Complex network prediction using deep learning

Yoshihisa Tanaka$^{1,2}$, Ryosuke Kojima$^{1,4*}$, Shoichi Ishida$^1$, Fumiyoshi Yamashita$^1$, and Yasushi Okuno$^2,3*$$^\dagger$

$^1$Graduate School of Pharmaceutical Sciences, Kyoto University
$^2$RIKEN Center for Computational Science, HPC and AI driven Drug Development Platform Division, Biomedical Computational Intelligence Unit
$^3$Graduate School of Medicine, Kyoto University

Abstract

Systematic relations between multiple objects that occur in various fields can be represented as networks. Real-world networks typically exhibit complex topologies whose structural properties are key factors in characterizing and further exploring the networks themselves. Uncertainty, modelling procedures and measurement difficulties raise often insurmountable challenges in fully characterizing most of the known real-world networks; hence, the necessity to predict their unknown elements from the limited data currently available in order to estimate possible future relations and/or to unveil unmeasurable relations. In this work, we propose a deep learning approach to this problem based on Graph Convolutional Networks for predicting networks while preserving their original structural properties. The study reveals that this method can preserve scale-free and small-world properties of complex networks when predicting their unknown parts, a feature lacked by the up-to-date conventional methods. An external validation realized by testing the approach on biological networks confirms the results, initially obtained on artificial data. Moreover, this process provides new insights into the retainability of network structure properties in network prediction. We anticipate that our work could inspire similar approaches in other research fields as well, where unknown mechanisms behind complex systems need to be revealed by combining machine-based and experiment-based methods.

network prediction | scale-free | complex network | GCNs

Introduction

By network we understand a complex system with multiple relations between its components described as a graph, which plays a role in a variety of fields such as biology (1–3), social sciences (4–6), disease transmission (7, 8) and the internet (9, 10). Many real-world networks have a complex topology (11); however, there are laws that govern the structural property of the entire network, and therefore, understanding the network as a whole structure is important for elucidating the systems underlying it.

Two common topological traits of real-world networks are known as ‘small-world’ (1) and ‘scale-free’ (12). The first of them means that the paths between any two nodes are short, and there are highly clustered connections (1) (e.g. your two friends are likely to be friends (13)). A network is scale-free when it has a few nodes with high connections and many nodes with low connections, where the degree distribution probability of nodes with $k$ connections, $P(k)$, follows a power law form as $p(k) \sim k^{-\gamma}$ ($\gamma$ is often between 2 and 4) (12). Especially in the field of biology, the network topology was found to be indispensable for understanding the systematic behaviour of life (14). In networks with these structural properties, highly connected nodes called ‘hubs’ plays a key role in fields when studying topics like network attack (15), immunization strategy (16), disease transmission (17) and airline transportation (18), irrespective of biological domain (19–22). Meanwhile, due to the uncertainty and difficulty in the measurement of real-world networks, it is still difficult to uncover true networks, and currently available networks are far from complete (23–25). Hence, there is a necessity for predicting whole networks from limited available data in order to estimate possible future relations or to complement unmeasured relations. This prediction task is known as link prediction in the field of information retrieval (26, 27). Several studies have argued that the addition of the predicted links to the network should not change the original network topology (25, 28–30); however, this conservative perspective is often overlooked in the pursuit of high accuracy in link prediction, which is one of the key aspects in the prediction of network structures as there is no guarantee that link prediction is equivalent to preserving properties of the overall network structure. For example, it is unclear whether just predicting individual links will lead to the retention of the structural features described above.

In this study, we focus on preserving the structural property of networks and propose for the first time a deep learning approach using Graph Convolutional Networks (GCNs) to predict networks (Fig. 1) (see ‘Methods’). We show that although conventional link prediction methods could not predict the entire network while preserving its innate structure characteristics, GCNs do this due to their ability of learning network structures. These theoretical achievements are confirmed by applying our approach to biological scale-free and small-world real-world networks. Our results suggest that GCNs could assist to extrapolate a true network from current network data while maintaining its structure.

