STRUCTURED EGONET TENSORS FOR ROBUST NODE EMBEDDING

Fatemeh Sheikholeslami and Georgios B. Giannakis,

Department of ECE and Digital Techology Center
University of Minnesota, Minneapolis, MN 55455

ABSTRACT

Recent advances in algorithmic and computational tools have led to an unprecedented growth in data mining over networks. However, partial knowledge of node connectivity (due to privacy concerns or the large number of nodes), as well as incomplete domain knowledge (as in e.g., biological applications), challenge learning tasks over real networks. For robust learning from incomplete data, node embedding over graphs is thus well motivated, and is pursued here by leveraging tensors as multi-dimensional data structures. To this end, a novel tensor-based network representation is advocated, over which node embedding is cast as a structured nonnegative tensor decomposition. The trilinear factorization involved is performed using an alternating least-squares approach. The extracted node embeddings are then utilized to predict the missing links. Performance is assessed via numerical tests on benchmark networks, corroborating the effectiveness and robustness of the proposed technique over incomplete graphs.

Index Terms— Node embedding, tensor decomposition, link prediction.

1. INTRODUCTION

In recent years, complex system analysis has experienced unprecedented growth, thanks to the emergence of graphs as an indispensable tool for data mining and pattern recognition in heterogeneous collections of data. Particularly, in many real-world networks emerging with social media and biological applications, a variety of tasks including regression, classification, clustering, and link prediction, all face challenges related to incomplete data as well as limited computational resources that must be accounted for, when designing efficient and robust learning schemes [1, 2].

Specifically, incomplete knowledge of graph connectivity may arise due to various reasons such as privacy and the large scale of social networks [3, 4]. In addition, incomplete domain knowledge arises in biomedical applications [5, 6]. To this end, tensors as multi-dimensional data structures with increased representational capabilities provide a valuable toolbox that has been investigated for a variety of challenging network analysis tasks such as anomaly detection and dynamic community identification, among others [7, 8, 9, 10].

Furthermore, there has been a growing research interest at the intersection of representation learning and network analysis, where the objective is to learn informative nodal features, or *embeddings* [11, 12]. Such embeddings are then utilized for a variety of tasks such as node-classification or link prediction using mature and popular data vector machines such as logistic regressions, multi-layer perceptrons, or support vector machines. Embedding techniques using random walks, neural networks, and nonnegative matrix factorization are among the recently proposed approaches [13, 14, 15, 16].

The objective of this work is to design an efficient node embedding technique, whose performance is robust in the face of partial knowledge on network connectivity. Inspired by [8], a novel tensor-based node embedding technique is developed, whose factorization provides robust embeddings in the presence of missing connections in the graph. Specifically, a *tensor of egonets* is formed to obtain a reinforced network representation with inherently structured redundancy. This tensor is then factorized using structured polyadic canonical decomposition (CPD), while maintaining sparsity for affordable memory and computational load [17]. One of the CPD factors subsequently yields the sought node embeddings, and is then utilized to predict the missing edges in the graph. Thanks to the reinforced graph representation as well as the multi-dimensional nature of tensors, the proposed method offers considerable improvement, as corroborated through numerical tests.

The rest of the paper is organized as follows. The proposed tensor representation is the subject of Section 2, while tensor-based node embedding along with its solver are presented in Section 3. Section 4 provides numerical tests, and Section 5 concludes the paper.

2. PRELIMINARIES AND TENSOR OF EGONETS

Given a network of N vertices (or nodes) $n \in \mathcal{V}$ with $|\mathcal{V}| = N$, and their edgeset $\mathcal{E}$, let us denote the observed subset of the edgeset as $\mathcal{E}_o \subset \mathcal{E}$, where the set difference $\mathcal{E}_M := \mathcal{E} \setminus \mathcal{E}_o$ corresponds to the missing network edges due to incomplete data. Let us also denote the often sparse binary adjacency matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$ with (i, j)th entry $[\mathbf{W}]_{i,j} := w_{ij} = 1$,

