

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/337719367>

LinkAUC: Unsupervised Evaluation of Multiple Network Node Ranks Using Link Prediction

Conference Paper · December 2019

DOI: 10.1007/978-3-030-36687-2_1

CITATIONS

0

READS

29

3 authors:



[Emmanouil Krasanakis](#)

The Centre for Research and Technology, Hellas

7 PUBLICATIONS 21 CITATIONS

[SEE PROFILE](#)



[Symeon Papadopoulos](#)

The Centre for Research and Technology, Hellas

175 PUBLICATIONS 2,321 CITATIONS

[SEE PROFILE](#)



[Ioannis \(Yiannis\) Kompatsiaris](#)

The Centre for Research and Technology, Hellas

773 PUBLICATIONS 8,007 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Dem@Care EU-FP7 [View project](#)



KRISTINA: A Knowledge-Based Information Agent with Social Competence and Human Interaction Capabilities [View project](#)

LinkAUC: Unsupervised Evaluation of Multiple Network Node Ranks using Link Prediction

Emmanouil Krasanakis, Symeon Papadopoulos, and Yiannis Kompatsiaris

CERTH-ITI, Thessaloniki, Greece
maniospas,papadop,ikom@iti.gr

Abstract. An emerging problem in network analysis is ranking network nodes based on their relevance to metadata groups that share attributes of interest, for example in the context of recommender systems or node discovery services. For this task, it is important to evaluate ranking algorithms and parameters and select the ones most suited to each network. Unfortunately, large real-world networks often comprise sparsely labelled nodes that hinder supervised evaluation, whereas unsupervised measures of community quality, such as density and conductance, favor structural characteristics that may not be indicative of metadata group quality. In this work, we introduce LinkAUC, a new unsupervised approach that evaluates network node ranks of multiple metadata groups by measuring how well they predict network edges. We explain that this accounts for relation knowledge encapsulated in known members of metadata groups and show that it enriches density-based evaluation. Experiments on one synthetic and two real-world networks indicate that LinkAUC agrees with AUC and NDCG for comparing ranking algorithms more than other unsupervised measures.

1 Introduction

It is well-known that network nodes can be organized into communities [1,2,3,4] identified through either ground truth structural characteristics or shared node attributes [5,6,7]. A common task in network analysis is to rank all network nodes based on their relevance to such communities, especially of the second type [8,9], which are commonly referred to as metadata groups. Ranking nodes is particularly important in large social networks, where metadata group boundaries can be vague [10,11]. Node ranks can also be used by recommender systems that combine them with other characteristics, in which case it is important to be of high quality across the whole network. Some of the most well-known algorithms that discover communities with only a few known members also rely on ranking mechanisms and work by thresholding their outcome [12,13].

Node ranks for metadata groups are a form of recommendation and their quality is usually (e.g. in [14]) evaluated with well-known recommender system measures [15,16,17], such as AUC and NDCG. Since calculating these measures requires knowledge of node labels, the efficacy of ranking algorithms needs be demonstrated on labeled networks, such as those of the SNAP repository¹. How-

¹<https://snap.stanford.edu/data/>

ever, different algorithms and parameters are more suited to different networks, for example based on how well their assumptions match structural or metadata characteristics. At the same time, large real-world networks are often sparsely labeled, prohibiting supervised evaluation. In such cases, there is a need to evaluate ranking algorithms on the network at hand using unsupervised procedures.

A first take on unsupervised evaluation would be to generalize traditional structural community measures, such as density [18], modularity [19] and conductance [20], to support ranks. However, these measures are designed with structural ground truth communities in mind and often fail to assess hierarchical dependencies or other meso-scale (instead of local) features [6,11,21] that may characterize metadata groups. To circumvent this problem, we propose utilizing the network’s structure and the existence of multiple metadata groups; under the assumption that network edges are influenced by node metadata similarity [6], a phenomenon known as homophily in social networks [22], we assess the quality of ranks for multiple metadata groups based on their ability to predict network edges. We show that this practice enriches density-based evaluation and that it agrees with supervised measures better than other unsupervised ones.

2 LinkAUC

The main idea behind our approach is that, if there is little information to help evaluate node ranks, we can evaluate other related structural characteristics instead. To this end, we propose using node rank distributions across metadata groups to derive link ranks between nodes. Link ranks can in turn be evaluated through their ability to predict the network’s edges.