Results

GCN reconstructed network topology of scale-free and small-world. To examine whether the GCN-predicted network had similar structural properties to the original one, we firstly carried out network prediction on the artificially-generated networks. We used three model networks: the most classic Erdős-Rényi (ER) model (random network) exhibiting Poisson distribution (31), the Watts-Strogatz (WS) model (small-world network) (1) and the Barabási-Albert
To confirm the validity and plausibility of the new links of the IP-predicted ones (Supplementary Fig. 4). Moreover, to evaluate the small-world effect for the pre-
dicted networks, we measured the clustering coefficient, an
dicator of small-world property (1). Among the three
model networks, the WS model showed this property, with a
high clustering coefficient (1) (Supplementary Table 2). We
found that, regardless of model network, the clustering
coefficient of the GCN-predicted networks was not as high as that of the IP-predicted networks (Supplementary Fig. 4). Moreover, the DistMult- and IP-predicted networks tended to show lower clustering coefficients than the GCN-predicted networks (Supplementary Table 1). After carrying out network predictions for each biomolecular network, we noted that the degree distribution of the predicted networks is affected as the number of links increases. While the DistMult- and IP-predicted networks gradually came near random graph-like degree distributions, the GCN-predicted networks consistently followed the power law (Supplementary Fig. 1-3). This suggests that GCNs preserve the small-world property unlike DistMult and IP, probably due to the differences in network structure learning.

Moreover, to evaluate the small-world effect for the predicted networks, we measured the clustering coefficient, an indicator of small-world property (1). Among the three model networks, the WS model showed this property, with a high clustering coefficient (1) (Supplementary Table 2). We found that, regardless of model network, the clustering coefficient of the GCN-predicted networks was not as high as that of the IP-predicted networks (Supplementary Fig. 4). Moreover, the DistMult- and IP-predicted networks displayed lower clustering coefficients than the GCN-predicted networks (Supplementary Table 1). After carrying out network predictions for each biomolecular network, we noted that the degree distribution of the predicted networks is affected as the number of links increases. While the DistMult- and IP-predicted networks gradually came near random graph-like degree distributions, the GCN-predicted networks consistently followed the power law (Supplementary Fig. 1-3). This suggests that GCNs preserve the small-world property unlike DistMult and IP, probably due to the differences in network structure learning.
Hence, the internal validation using real-world biological networks suggest that GCNs can reconstruct network while preserving their properties.

**External validation of the GCN-predicted networks using the other real-world networks.** To examine the true validity of the predicted network, we conducted an external validation, too, by comparing the predicted network above using the real-world biological network with another external interactome network obtained from experiments. We investigated the latest experimental interactome network of the human reference interactome (HuRI) (36), noting its scale-free and small-world characteristics (Supplementary Fig. 12a, Supplementary Table 1). Among the six considered biomolecular networks, the ‘interaction’ one is most similar to HuRI (Supplementary Fig. 12b). The analysis of the extent
Fig. 3. The degree distributions of the predicted networks for the six biomolecular networks (green labelled columns). The link threshold was adopted as 10000 for expression, 40000 for interaction, 7000 for phosphorylation, 20000 for state change, 30000 for complex, 40000 for catalysis. The original degree distributions of the individual biomolecular networks were displayed at the original network panel (magenta labelled columns), and their network properties were listed in Supplementary Table 1. The X axis represents the degree \( k \), and the Y axis the degree distribution probability \( P(k) \). The detailed network properties of the predicted networks were listed in Supplementary Table 3.
Fig. 4. Performance evaluation of the predicted network for the six biomolecular networks. (a-f) The enrichment analysis with the three methods: IP (grey), DistMult (blue), GCN (magenta). The X axis represents an arbitrary link threshold. (g) The comparison of matching links of the predicted ‘interaction’ biomolecular network with the HuRI network by the three methods: IP (grey), DistMult (blue), GCN (magenta). The X axis represents the link threshold for the predicted networks, while the Y axis represents the number of the matched links. Error bar: mean ± standard deviation (n = 3).