This work has been supported by USA NSF grants 1901134 and 1711471. E-mails: {sheik081,georgios}@umn.edu

if $(i,j) \in \mathcal{E}$, and $w_{ij} = 0$, otherwise. Correspondingly, the observed adjacency matrix $\mathbf{W}_o := \mathcal{P}_\Omega(\mathbf{W})$ sampled through the operator $\mathcal{P}_\Omega(.)$ is defined as

$$[\mathbf{W}_o]_{i,j} := [\mathcal{P}_\Omega(\mathbf{W})]_{i,j} = \begin{cases} 1 & \text{if } (i,j) \in \mathcal{E}_o \\ 0 & \text{otherwise.} \end{cases}$$

Aiming at a reinforced tensor-based network representation, we rely on the so-termed *egonets* of the observed binary network, where the egonet of node n is defined as the subgraph induced by node n and its one-hop observed neighbors $\mathcal{N}_o(n)$, as well as all their connections in edgeset $\mathcal{E}_o$ [7]. Specifically, the egonet of node n can be represented by the induced subgraph $\mathcal{G}^{(n)} := (\mathcal{V}, \mathcal{E}_o^{(n)})$, where $\mathcal{E}_o^{(n)}$ is the edgeset of the links in between nodes $\{n\} \cup \mathcal{N}_o(n)$ in the observed edgeset $\mathcal{E}_o$. Thus, the binary adjacency matrix $\mathbf{W}_o^{(n)} \in \mathbb{R}^{N \times N}$ corresponding to egonet $\mathcal{G}^{(n)}$ has entries

$$w_{o.ij}^{(n)} := \begin{cases} 1 & \text{if } (i,j) \in \mathcal{E}_0^{(n)} \text{and } i,j \in \mathcal{N}_o(n) \cup \{n\} \\ 0 & \text{otherwise.} \end{cases}$$

Subsequently, the *three-way observed egonet-tensor* $\underline{\mathbf{W}}_o \in \mathbb{R}^{N \times N \times N}$, or *EgoTen*, is constructed by stacking the observed sparse egonet adjacency matrices $\mathbf{W}_o^{(n)}$ for all nodes $n \in \mathcal{V}$ in the frontal slabs of $\underline{\mathbf{W}}_o$; that is, $\underline{\mathbf{W}}_{o_{:,:,n}} := \mathbf{W}_o^{(n)}$, where : is a free index that spans its range, in tensor parlance.

3. NODE EMBEDDING

Inspired by node-embedding techniques based on low-rank decomposition of the adjacency matrix, here we propose a joint low-rank decomposition of the observed egonet adjacencies, namely via the well-known canonical polyadic decomposition (CPD) [18]. This is mainly motivated by the fact that the 'structured redundancy' in the tensor of egonets provides a reinforced representation of the network, hence yielding providing with increased robustness, i.e., [8]. We approximate the observed EgoTen by a nonnegative rank-D decomposition

$$\min_{\mathbf{A} \geq \mathbf{0}, \mathbf{B} \geq \mathbf{0}, \mathbf{C} \geq \mathbf{0}} \|\underline{\mathbf{W}}_o - \sum_{d=1}^{D} \mathbf{a}_d \circ \mathbf{b}_d \circ \mathbf{c}_d\|_F^2$$

which, for the n-th slab corresponding to the n-th egonet, yields

$$\mathbf{W}_o^{(n)} \simeq \sum_{d=1}^{D} c_{nd} (\mathbf{a}_d \circ \mathbf{b}_d) \tag{1}$$

where vectors $\mathbf{a}_d$ and $\mathbf{b}_d$ correspond to the d-th columns of factors $\mathbf{A}$ and $\mathbf{B}$, and c_{nd} is the (n,d)-th entry of factor $\mathbf{C}$, while $\circ$ denotes vector outer product. By interpreting the rank of the decomposition as the number of "components" in the network, the joint decomposition yielding (1) can then be interpreted as a weighted sum over the D "components," where $(\mathbf{a}_d \circ \mathbf{b}_d)$ captures the d-th "component structure." This motivates further regularization of the CPD as

