One-Hop Linear Models. In *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)*.

[26] Guillaume Salha, Stratis Limnios, Romain Hennequin, Viet Anh Tran, and Michalis Vazirgiannis. 2019. Gravity-Inspired Graph Autoencoders for Directed Link Prediction. In *ACM International Conference on Information and Knowledge Management (CIKM)*.

[27] C Seshadhri, Ali Pinar, Nurcan Durak, and Tamara G Kolda. 2017. Directed closure measures for networks with reciprocity. *Journal of Complex Networks* 5, 1 (2017), 32–47.

[28] Johan Ugander, Lars Backstrom, and Jon Kleinberg. 2013. Subgraph frequencies: Mapping the empirical and extremal geography of large graph collections. In *Proceedings of the 22nd international conference on World Wide Web*. 1307–1318.

[29] Boris Weisfeiler and Andrei A Lehman. 1968. A reduction of a graph to a canonical form and an algebra arising during this reduction. *Nauchno-Technicheskaya Informatsia* 2, 9 (1968), 12–16.

[30] Yang Yang, Ryan N Lichtenwalter, and Nitesh V Chawla. 2015. Evaluating link prediction methods. *Knowledge and Information Systems* 45, 3 (2015), 751–782.

[31] Hao Yin, Austin R Benson, and Johan Ugander. 2019. Measuring directed triadic closure with closure coefficients. *arXiv preprint arXiv:1905.10683* (2019).

[32] Muhan Zhang and Yixin Chen. 2017. Weisfeiler-lehman neural machine for link prediction. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 575–583.