to which the predicted network generated from ‘interaction’ (Fig. 3, Supplementary Fig. 6) matches HuRI shows that while DistMult is inferior to IP and GCNs, the degree of matching of the IP- and GCN-predicted networks seems to be almost comparable (Fig. 4g). However, given that the IP-predicted networks do not show scale-free on the structural level (Fig. 3, Supplementary Fig. 6), even if their degree of matching is better to some extent (Fig. 4g), they do not reflect the intrinsic structure of the HuRI network. In contrast, the GCN-predicted network satisfies both the high degree of matching and structural homology with the HuRI network (Fig. 3, Fig. 4g, Supplementary Fig. 6). This evidence further supports that network prediction by GCNs is more effective for predicting real-world networks than its currently known counterparts.

Discussion

We propose a deep learning approach for network prediction (Fig. 1) based on GCNs that predicts complex networks while preserving their structural properties (Fig. 2, Fig. 3). Experiments with biological networks guarantee the robust validity of this new method, also showing that it outperforms
its currently standard counterparts (Fig. 4). As most real-world networks exhibit complex structural properties (11), we focused in our study on the importance of their retention in order to capture the hidden mechanisms behind them. Our results indicate that the deep learning-based GCNs are able to capture the structural features of the original networks by learning their structural information. Previous work demonstrated that it is difficult to infer the scale-free property of a network by partial network sampling due to the small coverage of the protein–protein interaction networks (29). Our analysis reveals that the emerging deep learning techniques can contribute to overcome the challenge of predicting networks while preserving their topology. Moreover, we also expect the graph embedding used in GCNs to be useful in machine learning tasks, such as graph or node classification (27). In the process of expanding the applicability range of the graph embedding, GCN could prove to be useful in achieving better graph representations by learning the network structure information (37, 38), as suggested by our study. However, the limitations of our approach in predicting networks essentially lie in the training dataset. New external nodes cannot be added to the predicted network and consequently it is almost infeasible to predict nodes not included in the current data that could be discovered in the future. The scale-free network model is also known as the preferential attachment model (12) in which new nodes are preferentially attached to highly connected ones forming nodes with higher degree, which eventually grows the network. Although GCNs might potentially learn these processes, this point of handling new nodes is slightly different from the BA model. By extending the framework to address this issue, network prediction would be more in line with the nature of scale-free networks. Our work sheds light on a landscape for investigating unknown mechanisms behind complex systems by combining machine-based and experiment-based methods.

Methods

Development of an approach for network prediction. The proposed approach for network prediction is presented in Fig. 1. The key idea was inspired by graph embeddings, where graph-structured data is projected onto vector spaces. A node receives a low-dimensional vector representation to reflect the relational data through a process known as representation learning (27).

While different embedding approaches have been studied in link prediction (32, 39–42), the remarkable progress of artificial intelligence (AI) research in recent years, in particular of deep learning (37, 43–45), raises hopes towards more accurate predictions. In our investigation, we employ to this end GCNs (37) due to their performance in various prediction tasks (38).

The principle behind GCNs is conceptually similar to the one of Convolutional Neural Networks (CNNs), which brought a breakthrough in imaging. A CNN learns image features by convolving surrounding pixel information for a pixel (46), whereas a GCN learns graph topological features by convolving adjacent node information of a node (37). Inspired by CNNs’ ability to restore original images (47), we presumed that GCNs can reconstruct network topological characteristics by learning the network structure. To test this hypothesis, we implemented a GCN and developed an approach for network prediction, which is explained step-by-step below and illustrated in Fig. 1. The input is a network data with its characteristic structure (Fig. 1a). The topological structure of the network is learned by the GCN, and this topological information is transformed into a vector space, the individual nodes of which are represented as feature vectors (Fig. 1b). Next, by using these feature vectors, link scores are calculated for unknown links (Fig. 1c). As the links with high scores are more likely to appear in the future, they are extracted in a descending order and integrated into the input network, constructing thus the predicted network (Fig. 1d). We defined a set of these sequential steps (starting with the network as input and ending with the predicted network as output) as network prediction.