$$\{\widehat{\mathbf{A}}, \widehat{\mathbf{B}}, \widehat{\mathbf{C}}\} = \arg\min_{\mathbf{A} \geq \mathbf{0}, \mathbf{B} \geq \mathbf{0}, \mathbf{C} \geq \mathbf{0}} \Big\{ \|\underline{\mathbf{W}}_o - \sum_{d=1}^{D} \mathbf{a}_d \circ \mathbf{b}_d \circ \mathbf{c}_d\|_F^2 + \lambda(\|\mathbf{A}\|_1 + \|\mathbf{B}\|_1) \Big\} \tag{2}$$

in which sparsity of the "components" in terms of constituent nodes is promoted. Finally, the D-dimensional rows of the factor $\mathbf{C} := [\mathbf{c}_1^\top, \mathbf{c}_2^\top, \cdots, \mathbf{c}_n^\top]^\top$ will provide the embedded nodal features in the graph. Optimization in (2) is a trilinear non-smooth block-convex minimization, that can be effectively solved by block-doordinate descent. The following subsection delineates the solver in detail.

3.1. Solver

The proposed minimization in (2) can be effectively handled by block-coordinate descent, where the blocks of the optimization variables are $\mathbf{A}, \mathbf{B}$, and $\mathbf{C}$, yielding the following subproblems.

Subproblem A. At iteration $k+1$, by fixing factors $\mathbf{B}^{(k)}$ and $\mathbf{C}^{(k)}$ at their current values, factor $\mathbf{A}$ is updated by minimizing the nonnegative regularized least-squares (LS) cost as

$$\mathbf{A}^{(k+1)} = \arg\min_{\mathbf{A} \geq \mathbf{0}} \|\mathbf{W}_A - \mathbf{H}_A^{(k)} \mathbf{A}^\top\|_F^2 + \lambda \|\mathbf{A}\|_1 \tag{3}$$

where $\mathbf{W}_A := [\text{vec}(\underline{\mathbf{W}}_{o_{1,:,:}}), \ldots, \text{vec}(\underline{\mathbf{W}}_{o_{N,:,:}})] \in \mathbf{R}^{N^2 \times N}$ is a matricized reshaping of the tensor $\underline{\mathbf{W}}_o$, and

$$\mathbf{H}_A^{(k)} := \left[\mathbf{b}_1^{(k)} \otimes \mathbf{c}_1^{(k)}, \ldots, \mathbf{b}_D^{(k)} \otimes \mathbf{c}_D^{(k)}\right]$$

with $\mathbf{b}_d^{(k)}$ ($\mathbf{c}_d^{(k)}$) denoting column d of $\mathbf{B}^{(k)}$ (respectively $\mathbf{C}^{(k)}$), and $\otimes$ the Kronecker product operator; see also [18].

Subproblem B. Similarly, update $(k+1)$ of factor $\mathbf{B}$ is carried out by minimizing

$$\mathbf{B}^{(k+1)} = \arg\min_{\mathbf{B} \geq \mathbf{0}} \|\mathbf{W}_B - \mathbf{H}_B^{(k)} \mathbf{B}^\top\|_F^2 + \lambda \|\mathbf{B}\|_1 \tag{4}$$

where $\mathbf{W}_B := [\text{vec}(\underline{\mathbf{W}}_{o_{:,1,:}}), \ldots, \text{vec}(\underline{\mathbf{W}}_{o_{:,N,:}})] \in \mathbb{R}^{N^2 \times N}$, and

$$\mathbf{H}_B^{(k)} := \left[\mathbf{a}_1^{(k+1)} \otimes \mathbf{c}_1^{(k)}, \ldots, \mathbf{a}_D^{(k+1)} \otimes \mathbf{c}_D^{(k)}\right]$$

yielding a similar optimization problem as in (3).