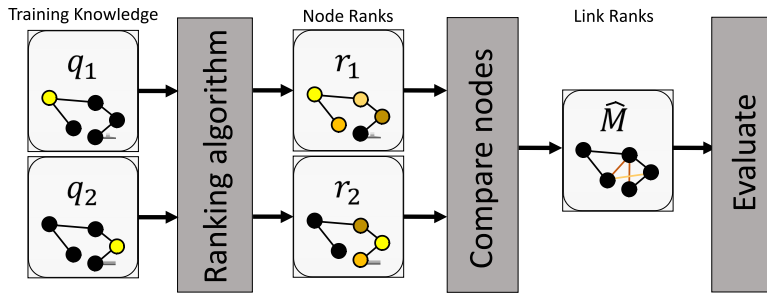


Fig. 1. Proposed scheme for evaluating ranking algorithms. Lighter colored nodes have higher ranks and lighter colored edges have lower link ranks.

An overview of the proposed scheme is demonstrated in Figure 1. In this section, we first justify why we expect node rank quality to follow link rank quality (Subsection 2.1) and formally describe the evaluation process of the latter using AUC (Subsection 2.2). We then show that link rank quality enriches density-based evaluation (Subsection 2.3).

2.1 Link Ranks

Let r_i be vectors whose elements r_{ij} estimate the relevance of network nodes j to metadata groups $i = 1, \dots, n$. Motivated by latent factor models for link prediction [23] and collaborative filtering [24], we consider $R = [r_1 \dots r_n]$ a matrix factorization of the network. Its rows $R_j = [r_{1j} \dots r_{nj}]$ represent the distribution of ranks of network nodes j across metadata groups. Following the principles of previous link prediction works [25,26], if network construction is influenced predominantly by structure-based and metadata-based characteristics, this factorization can help predict network edges by linking nodes with similar rank distributions. We calculate the similarities of rank distributions between nodes j, k using the dot product² as $\widehat{M}_{jk} = R_j \cdot R_k$. These form a matrix of link ranks:

$$\widehat{M} = RR^T \quad (1)$$

Accurate link prediction using link ranks implies good metadata group representations. To empirically understand this claim, let us consider ranking algorithms that can be expressed as network filters $f(M) = \sum_{n=0}^{\infty} a_n M^n$ [27] of the network’s adjacency matrix M , where a_n are the weights placed on random walks of length n . For example, personalized PageRank and Heat Kernels arise from exponentially degrading weights and the Taylor expansion coefficients of an exponential function respectively. If applied on query vectors q_i , where q_{ij} are proportional to probabilities that nodes j belong to metadata groups i , network filters produce ranks $r_i = f(M)q_i$ of how much nodes pertain to the metadata groups. Organizing multiple queries into a matrix $Q = [q_1 \dots q_n]$:

$$R = f(M)Q \Rightarrow \widehat{M} = f(M)QQ^T f^T(M) \quad (2)$$

This is a quadratic form of $f(M)$ around the kernel QQ^T and, as such, propagates link ranks between queries to the rest of link candidates. Therefore, if queries adequately predict the links between involved query nodes and link ranks can predict the network’s edges, then the algorithm with filter $f(M)$ is a good rank propagation mechanism. At best, queries form an orthonormal basis of ranks $QQ^T = I$ and this process can express any symmetric link prediction filter [25,26,28] by decomposing it to $f(M)f^T(M)$.

2.2 Link Rank Evaluation using AUC

When evaluating link ranks, it is often desirable to exclude certain links, such as withheld test edges or those absent due to systemic reasons (e.g. users may not be allowed to befriend themselves in social networks). To model this, we devise the notion of a *network group* that uses a binary matrix \mathbb{M} to remove non-comparable links of the network’s adjacency matrix M by projecting the latter to $M \odot \mathbb{M}$, where \odot is the Hadamard product performing elementwise multiplication. For

²Cosine similarity would arise by a fixed-flow assumption of the ranking algorithm that performs row-wise normalization of R before the dot product.

example, network groups of zero-diagonal networks correspond to $\mathbb{M} = 1 - I$, where 1 are matrices of ones and I identity matrices.

To help evaluate link ranks \widehat{M} against a network with adjacency matrix M within a network group \mathbb{M} , we introduce a transformation $vec_{\mathbb{M}}(M)$ that creates a vector containing all elements of M for which $\mathbb{M} \neq 0$ in a predetermined order. Then, link ranks can be evaluated by comparing $vec_{\mathbb{M}}(\widehat{M})$ with $vec_{\mathbb{M}}(M)$.