Learning architecture. The core learning frame of network prediction followed an already established path (38). The learning architecture mainly consists of two parts, graph embedding as encoder and link scoring as decoder. In the process of graph embedding (Fig. 1b), the graph convolution technique was used for the GCN to encode an input network as a vector. Our implementation followed Kipf’s model (37) with two graph convolution layers and the activation function ReLU. In the process of link scoring (Fig. 1c), the encoded vector data is decoded by the scoring function as a score that represents how likely an unknown link is to exist. For the GCN and IP models, the score of a possible link is computed as a dot product of the feature vectors of the two nodes at a given link (33, 48). For the DistMult model, the weighted score is computed following the matrix factorization algorithm, which is a standard benchmark for link prediction (32). To allow nodes acquire more accurate representations, the model is trained to gain a larger score for an existing link (positive link) than for a non-existing one (negative link). As it is impossible to know whether non-existing links in a current network will appear in the future or not, negative links were randomly sampled as previously established (32, 38, 41). Since we assumed that the quantity of available data is limited and small (23, 24), the input network data were divided into 5000 links for training and the remaining for testing. In the process of constructing a predicted network (Fig. 1d), the links with high scores were integrated into the training data in a descending order. The predicted network is determined as a network consisting of this set of links. The number of links to be assigned to the predicted network can be a priori set as a link threshold. The procedure of network prediction was independently performed three times with different datasets. The learning parameters are listed in Supplementary Table 4. We implemented the method in python, integrating it into our open source GCN platform, kGCN (49).
Enrichment analysis. Enrichment was defined as the ratio between the number of test links in top N links and the total number of links, where N is the arbitrary link threshold, and the total number of links is $n(n-1)/2$, $n$ being the number of total nodes in the input network. This is an indicator of the precision performance of link prediction. The higher the value, the more precisely the links are predicted.

Network dataset. To test the proposed approach we consider 10 networks, both artificially generated and real world. Three model networks (BA model, WS model and ER model) were generated using the python package networkx (50) in order to share approximately the same graph density. We also consider some biomolecular networks whose original dataset (version 11) was downloaded from Pathway Commons (34, 35). We selected the six types of graph datasets: ‘control expression of’, ‘interact with’, ‘control phosphorylation of’, ‘control state change of’, ‘in complex with’, ‘catalysis precedes’. Each dataset was preprocessed by conversion into an undirected graph and removal of selfloops. The HuRI network was extracted from The Human Reference Protein Interactome Mapping Project (36). Since the names of nodes were written in Ensembl gene identifier, these were converted into gene symbols to correspond with the descriptor of Pathway Commons using the python package mygene descriptor of Pathway Commons using the python package powerlaw (52). To confirm the fitting state of power law, the exponential function was employed for the fitting comparison. The power law fitting was determined when the likelihood value was positive. Network properties (53) were calculated using python package networkx (50) of the largest one. Network visualization was performed using python package networkx (50). When multiple connected components were observed in a network, the network property analysis were performed on the largest one. Network visualization was performed using Cytoscape (54).

Data availability. The public network datasets used in this study are freely downloaded at Pathway Commons (https://www.pathwaycommons.org/) and The Human Reference Protein Interactome Mapping Project (http://www.interactome-atlas.org/). The three model networks generated in this study are included in the GitHub repository.

Code availability. The whole code for network prediction is available at our open source GCN platform for lifesiscience, kGCN (https://github.com/clinfo/kGCN).