Subproblem C. Finally, factor $\mathbf{C}$ is updated by solving

$$\mathbf{C}^{(k+1)} = \arg\min_{\mathbf{C} \geq \mathbf{0}} \|\mathbf{W}_C - \mathbf{H}_C^{(k)} \mathbf{C}^\top\|_F^2 \tag{5}$$

where $\mathbf{W}_C := [\text{vec}(\underline{\mathbf{W}}_{o_{:,:,1}}), \ldots, \text{vec}(\underline{\mathbf{W}}_{o_{:,:,N}})]$ is the matricized version of $\underline{\mathbf{W}}$ along the 3-rd mode, and

$$\mathbf{H}_C^{(k)} := \left[\mathbf{a}_1^{(k+1)} \otimes \mathbf{b}_1^{(k+1)}, \ldots, \mathbf{a}_D^{(k+1)} \otimes \mathbf{b}_D^{(k+1)}\right].$$

Algorithm 1 ADMM solver for regularized nonnegative LS

Input $\mathbf{H}_x, \mathbf{W}_x, \mathbf{X}_{\text{init}}$

Set $\rho = \frac{\|\mathbf{X}_{\text{init}}\|_F^2}{D}, \mathbf{X}^{(0)} = \mathbf{X}_{\text{init}}, \bar{\mathbf{X}}^{(0)} = \mathbf{0}_{N\times D}, \mathbf{Y}^{(0)} = \mathbf{0}_{N\times D}, r = 0$

while $r < I_{\text{max,ADMM}}$ **do**

$$\mathbf{X}^{(r)} = \min_{\mathbf{X}} \|\mathbf{W}_x - \mathbf{H}_x\mathbf{X}^\top\|_F^2 + \rho\|\mathbf{Y}^{(r-1)} - \mathbf{X} + \bar{\mathbf{X}}^{(r-1)}\|_F^2 + \gamma\|\mathbf{X}\|_1$$

$$\bar{\mathbf{X}}^{(r)} = \mathcal{P}_+\Big(\mathbf{X}^{(r)} + \mathbf{Y}^{(r-1)}\Big)$$
$$\mathbf{Y}^{(r)} = \mathbf{Y}^{(r-1)} - \rho(\mathbf{X}^{(r)} - \bar{\mathbf{X}}^{(r)})$$
$$r = r + 1$$

end while

Retrun $\bar{\mathbf{X}}^{(r)}$

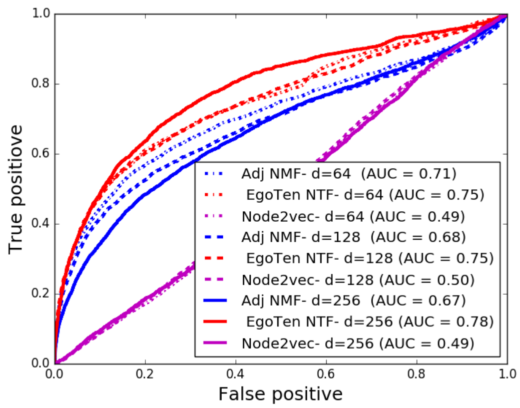

Fig. 1. ROC curve for link prediction with 10% missing links

All the emerging subproblems involve nonnegative (regularized) LS minimization, which can be effectively solved by the altenating direction method of multipliers (ADMM) [19]. To describe the latter, consider the general problem

$$\mathbf{X}^{(k+1)} = \underset{\mathbf{X}\geq\mathbf{0}}{\text{arg min}} \; \|\mathbf{W}_x - \mathbf{H}_x\mathbf{X}^\top\|_F^2 + \gamma\|\mathbf{X}\|_1 \quad (6)$$

where appropriate selection of matrices $\mathbf{X}, \mathbf{W}_x$, and $\mathbf{H}_x$ together with regularization parameter γ specializes (6) to the subproblems in (3), (4), and (5), e.g. setting $\mathbf{X} = \mathbf{C}$, $\mathbf{H}_x = \mathbf{H}_C^{(k)}$, $\mathbf{W}_x = \mathbf{W}_C$, and $\gamma = 0$ yields (5). The ADMM steps for solving (6) are described in pseudocode of Alg. 1. Note that in Alg. 1 updating $\mathbf{X}^{(r)}$ for $\gamma = 0$ boils down to the LS solution that is available in closed form, whereas for $\gamma \neq 0$ the LS cost is augmented with the Lasso regularization term, and can be solved iteratively. To handle sparsity, we have utilized SPLATT to carry out the proposed ADMM-based solvers [17].