A robust measure that compares operating characteristic trade-offs at different decision thresholds is the Area Under Curve (AUC) [29], which has been previously used to evaluate link ranks [26]. When network edges are not weighted, if $TPR(\theta)$ and $FPR(\theta)$ are the true positive and false positive rates of a decision threshold θ on $vec_{\mathbb{M}}(\widehat{M})$ predicting $vec_{\mathbb{M}}(M)$, the AUC of link ranks becomes:

$$LinkAUC = \int_{-\infty}^{\infty} TPR(\theta)FPR'(\theta) d\theta \quad (3)$$

This evaluates whether actual linkage is assigned higher ranks across the network [30] without being affected from edge sparsity. These properties make LinkAUC preferable to precision-based evaluation of link ranks, which assesses the correctness of only a fixed number of top predictions [26].

2.3 Relation to Rank Density

The density of a network is defined as the portion of edges compared to the maximum number of possible ones [31,32]. Using the notion of volume $vol(M)$ to annotate the number of edges in a network with adjacency matrix M , the density of its projection inside the network group \mathbb{M} becomes $D_{\mathbb{M}}(M) = \frac{vol(M \odot \mathbb{M})}{vol(\mathbb{M})}$. We similarly define rank density by substituting the volume with the expected volume $vol(M, r)$ of the fuzzy set of subgraphs arising from ranks being proportional to the probabilities that nodes are involved in links:

$$\begin{aligned} vol(M, r) &= \mathbb{E}_{v \sim r} [v^T M v] = \frac{r^T M r}{\|r\|_1^2} \\ \Rightarrow D_{\mathbb{M}}(M, r) &= \frac{vol(M \odot \mathbb{M}, r)}{vol(\mathbb{M}, r)} = \frac{r^T (M \odot \mathbb{M}) r}{r^T \mathbb{M} r} \end{aligned} \quad (4)$$

where $\|\cdot\|_1$ is the L1 norm, calculated as the sum of vector elements, and v are binary vectors of vertices sampled with probabilities r .

We first examine the qualitative relation between link ranks and rank density for a single metadata group $R = r_1$. Annotating as $m \geq \theta$ the vectors arising from binary thresholding on the elements of $m = \frac{vec_{\mathbb{M}}(\widehat{M})}{\|vec_{\mathbb{M}}(\widehat{M})\|_1}$ and selecting thresholds $\theta[k]$ that determine the top-k link ranks up to all K link candidates ($\theta[K] = 0$):

$$\begin{aligned} m &= \sum_{k=1}^{K-1} (m \geq \theta[k])(\theta[k] - \theta[k+1]) \\ \Rightarrow D_{\mathbb{M}}(M, r_1) &= \frac{vec_{\mathbb{M}}^T(M) vec_{\mathbb{M}}(\widehat{M})}{\|vec_{\mathbb{M}}(\widehat{M})\|_1} = \int_{-\infty}^{\infty} TP(\theta)P'(\theta)d\theta \end{aligned}$$

where TP and P denote the number of true positive and positive number of thresholded link ranks respectively. At worst, every new positive link after a certain point would be a false positive. Using the big-O notation this can be written as $\frac{\partial FPR(\theta)}{\partial P(\theta)} \in O(1)$ and hence:

$$LinkAUC \in O(D_{\mathbb{M}}(M, r_1)) \quad (5)$$

We next consider the case where discovered ranks form non-overlapping metadata groups, i.e. each node has non-zero rank only for one group. This may happen when query propagation stops before it reaches other metadata groups. Annotating $\widehat{M}_i = r_i r_i^T$, for non-overlapping ranks $r_i \cdot r_j = 0$ for $i \neq j$, we rewrite (1) as $\widehat{M} = \sum_i \widehat{M}_i \Rightarrow vec_{\mathbb{M}}(\widehat{M}) = \sum_i vec_{\mathbb{M}}(\widehat{M}_i)$, similarly to before:

$$LinkAUC \in O\left(\sum_i D_{\mathbb{M}}(M, r_i) vol(\mathbb{M}, r_i) \|r_i\|_1^2\right)$$

This averages group densities and weights them by $vol(\mathbb{M}, r_i) \|r_i\|_1^2$. Hence, when metadata groups are non-overlapping, high LinkAUC indicates high rank density.

Finally, for overlapping metadata groups, LinkAUC involves inter-group links in its evaluation. Since averaging density-based evaluations across groups ignores these links, LinkAUC can be considered an enrichment of rank density in the sense that it bounds it when metadata groups do not overlap but accounts for more information when they do.