ACKNOWLEDGEMENTS

We thank Y. Tamada for helpful discussions. This study was supported by RIKEN Junior Research Associate Program; Medical Science Innovation Hub Program of RIKEN: AMED under Grant Number P20k0205013; Cabinet Office, Government of Japan, Public/Private R&D Investment Strategic Expansion Program (PRISIM). This paper format was generated through self-modification of the original template designed by Ricardo Henriqueis.

AUTHOR CONTRIBUTIONS

Y.T., R.K. and Y.O. designed the experiments. Y.T. and R.K. conducted the experiments. Y.T., R.K. and S.I. wrote the codes. Y.T., R.K., S.I. and Y.O. analysed the results. Y.T. and Y.O. wrote the original manuscript. R.K., S.I. and F.Y. reviewed and edited the manuscript. F.Y. and Y.O. supervised the study.

COMPETING FINANCIAL INTERESTS

The authors declare no competing or financial interests.

Bibliography

1. D J Watts and S H Strogatz. Collective dynamics of ‘small-world’ networks. Nature, 393(6684):440–442, June 1998. ISSN 0028-0836. doi: 10.1038/30918.
2. H Jeong, B Tombor, R Albert, Z N Oltvai, and A L Barabási. The large-scale organization of metabolic networks. Nature, 407(6804):651–654, October 2000. ISSN 0028-0836. doi: 10.1038/35093627.
3. David A Fell and Andreas Wagner. The small world of metabolism. Nat. Biotechnol., 18(11):1121–1122, November 2000. ISSN 1087-0156, 1546-1696. doi: 10.1038/nbt1025.
4. L A Amaral, A Scala, M Barthelemy, and H E Stanley. Classes of small-world networks. Proc. Natl. Acad. Sci. U. S. A., 97(21):11148–11152, October 2000. ISSN 0027-8424. doi: 10.1073/pnas.023121797.
5. M E J Newman. The structure of scientific collaboration networks. Proc. Natl. Acad. Sci. U. S. A., 98(2):404–409, January 2001. ISSN 0027-8424, 1091-6480. doi: 10.1073/pnas.98.2.404.
6. Sofia Aparicio, Javier Villazín-Terrazas, and Gonzalo Álvarez. A model for Scale-Free networks: Application to twitter. Entropy, 17(8):8548–8567, August 2015. doi: 10.3390/entropy170808548.
7. F Liljeros, C R Edling, L A Amaral, H E Stanley, and Y Aberg. The web of human sexual contacts. Nature, 411(6840):907–909, June 2001. ISSN 0028-0836. doi: 10.1038/35082140.
8. Helena A Herrmann and Jean-Marc Schwartz. Why COVID-19 models should incorporate the network of social interactions. Phys. Biol., 17(6):065008, October 2020. ISSN 1478-3975, 1478-3967. doi: 10.1088/1478-3975/abaec5.
9. Rükka Albert, Haoqiong Jeong, and Albert-László Barabási. Diameter of the World-wide web. Nature, 401(6749):135–131, September 1999. ISSN 0028-0836, 1476-4687. doi: 10.1038/46301.
10. A Vázquez, R Pastor-Satorras, and A Vespignani. Internet topology at the router and autonomous system level. June 2002.
11. Rükka Albert and Albert-László Barabási. Statistical mechanics of complex networks. Rev. Mod. Phys., 74(1):47–97, January 2002. ISSN 0034-6861. doi: 10.1103/RevModPhys.74.47.
12. A L Barabasi and R Albert. Emergence of scaling in random networks. Science, 286(5439): 509–512, October 1999. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.286.5439.509.
13. Lars Backstrom, Paolo Boldi, Marco Rosa, Johan Ugander, and Sebastiano Vigna. Four degrees of separation. In Proceedings of the 4th Annual ACM Web Science Conference, WebSci ’12, pages 33–42, New York, NY, USA, June 2012. Association for Computing Machinery. ISBN 9781450321288. doi: 10.1145/2308718.2308723.
14. Albert-László Barabási and Zoltán N Oltvai. Network biology: understanding the cell’s functional organization. Nat. Rev. Genet., 5(2):101–113, February 2004. ISSN 1471-0056. doi: 10.1038/nrg1272.
15. R Albert, H Jeong, and A L Barabási. Error and attack tolerance of complex networks. Nature, 406(6794):378–382, July 2000. ISSN 0028-0836, 1476-4687. doi: 10.1038/35101909.
16. Romualdo Pastor-Satorras and Alessandro Vespignani. Immunization topology at the router and autonomous system level. June 2002.
17. Matt J Keeling and Ken T D Eames. Networks and epidemic models. J. R. Soc. Interface, 2(4):295–307, September 2005. ISSN 1742-5689, 1742-5662. doi: 10.1098/rsif.2005.0051.
18. Massimiliano Zanin and Fabrizio Lillo. Modelling the air transport with complex networks: A short review. Eur. Phys. J. Spec. Top., 215(1):5–21, January 2013. ISSN 1951-6355, 2056-5950. doi: 10.1186/1478-3975-215-1.
19. H Jeong, S P Mason, A L Barabási, and Z N Oltvai. Lethality and centrality in protein networks. Nature, 411(6833):440–442, May 2001. ISSN 0027-8424. doi: 10.1038/35075138.
20. Katja Basso, Adam A Margolin, Gustavo Stolovitzky, Ulf Klein, Ricardo Dalla-Favera, and Andrea Califano. Reverse engineering of regulatory networks in human B cells. Nat. Genet., 37(4):382–390, April 2005. ISSN 1061-4036. doi: 10.1038/ng1552.
21. Xiaoning He and Jianzhi Zhang. Why do hubs tend to be essential in protein networks? Proc. Natl. Acad. Sci. U. S. A, 105(19):6959–6964, May 2008. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0708078105.
22. Michael P H Stumpf, Thomas Thorne, Eric de Silva, Ronald Stewart, Hyeong Jun An, Michael Lapke, and Carsten Wul. Estimating the size of the human interactome. Proc. Natl. Acad. Sci. U. S. A., 105(19):6959–6964, May 2008. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.0708078105.