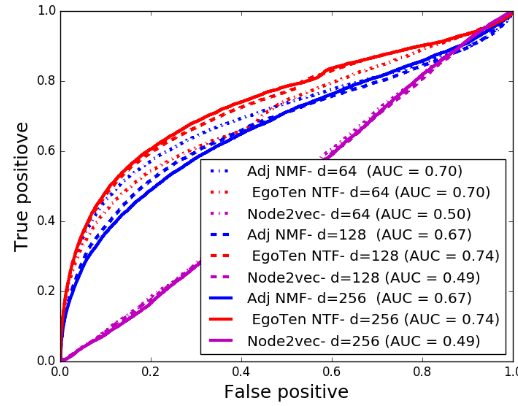

Fig. 2. ROC curve for link prediction with 30% missing links

4. NUMERICAL TESTS

In order to investigate the performance and resilience of the proposed EgoTen-based node embedding, synthetic Lancicchinetti-Fortunatoand-Radicci (LFR) networks [20] are utilized in this section. LFR graphs serve as benchmark networks, in which certain real-world properties, namely power-law distribution for nodal degree and community sizes, are preserved. Such networks are configured by a total number of N nodes, $\bar{d}$ average degree, power-law distribution exponents for degree and community sizes γ_1 and γ_2, and parameter μ to control the community mixing index.

In order to assess the performance under incomplete knowledge of the network connectivity, we have generated networks with $N = 1,000$, $\gamma_1 = 2$, $\gamma_2 = 1$, $\mu = 0.5$, and sampled $m\%$ of the network edges uniformly, while ensuring that all nodes remain connected to the main component. That is, if a node has become disconnected after edge sampling, we randomly retain one of its edges to preserve graph connectivity. Also, parameter λ is set to 0.01.

Performance of the resultant embedded nodal vectors is assessed by utilizing them for link prediction. That is, a possible link between nodes i and j is scored by the inner product of their embedded nodal vectors in the corresponding D dimensional space, i.e., $s_{i,j} := \mathbf{c}_i^\top \mathbf{c}_j$, and the links with higher scores $s_{i,j}$ are considered more likely to appear (or have been missing). By sorting the non-existing edges accordingly to their link scores $s_{i,j}$ in the observed graph in a decreasing order, and comparing with the ground truth edge-set, the receiver operating characteristic curve (ROC) and the area-under-curve (AUC) figures of merit are obtained.

The tests are carried out over LFR networks with $\bar{d} = 50$, with $m = 10\%$, 30% and 50% missing edges, and the ROC curves for four random realizations are depicted in Figures 1-4, and also compared with nonnegative matrix decomposition using the adjacency matrix as well as Node2Vec embedding [13] with $D = 256$. Nonnegative tensor- and matrix-based