3 Experiments

To assess the merit of evaluating node ranks using LinkAUC, we devise a series of experiments where we test a number of different algorithms on several ranking tasks of varying degrees of difficulty across labeled networks. We use the ranks produced by these experiments to compare various unsupervised measures with supervised ones; the latter form the ground truth unsupervised measures need reproduce, but would not be computable if node labels were sparse or missing.

For every network, we start with known binary vectors c_i , whose elements c_{ij} show whether nodes j are members of metadata groups i . We use a uniform sampling process U to withhold a small set of evaluation nodes $eval_i \sim U(c_i, 1\%)$ and edges $(eval_i \times eval_i) \odot M$ that -by merit of their small number- do not significantly affect the ranking algorithm outcomes. We also procure varied length query vectors $q_i \sim U(c_i - eval_i, f)$ that serve as inputs to the ranking algorithms, where their relative size compared to the group is selected amongst $f \in \{0.1\%, 1\%, 10\%\}$. Depending on whether query nodes are adequately many or too few, we expect algorithms to encounter high and low difficulty respectively.

3.1 Networks

Experiments are conducted on three networks; a synthetic one constructed through a stochastic block model [33] and two real-world ones often used to evaluate

metadata group detection; the *Amazon* co-purchasing [34] and the *DBLP* author co-authorship networks. These networks were selected on merit of being fully labeled, hence enabling supervised evaluation to serve as ground truth. They also comprise multiple metadata groups and unweighted edges needed for LinkAUC.

The stochastic block model is a popular method to construct networks of known communities [35,36], where the probability of two nodes being linked is determined by which communities they belong to. Our synthetic network uses the randomly generated 5×5 block probability matrix of Figure 2 with blocks of 2K-5K nodes. The Amazon network comprises links between frequently co-purchased products³ that form communities based on their type (e.g. Book, CD, DVD, Video). We use the 2011 version of the DBLP dataset⁴, which comprises 1.6M papers from the DBLP database, from which we extracted an author network based on co-authorship relations. In this network, authors form overlapping metadata groups based on academic venues (journals, conferences) they have published in. To experiment with smaller portions of query nodes and limit the running time of experiments, we select only the metadata groups with $\geq 5K$ nodes for the real-world networks. A summary of these is presented in Table 1.

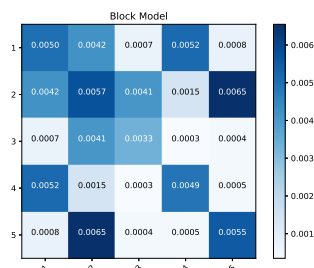


Fig. 2. Stochastic block model used to create the synthetic network.

Network	Nodes	Edges	Groups
Synthetic	15K	0.4M	5
Amazon [34]	0.5M	1.8M	4
DBLP [37]	1.0M	11.3M	52

Table 1. Networks and the number of metadata groups used in experiments.

3.2 Ranking Algorithms

We use both heuristic and established algorithms to rank the relation of network nodes to metadata groups. Our goal is not to select the best algorithm but to obtain ranks with many different methods and then use these ranks to compute the evaluation measures to be compared. The considered algorithms are:

PPR [12,38]. Personalized PageRank with symmetric matrix Laplacian normalization arising from a random walk with restart strategy. It iterates $r_i \leftarrow aD^{-1/2}MD^{-1/2}r_i + (1-a)q_i$, where D is the diagonal matrix of node degrees. Throughout our experiments, we select the well-performing parameter $a = 0.99$.

PPR+Inflation [13]. Adds all neighbors of the original query nodes to the query to further spread PPR.

PPR+Oversampling [39]. Adding nodes with high PPR ranks to the query vector before rerunning the algorithm.

³<https://snap.stanford.edu/data/amazon-meta.html>

⁴DBLP-Citation-network V4 from <https://aminer.org/citation>

HK [40]. Heat Kernel ranks obtained through an exponential degradation filter $r_i = e^{-t} \left(\sum_{k=0}^N \frac{t^k}{k!} (D^{-1/2} M D^{-1/2})^k \right) q_i$. This places higher weights on shorter paths instead of uniformly spreading them across longer random walks. Hence, it discovers denser local structures at the cost of not spreading ranks too much. We selected $t = 5$ and stopped iterations when $(D^{-1/2} M D^{-1/2})^k q_i$ converged.

HPR. A heuristic adaptation of PPR that borrows assumptions of heat kernels to place emphasis on short random walks $r_i \leftarrow \frac{t}{k} a (D^{-1/2} M D^{-1/2} - I) r_i + (1 - a) q_i$, where k is the current iteration, $t = 5$ and $a = 0.99$.