Tanaka et al. | complex network prediction
arXiv | 7
25. Aaron Clauset, Cristopher Moore, and M E J Newman. Hierarchical structure and the prediction of missing links in networks. Nature. 453(7191):98–101, May 2008. ISBN 0028-0836, 1476-4847. doi: 10.1038/nature06830.

26. Linyuan Lü and Tao Zhou. Link prediction in complex networks: A survey. Physica A: Statistical Mechanics and its Applications. 389(6):1150–1170, March 2011. ISSN 0378-4371. doi: 10.1016/j.physa.2010.11.072.

27. William L Hamilton, Rex Ying, and Jure Leskovec. Representation learning on graphs: Methods and applications. September 2017.

28. M Gomez, S H Li, and A Rotshtein. Probabilistic prediction of unknown metabolic and signal-transduction networks. Genetics. 159(3):1291–1298, November 2001. ISSN 0016-7511. doi: 10.1101/gr.1239303

29. Jing-Dong J Han, Denis Dupuy, Nicolas Bertin, Michael E Cusick, and Marc Vidal. Effect of sampling on topology predictions of protein-protein interaction networks. Nat. Biotechnol., 23(7):839–844, July 2005. ISSN 1087-0156. doi: 10.1101/jcb.011371.116.

30. Roger Guimerà and Marta Sales-Pardo. Missing and spurious interactions and the reconstruction of complex networks. Proc. Natl. Acad. Sci. U. S. A., 106(52):22073–22078, December 2009. ISSN 0027-8424. doi: 10.1073/pnas.0908366106.

31. Paul Erdős and Alfréd Rényi. On the evolution of random graphs. Adv. Math., 51(2–3):17–60, 1960.

32. Bishan Yang, Wen-Tau Yih, Xiaodong He, Jianting Gao, and Li Deng. Embedding entities and relations for learning and inference in knowledge bases. December 2014.