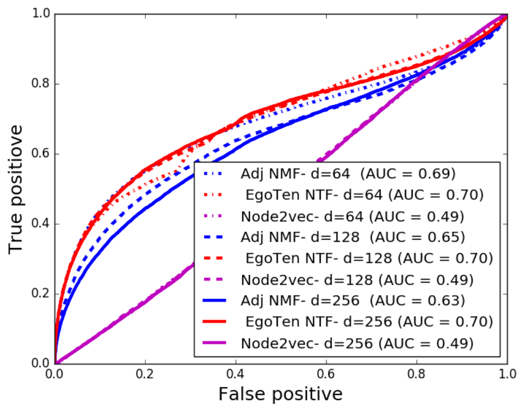

Fig. 3. ROC curve for link prediction with 50% missing links

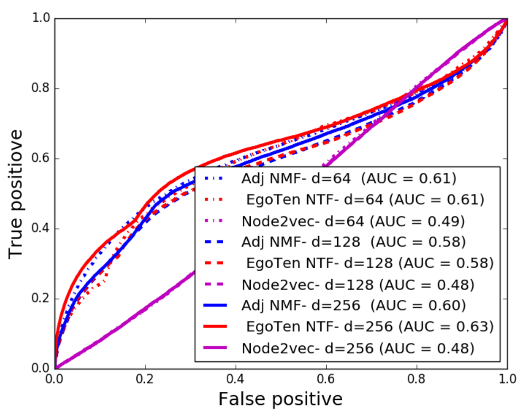

Fig. 4. ROC curve for link prediction with 70% missing links

node embedding methods are denoted as NTM and NMF in the figures. Furthermore, Tables I and II also report the AUC for NMF and NTM methods averaged over 20 network realizations with average degree $\bar{d} = 20$ and 50, respectively, corroborating the trend in the curves of Figures 1-4, for various values of m and D.

As the results demonstrate, under incomplete network knowledge, utilizing the reinforced structure of the proposed Egonet tensor endows the resultant node embeddings with increased robustness, leading to a resilient performance in terms of AUC as the missing ratio increases. Furthermore, the performance of nonnegative matrix factorization deteriorates more by increasing the missing ratio m, and Node2Vec performs poorly in incomplete networks for link prediction. It is further observed that tensor-based embeddings perform better as the embedding dimension D grows, while matrix-based embeddings provide smaller or no improvement.

Table 1. ROC-AUC for link prediction for various ratios of missing links ($m\%$) as well as embedding dimension D for LFR networks with $\bar{d} = 50$ averaged over 20 realizations.

	m%	D=64	D=128	D=256
Egonet	10	0.76 ± 0.02	0.76 ± 0.02	0.76 ± 0.03
Adj.	10	0.71 ± 0.01	0.70 ± 0.02	0.71 ± 0.02
Egonet	30	0.71 ± 0.01	0.72 ± 0.02	0.72 ± 0.02
Adj.	30	0.67 ± 0.01	0.67 ± 0.01	0.67 ± 0.02
Egonet	50	0.65 ± 0.01	0.65 ± 0.02	0.66 ± 0.01
Adj.	50	0.63 ± 0.02	0.62 ± 0.01	0.63 ± 0.01

Table 2. ROC-AUC for link prediction for various ratios of missing links ($m\%$) as well as embedding dimension D for LFR networks with $\bar{d} = 50$ averaged over 20 realizations.

	m%	D=64	D=128	D=256
Egonet	10	0.73 ± 0.01	0.75 ± 0.01	0.77 ± 0.02
Adj.	10	0.73 ± 0.01	0.70 ± 0.01	0.68 ± 0.02
Egonet	30	0.71 ± 0.01	0.73 ± 0.01	0.74 ± 0.01
Adj.	30	0.70 ± 0.01	0.68 ± 0.02	0.68 ± 0.01
Egonet	50	0.67 ± 0.01	0.68 ± 0.01	0.69 ± 0.01
Adj.	50	0.67 ± 0.02	0.65 ± 0.01	0.63 ± 0.01

5. CONCLUSIONS

In this work, a novel tensor-based node embedding technique is introduced for network analytics relying on incomplete data. A tensor of egonets is formed to offer a reinforced network representation, which is subsequently factorized using a structured canonical polyadic decomposition. The sought embedding is then obtained from one of the factors, and can be utilized to predict the missing edges in the graph. Thanks to the multi-dimensional nature of the tensors, as well as the reinforced graph representation with carefully induced structured redundancy, the novel embedding approach offers considerable improvement in link prediction tasks.

Related tensor-based embedding approaches can be beneficial for other learning tasks, including (semi)-supervised classification and clustering. Exploring these directions and carrying out extensive tests over real-world networks are among the envisioned future directions of this work.

6. REFERENCES

[1] P. Wang, B. Xu, Y. Wu, and X. Zhou, "Link prediction in social networks: the state-of-the-art," *Science China Information Sciences*, vol. 58, no. 1, pp. 1–38, 2015.