3.3 Measures

The following measures are calculated for the node ranks of metadata groups produced in each experiment. We remind that, when network labels are sparse, supervised measures that serve as the ground truth of evaluation may be inapplicable. Unsupervised measures other than LinkAUC are computed on the training edges, as the sparsity of withheld group members $eval_i$ does not allow meaningful structural scores. LinkAUC, on the other hand is applicable regardless of the evaluation edge set’s sparsity. To avoid data overlap between rank calculation and evaluation, which could overestimate the latter, supervised measures and LinkAUC use only the test group members and edges.

Unsupervised Measures

Conductance - Compares the probability of a random walk to move outside a community vs. to return to it [41]. Using the same probabilistic formulation as for rank density we define rank conductance:

$$\phi_{\mathbb{M}}(M, r) = \frac{r^T (M \odot \mathbb{M})(C - r)}{r^T (M \odot \mathbb{M})r} \quad (6)$$

where $C = 1$ is a max-probability parameter. (Comparisons are preserved for any value.) Lower conductance indicates better community separation.

Gap Conductance - Conductance of binarily cutting the network on the maximal percentage gap between $\frac{r_{ij}}{degree(j)}$ for each community i [42,43]. We use this as an alternative to sweeping strategies [12,13], which took too long to run.

Density - The rank-based extension of density in (4).

LinkAUC - AUC of links ranks calculated through (1), where columns are divided with their maximal value and then each node’s row representation is L2-normalized, making link ranks represent cosine similarity between edge nodes. This is our proposed unsupervised measure.

Supervised Measures (Ground Truth)

NodeAUC - AUC of node ranks, averaged across metadata groups i .

NDCG - Normalized discounted cumulative gain across all network nodes. For this non-parametric statistic, ranks derive ordinalities $ord[j]$ for nodes j (i.e. the highest ranked node is assigned $ord[j] = 1$). For each metadata group i , assigning to nodes j relevance scores of 1 if they belongs to it and 0 otherwise:

$$NDCG_i = \frac{\sum_{j:j \in eval_i} 1/\log_2(ord[j] + 1)}{\sum_{c=1}^{|eval_i|} 1/\log_2(c + 1)} \quad (7)$$

NDCG is usually used to evaluate whether a fixed top-k nodes are relevant to the metadata group. However, we are interested in evaluating the relevant nodes of the whole network and hence we make this measure span all nodes. This makes it similar to AUC in that values closer to 1 indicate that metadata group members are ranked as more relevant to the group compared to non-group members. Its main difference is that more emphasis is placed on the top discoveries.

3.4 Results

In Figure 3 we present the outcome of evaluating different algorithms on the various experiment setups, i.e. tuples of networks, seed node fractions and ranking algorithms. Each point corresponds to a different unsupervised (vertical axes) - supervised (horizontal axes) measure pair calculated for a different experiment setup (i.e. combination of seed node sizes and ranking algorithms) and is obtained by averaging the measures across 5 repetitions of the setup. Unsupervised measures are considered to yield descriptive evaluations when they correlate to supervised ones for the same network (each network is involved in 15 experiment setups arising from the combination of $|f| = 3$ different seed node sizes with one of the 5 different ranking algorithms).

We can see that LinkAUC is the unsupervised measure whose behavior most closely resembles that of the supervised ones. In particular, Table 2 shows that LinkAUC has a strong positive correlation with NodeAUC and a positive correlation with NDCG for all three networks, outperforming the other metrics in all but one experiments. To make sure that these findings cannot be attributed to non-linear relations with other measures, we confirm them using both Pearson and Spearman correlation, where the latter is a non-parametric metric that compares the ordinality of measure outcomes. The slightly weaker correlation of LinkAUC with NDCG can be attributed to the latter’s tendency to place more emphasis on the top predictions, which makes it overstate the correctness of rank quality compared to AUC when the rest of ranks are inaccurate.

Looking at the other unsupervised measures, fuzzy definitions of conductance and density sometimes degrade for higher NodeAUC values. This can be attributed to these metrics measuring local-scale features, which are not always a good indication of the quality of larger metadata groups. It must be noted that gap conductance also exhibits strong correlation with the supervised measures on the real-world networks. However, especially for the synthetic network, it frequently assumes a value of 1 that reflects its inability to discover clear-cut boundaries. This sheds doubt on the validity of using it for evaluating ranks in new networks, since similar structural deficiencies can render it uninformative.