33. Yehuda Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD ’08, pages 426–434, New York, NY, USA, August 2008. Association for Computing Machinery. ISBN 9781605581934. doi: 10.1145/1418906.1419344.

34. Ethan G Cerami, Benjamin E Gross, Christoph Gembicki, Ozgün Babur, Nadia Awar, Nicolas Schulz, Gary D Bader, and Chris Sander. Pathway commons, a web resource for biological pathway data. Nucleic Acids Res., 39(Database issue):D685–90, January 2011. ISSN 0001-8732. doi: 10.1101/ijr.kj0139.

35. Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? October 2018.

36. Katja Luck, Dae-Kyum Kim, Luke Lambourne, Kerstin Spirohn, Bridget Teeking, Anjali Gopal, Ghazal Haddad, Eidite Hatchi, Mohamed Helmy, Yves Jacob, Joseph Kassa, Serena Landini, Roujia Li, Natascha van Lieshout, Andrew MacWilliams, Dylan Markey, Joseph N Richardson, Sadie Schlabach, Omer Basha, Christian Bowman-Colin, Suet-Feung Chin, Soon Gang Choi, Claudia Colabella, Georges Coppin, Cassandra D’Amata, David De Ridder, Steffi De Rouck, Miquel Duran-Frigola, Hanane Ennajdaoui, Florian Goebels, Liana Goerbing, Aric Hagberg, Irma Kovacs, Miles W Mee, Joseph C Mellor, Carl Pollis, Carls Pons, Aaron D Richardson, Sadie Schlabach, Bridge Tseung, Anupama Yadav, Mariana Babor, Dongsic Choi, Atina G Coté, Meaghan Dalago, Ethan Cerami, Benjamin Gross, Ugur Dogrusoz, Emek Demir, Gary D Bader, and Chris Sander. Pathway commons 2019 update: integration and analysis and exploration of pathway data. Nucleic Acids Res., 48(O1):D489–D497, January 2019. ISSN 0001-8732. doi: 10.1101/ijr.kj0139.

37. Thomas N Kipf and Max Welling. Semi-Supervised classification with graph convolutional networks. Paper presented at the International Conference on Machine Learning, 2016.

38. Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling. Modeling relational data with graph convolutional networks. March 2017.

39. Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhrken. Translating embeddings for modeling multi-relational data. In J C J Burges, L Bottou, M Welling, Z Ghahramani, and K Q Weinberger, editors, Advances in Neural Information Processing Systems 26, pages 2787–2795, Curran Associates, Inc., 2013.

40. M Nickel, L Rosasco, and T Poggio. Holographic embeddings of knowledge graphs. Advances in Neural Information Processing Systems, 25:1097–1105, 2012. ISSN 1049-5258.

41. Jonathan Masci, Ueli Meier, Dan Ciresan, and Jürgen Schmidhuber. Stacked convolutional Auto-Encoders for hierarchical feature extraction. In Artificial Neural Networks and Machine Learning – ICANN 2011, pages 52–59. Springer Berlin Heidelberg, 2011. doi: 10.1007/978-3-642-21735-7_7.

42. Thomas N Kipf and Max Welling. Variational graph Auto-Encoders. November 2016.

43. Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Veličković, and Geoffrey E Hinton. Imagener classifier with deep convolutional neural networks. Adv. Neural Inf. Process. Syst., 25:1097–1105, 2012. ISSN 1049-5258.

44. Jonathan Masci, Ueli Meier, Dan Ciresan, and Jürgen Schmidhuber. Stacked convolutional Auto-Encoders for hierarchical feature extraction. In Artificial Neural Networks and Machine Learning – ICANN 2011, pages 52–59. Springer Berlin Heidelberg, 2011. doi: 10.1007/978-3-642-21735-7_7.