[2] L. Getoor and C. P. Diehl, "Link mining: A survey,"

Acm SIGKDD Explorations Newsletter, vol. 7, no. 2, pp. 3–12, 2005.

[3] Y. R. Lin, Y. Chi, S. Zhu, H. Sundaram, and B. L. Tseng, “Analyzing communities and their evolutions in dynamic social networks,” *ACM Trans. on Knowledge Discovery from Data*, vol. 3, no. 2, p. 8, Paris, France, April 2009.

[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, Beijing, China, 2012.

[5] P. Du, J. Gong, E. S. Wurtele, and J. A. Dickerson, “Modeling gene expression networks using fuzzy logic,” *IEEE Trans. on Systems, Man, and Cybernetics, Part B*, vol. 35, no. 6, pp. 1351–1359, Dec 2005.

[6] J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Miezin, B. L. Schlaggar, and S. E. Petersen, “Functional network organization of the human brain,” *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.

[7] L. Akoglu, M. McGlohon, and C. Faloutsos, “Oddball: Spotting anomalies in weighted graphs,” *Advances in Knowledge Discovery and Data Mining*, pp. 410–421, Washington, DC, USA, 2010.

[8] F. Sheikholeslami and G. B. Giannakis, “Identification of overlapping communities via constrained egonet tensor decomposition,” *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5730–5745, Nov. 2018.

[9] M. Araujo, S. Papadimitriou, S. Günnemann, C. Faloutsos, P. Basu, A. Swami, E. E. Papalexakis, and D. Koutra, “Com2: Fast automatic discovery of temporal (comet) communities,” *Pacific-Asia Conf. on Knowledge Discovery and Data Mining*, pp. 271–283, Tairan Taiwan, Aug. 2014.

[10] F. Huang, U. Niranjan, M. U. Hakeem, and A. Anandkumar, “Online tensor methods for learning latent variable models,” *J. of Machine Learning Research*, vol. 16, pp. 2797–2835, 2015.

[11] H. Cai, V. W. Zheng, and K. C.-C. Chang, “A comprehensive survey of graph embedding: Problems, techniques and applications,” *Trans. on Knowledge and Data Engineering*, vol. 30, no. 9, pp. 1616–1637, 2018.

[12] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, “Translating embeddings for modeling multirelational data,” *Advances in Neural Inf. Proc. Systems*, vol. 2, p. 27872795, Lake Tahoe, CA, 2013.

[13] A. Grover and J. Leskovec, “Node2vec: Scalable feature learning for networks,” *Proc. of ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining*, p. 855864, San Francisco, CA, 2016.

[14] A. Tsitsulin, D. Mottin, P. Karras, and E. Müller, “Verse: Versatile graph embeddings from similarity measures,” *Proc. of the World Wide Web Conf.*, p. 539548, Lyon, France, 2018.

[15] J. Qiu, Y. Dong, H. Ma, J. Li, K. Wang, and J. Tang, “Network embedding as matrix factorization: Unifying deepwalk, line, pte, and node2vec,” *Proc. of Intl. Conf. on Web Search and Data Mining*, p. 459467, Los Angeles, CA, 2018.

[16] D. Berberidis and G. B. Giannakis, “Node embedding with adaptive similarities for scalable learning over graphs,” *IEEE Transactions on Knowledge and Data Engineering*, to appear October 2019.

[17] S. Smith and G. Karypis, “SPLATT: The Surprisingly ParalleL spArse Tensor Toolkit,” http://cs.umn.edu/~splatt/.

[18] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.

[19] K. Huang, N. D. Sidiropoulos, and A. P. Liavas, “A flexible and efficient algorithmic framework for constrained matrix and tensor factorization,” *IEEE Trans. on Signal Processing*, vol. 64, no. 19, pp. 5052–5065, 2016.

[20] A. Lancichinetti, S. Fortunato, and F. Radicchi, “Benchmark graphs for testing community detection algorithms,” *Physical Review E*, vol. 78, no. 4, p. 046110, 2008.