4 Conclusions and Future Work

In this work we proposed a new unsupervised procedure that evaluates node ranks of multiple metadata groups based on how well they predict network edges. We explained the intuitive motivation behind this approach and experimentally

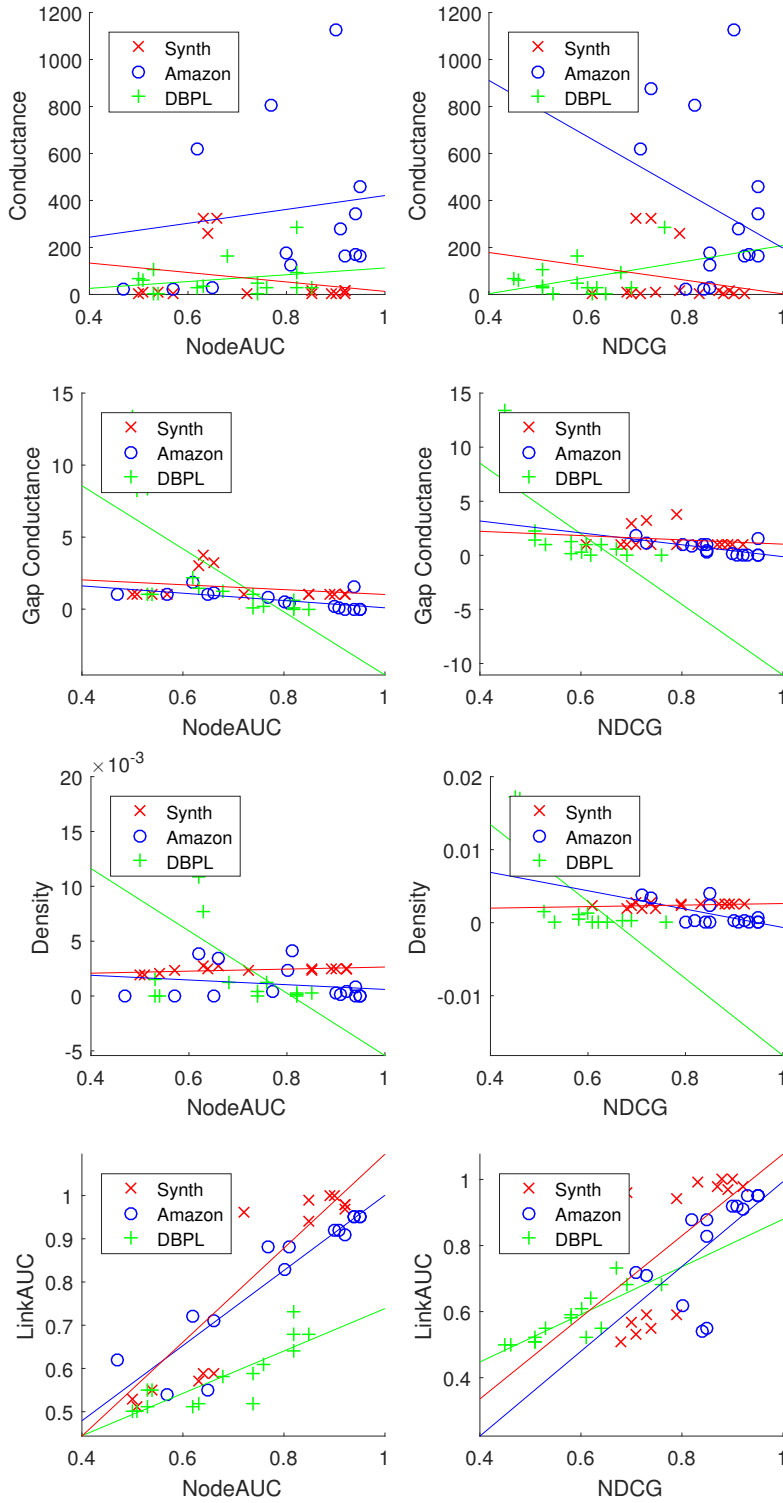


Fig. 3. Scatter plots and least square lines of unsupervised vs. supervised measures. Each point corresponds to a different experiment setup.