45. Thomas N Kipf and Max Welling. Variational graph Auto-Encoders. November 2016.
Supplementary Fig. 1. The degree distributions of the predicted networks with the sequential link thresholds for the BA model network. The panels with dotted lines (blue) signal that the degree distribution follows the power law. The marked panels were used in Fig. 2. The detailed network properties of the predicted networks were listed in Supplementary Table S2.
Supplementary Fig. 2. The degree distributions of the predicted networks with the sequential link thresholds for the WS model network. The panels with dotted lines (blue) signal that the degree distribution follows the power law. The marked panels were used in Fig. 2. The detailed network properties of the predicted networks is listed in Supplementary Table S2.
Supplementary Fig. 3. The degree distributions of the predicted networks with the sequential link thresholds for the ER model network. The panels with dotted lines (blue) signal that the degree distribution follows the power law. The marked panels were used in Fig. 2. The detailed network properties of the predicted networks is listed in Supplementary Table S2.
Supplementary Fig. 4. The transition of clustering coefficient for the three model networks. The X axis represents the link threshold for the predicted networks. The detailed value is listed in Supplementary Table 2. Error bar: mean ± standard deviation (n = 3).
Supplementary Fig. 5. The degree distributions of the predicted networks with the sequential link thresholds for the ‘expression’ biomolecular network. The panels with dotted lines (blue) signal that the degree distribution follows the power law. The marked panels were used in Fig. 3. The detailed network property is listed in Supplementary Table 3.
Supplementary Fig. 6. The degree distributions of the predicted networks with the sequential link thresholds for the ‘interaction’ biomolecular network. The panels with dotted lines (blue) signal that the degree distribution follows the power law. The marked panels were used in Fig. 3. The detailed network property is listed in Supplementary Table 3.
Supplementary Fig. 7. The degree distributions of the predicted networks with the sequential link thresholds for the ‘phosphorylation’ biomolecular network. The panels with dotted lines (blue) signal that the degree distribution follows the power law. The marked panels were used in Fig. 3. The detailed network property is listed in Supplementary Table 3.
Supplementary Fig. 8. The degree distributions of the predicted networks with the sequential link thresholds for the ‘state change’ biomolecular network. The panels with dotted lines (blue) signal that the degree distribution follows the power law. The marked panels were used in Fig. 3. The detailed network property is listed in Supplementary Table 3.
Figure S9

| Link Thresholds | GCN | DistMult | IP |
|-----------------|-----|----------|----|
| 10000           | ![Graph](image1) | ![Graph](image2) | ![Graph](image3) |
| 20000           | ![Graph](image4) | ![Graph](image5) | ![Graph](image6) |
| 30000 *         | ![Graph](image7) | ![Graph](image8) | ![Graph](image9) |
| 40000           | ![Graph](image10) | ![Graph](image11) | ![Graph](image12) |
| 50000           | ![Graph](image13) | ![Graph](image14) | ![Graph](image15) |

Supplementary Fig. 9. The degree distributions of the predicted networks with the sequential link thresholds for the 'complex' biomolecular network. The panels with dotted lines (blue) signal that the degree distribution follows the power law. The marked panels were used in Fig. 3. The detailed network property is listed in Supplementary Table 3.
Supplementary Fig. 10. The degree distributions of the predicted networks with the sequential link thresholds for the 'catalysis' biomolecular network. The panels with dotted lines (blue) signal that the degree distribution follows the power law. The marked panels were used in Fig. 3. The detailed network property is listed in Supplementary Table 3.
Supplementary Fig. 11. The transition of clustering coefficient for the six biomolecular networks. The X axis represents the link threshold for the predicted networks. The detailed value is listed in Supplementary Table 3. Error bar: mean ± standard deviation (n = 3).
Supplementary Fig. 12. Characterization of the HuRI network. (a) The degree distribution probability of the HuRI network. The X axis represents the degree k, and the Y axis the degree distribution probability P(k). The detailed network property is listed in Supplementary Table 1. (b) The coverage rate of nodes and links between the HuRI network and the six biomolecular networks.