	Pearson Correlation			Spearman Correlation		
	Synth	Amazon	DBLP	Synth	Amazon	DBLP
	with NodeAUC					
Conductance	-27%	14%	25%	1%	24%	5%
Gap Cond/nce	-28%	-67%	-70%	-23%	-72%	-88%
Density	58%	-22%	-59%	55%	6%	-45%
LinkAUC	84%	92%	84%	85%	95%	90%
	with NDCG					
Conductance	-23%	-26%	40%	1%	-3%	-9%
Gap Cond/nce	-19%	-69%	-71%	-19%	-68%	-85%
Density	38%	-63%	-74%	45%	-21%	-75%
LinkAUC	56%	65%	84%	71%	85%	88%

Table 2. Correlations between unsupervised and supervised measures. The strongest correlations for each dataset are bolded.

showed that it closely follows supervised rank evaluation across a number of different experiments, many of which are inadequately evaluated by other unsupervised community quality measures. Based on our findings, our approach can be a better alternative to existing rank evaluation strategies in unlabeled networks whose metadata propagation mechanisms are unknown. This indicates that network structure and awareness of multiple metadata groups are two promising types of ground truth that can help evaluate metadata group ranks.

In the future, we are interested in performing experiments across more networks and compare our approach with additional unsupervised measures.

Acknowledgements

This work was partially funded by the European Commission under contract numbers H2020-761634 FuturePulse and H2020-825585 HELIOS.

References

1. S. Fortunato and D. Hric, “Community detection in networks: A user guide,” *Physics Reports*, vol. 659, pp. 1–44, 2016.
2. J. Leskovec, K. J. Lang, and M. Mahoney, “Empirical comparison of algorithms for network community detection,” in *Proceedings of the 19th international conference on World wide web*, pp. 631–640, ACM, 2010.
3. J. Xie, S. Kelley, and B. K. Szymanski, “Overlapping community detection in networks: The state-of-the-art and comparative study,” *Acm computing surveys (csur)*, vol. 45, no. 4, p. 43, 2013.
4. S. Papadopoulos, Y. Kompatsiaris, A. Vakali, and P. Spyridonos, “Community detection in social media,” *Data Mining and Knowledge Discovery*, vol. 24, no. 3, pp. 515–554, 2012.
5. D. Hric, R. K. Darst, and S. Fortunato, “Community detection in networks: Structural communities versus ground truth,” *Physical Review E*, vol. 90, no. 6, p. 062805, 2014.

6. D. Hric, T. P. Peixoto, and S. Fortunato, “Network structure, metadata, and the prediction of missing nodes and annotations,” *Physical Review X*, vol. 6, no. 3, p. 031038, 2016.
7. L. Peel, D. B. Larremore, and A. Clauset, “The ground truth about metadata and community detection in networks,” *Science advances*, vol. 3, no. 5, p. e1602548, 2017.
8. A. Perer and B. Shneiderman, “Balancing systematic and flexible exploration of social networks,” *IEEE transactions on visualization and computer graphics*, vol. 12, no. 5, pp. 693–700, 2006.
9. M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, and A. Arenas, “Ranking in interconnected multilayer networks reveals versatile nodes,” *Nature communications*, vol. 6, p. 6868, 2015.
10. J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney, “Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters,” *Internet Mathematics*, vol. 6, no. 1, pp. 29–123, 2009.
11. A. Lancichinetti, S. Fortunato, and J. Kertész, “Detecting the overlapping and hierarchical community structure in complex networks,” *New Journal of Physics*, vol. 11, no. 3, p. 033015, 2009.
12. R. Andersen, F. Chung, and K. Lang, “Local graph partitioning using pagerank vectors,” in *null*, pp. 475–486, IEEE, 2006.
13. J. J. Whang, D. F. Gleich, and I. S. Dhillon, “Overlapping community detection using neighborhood-inflated seed expansion,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 5, pp. 1272–1284, 2016.
14. C.-C. Hsu, Y.-A. Lai, W.-H. Chen, M.-H. Feng, and S.-D. Lin, “Unsupervised ranking using graph structures and node attributes,” in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pp. 771–779, ACM, 2017.
15. G. Shani and A. Gunawardana, “Evaluating recommendation systems,” in *Recommender systems handbook*, pp. 257–297, Springer, 2011.
16. Y. Wang, L. Wang, Y. Li, D. He, W. Chen, and T.-Y. Liu, “A theoretical analysis of ndcg ranking measures,” in *Proceedings of the 26th annual conference on learning theory (COLT 2013)*, vol. 8, p. 6, 2013.
17. F. Isinkaye, Y. Folajimi, and B. Ojokoh, “Recommendation systems: Principles, methods and evaluation,” *Egyptian Informatics Journal*, vol. 16, no. 3, pp. 261–273, 2015.
18. L. Kowalik, “Approximation scheme for lowest outdegree orientation and graph density measures,” in *International Symposium on Algorithms and Computation*, pp. 557–566, Springer, 2006.
19. M. E. Newman, “Modularity and community structure in networks,” *Proceedings of the national academy of sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.
20. D. Chalupa, “A memetic algorithm for the minimum conductance graph partitioning problem,” *arXiv preprint arXiv:1704.02854*, 2017.
21. L. G. Jeub, P. Balachandran, M. A. Porter, P. J. Mucha, and M. W. Mahoney, “Think locally, act locally: Detection of small, medium-sized, and large communities in large networks,” *Physical Review E*, vol. 91, no. 1, p. 012821, 2015.
22. M. McPherson, L. Smith-Lovin, and J. M. Cook, “Birds of a feather: Homophily in social networks,” *Annual review of sociology*, vol. 27, no. 1, pp. 415–444, 2001.
23. L. Duan, S. Ma, C. Aggarwal, T. Ma, and J. Huai, “An ensemble approach to link prediction,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 11, pp. 2402–2416, 2017.

24. Y. Koren and R. Bell, “Advances in collaborative filtering,” in *Recommender systems handbook*, pp. 77–118, Springer, 2015.
25. D. Liben-Nowell and J. Kleinberg, “The link-prediction problem for social networks,” *Journal of the American society for information science and technology*, vol. 58, no. 7, pp. 1019–1031, 2007.
26. L. Lü and T. Zhou, “Link prediction in complex networks: A survey,” *Physica A: statistical mechanics and its applications*, vol. 390, no. 6, pp. 1150–1170, 2011.
27. A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, “Graph signal processing: Overview, challenges, and applications,” *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.
28. V. Martínez, F. Berzal, and J.-C. Cubero, “A survey of link prediction in complex networks,” *ACM Computing Surveys (CSUR)*, vol. 49, no. 4, p. 69, 2017.
29. J. A. Hanley and B. J. McNeil, “The meaning and use of the area under a receiver operating characteristic (roc) curve.,” *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.
30. S. J. Mason and N. E. Graham, “Areas beneath the relative operating characteristics (roc) and relative operating levels (rol) curves: Statistical significance and interpretation,” *Quarterly Journal of the Royal Meteorological Society*, vol. 128, no. 584, pp. 2145–2166, 2002.
31. S. E. Schaeffer, “Graph clustering,” *Computer science review*, vol. 1, no. 1, pp. 27–64, 2007.
32. R. Görke, A. Kappes, and D. Wagner, “Experiments on density-constrained graph clustering,” *Journal of Experimental Algorithmics (JEA)*, vol. 19, pp. 3–3, 2015.
33. P. W. Holland, K. B. Laskey, and S. Leinhardt, “Stochastic blockmodels: First steps,” *Social networks*, vol. 5, no. 2, pp. 109–137, 1983.
34. J. Leskovec, L. A. Adamic, and B. A. Huberman, “The dynamics of viral marketing,” *ACM Transactions on the Web (TWEB)*, vol. 1, no. 1, p. 5, 2007.
35. K. Rohe, S. Chatterjee, B. Yu, *et al.*, “Spectral clustering and the high-dimensional stochastic blockmodel,” *Annals of Statistics*, vol. 39, no. 4, pp. 1878–1915, 2011.
36. E. Abbe, A. S. Bandeira, and G. Hall, “Exact recovery in the stochastic block model,” *IEEE Trans. on Information Theory*, vol. 62, no. 1, pp. 471–487, 2016.
37. J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su, “Arnetminer: extraction and mining of academic social networks,” in *Proceedings of the 14th ACM SIGKDD int. conf. on Knowledge discovery and data mining*, pp. 990–998, ACM, 2008.
38. P. Lofgren, S. Banerjee, and A. Goel, “Personalized pagerank estimation and search: A bidirectional approach,” in *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pp. 163–172, ACM, 2016.
39. E. Krasanakis, E. Schinas, S. Papadopoulos, Y. Kompatsiaris, and A. Symeonidis, “Boosted seed oversampling for local community ranking,” *Information Processing & Management*, p. 102053, 2019.
40. K. Kloster and D. F. Gleich, “Heat kernel based community detection,” in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1386–1395, ACM, 2014.
41. R. Andersen, F. Chung, and K. Lang, “Local partitioning for directed graphs using pagerank,” *Internet Mathematics*, vol. 5, no. 1-2, pp. 3–22, 2008.
42. C. Borgs, J. Chayes, M. Mahdian, and A. Saberi, “Exploring the community structure of newsgroups,” in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 783–787, ACM, 2004.
43. D. Gleich and K. Kloster, “Seeded pagerank solution paths,” *European Journal of Applied Mathematics*, vol. 27, no. 6, pp. 812–845, 2016.