MEDINFO 2017: Precision Healthcare through Informatics
A.V. Gundlapalli et al. (Eds.)

doi:10.3233/978-1-61499-830-3-422

Visualizing Article Similarities via Sparsified Article Network and Map Projection for Systematic Reviews

Xiaonan Ji[1, 2], Raghu Machiraju[2], Alan Ritter[2], Po-Yin Yen[1]

[1] Department of Biomedical Informatics, [2] Department of Computer Science and Engineering,
The Ohio State University, Columbus, OH, USA

Abstract

Systematic Reviews (SRs) of biomedical literature summarize evidence from high-quality studies to inform clinical decisions, but are time and labor intensive due to the large number of article collections. Article similarities established from textual features have been shown to assist in the identification of relevant articles, thus facilitating the article screening process efficiently. In this study, we visualized article similarities to extend its utilization in practical settings for SR researchers, aiming to promote human comprehension of article distributions and hidden patterns. To prompt an effective visualization in an interpretable, intuitive, and scalable way, we implemented a graph-based network visualization with three network sparsification approaches and a distance-based map projection via dimensionality reduction. We evaluated and compared three network sparsification approaches and the visualization types (article network vs. article map). We demonstrated the effectiveness in revealing article distribution and exhibiting clustering patterns of relevant articles with practical meanings for SRs.

Keywords:

Information Storage and Retrieval; Data Display.

Introduction

Research findings in the biomedical literature are used to guide and improve clinical practice. Systematic reviews (SRs) summarize evidence drawn from high-quality and up-to-date studies to inform clinical decisions and are considered the preferred source of Evidence-based Practice (EBP) [1]. However, the number of studies published each week is over 12,000, including more than 300 randomized trials [2].

During the SR process, researchers conduct an exhaustive literature search and appraise the best evidence to answer a clinical question. With the overwhelming volume of studies (aka. articles), the article screening process has become the most burdensome aspect and often causes information overload [3]. Typically, an exhaustive literature search would yield hundreds or thousands of articles. SR researchers need to spend weeks or even months screening articles to identify relevant ones for inclusion. In general, only 2% to 30% of searched articles are included for evidence synthesization, which means researchers spend most of their efforts excluding irrelevant studies [4].

Existing studies have shown that automatic article classification with supervised machine learning (ML) is a valuable tool to facilitate the identification of relevant articles for SRs [4; 5]. However, such approaches have limited generalization to new SRs due to the dependency on prior supervised training data or manual annotations by domain experts. In our previous studies [6; 7], we demonstrated using established article similarities to assist in article screening for SRs in an unsupervised or semi-supervised manner. Article similarities were established in a feature space derived from several article elements i.e., title and abstract. Our approach achieved competitive performance in reducing SR workloads, and is highly generalizable [6]. Article similarities were also shown to capture article distribution (the structure of an article collection) and the clustering of relevant articles based on their strong similarities [7].

To extend the utilization of article similarities, we proposed to delve into the visualization of article similarities, and compared the effectiveness of different visualization approaches in revealing article distribution and clustering patterns. Under the notion of information visualization [8], articles are represented as visual elements, and their similarities are encoded by connections or visual channels. As pictures can provide more information with less clutter in less space [9], transforming an article collection into graphical representations can enable human's insights into article distribution and clusters of similar articles. With the "visible article distribution", SR researchers can identify studies of interest more efficiently. While there are many visualization approaches [8], we considered that an effective visualization should display articles along with their similarities in a human interpretable, attribute intuitive, and spatial scalable manner [10]. In this study, we focused on two types of visualization: graph-based network visualization (article network) and distance-based (aka. geometry-based) map projection (article map). Other visualization types, such as adjacency matrix, hypergraph, and circular graph, were not included because of their limited structural analytics or spatial scalability.

Figure 1. ADHD article

For graph-based network visualization, we employed the node-link diagram with a force-directed layout to draw an article network, where articles are represented as nodes, and similarities are represented as weighted edges. Network visualization enables the exploration of graph topology and graph-based algorithms for advanced analytics [11]. However, an article network is almost a complete network due to the existence of non-zero similarities between most article pairs retrieved in an SR. Direct visualization of such a network is limited by human perception and important structural patterns are inaccessible, because of the extreme clutter presentation ("hairball") [12]. To provide an effective network visualization, we implemented three network sparsification methods to reduce the network size via edge sampling [13; 14], expecting to preserve edges bearing important conceptual or structural information. The three network sparsification methods include 1) the established article similarity (AS) that predominantly retains edges for strong article similarities, 2) the derived algebraic distance (AD), first proposed by Chen 2011 [15] and used in John 2016 [13], values edges within neighborhoods and tends to preserve a network's local structure, and 3) the derived local degree (LD), first proposed by Lindner 2015 [14], values edges leading to hub nodes and tends to preserve the global structure. Both AD and LD have been shown to result in effective network sparsification [13; 14].

For distance-based map projection, we utilized the t-distributed Stochastic Neighbor Embedding (t-SNE) [16] to generate an article map with article features. The t-SNE technique converts the high-dimensional features into a matrix of pairwise similarities and visualizes it by projecting article points into a two-dimensional space, where article similarities are encoded as their spatial positions. t-SNE reveals local structures of the data and some important global structures such as clusters [16], though topological properties are not available in this map projection.

In summary, to effectively visualize article similarities, the purposes of the study are to 1) compare three network sparsification approaches, AS, AD, and LD; and 2) compare two types of visualization, graph-based network visualization (article network) and distance-based map projection (article map).

Methods

Dataset

We used publicly available data from 15 completed SRs [4; 17] produced by the Drug Effectiveness Review Project team (DERP). These SRs consist of article collections with coded decisions (inclusion or exclusion), which served as the gold standard to evaluate our performance. The size of these SRs ranges from 310 articles to 3465 articles (1249 articles on average), with the full-text level inclusion rate ranges from 0.55% to 27.04% (7.67% on average). We considered the full-text level included articles as relevant articles.

Article Features

Article similarities were established by article features. We generated (lexical) article features from several article elements, which were standardized by MEDLINE, including title (TI), abstract (AB), MeSH (MH), publication type (PT) and author (AU) [6]. We preprocessed the free-text of TI and AB by removing stop words, and stemming the remaining words with the classic Porter Stemmer. For MH, PT and AU, we used the exact strings as they were already standardly encoded. With the bag-of-words approach, we recorded the term frequency of each unigram word (TI and AB) and multi-word string (MH, PT and AU), and generated a feature vector for each article in the feature space. With article features, we calculated the article similarities to generate article networks. Alternatively, t-SNE calculated its own article similarities to generate article maps.

Article Network

An article network $G(V, E)$ consists of a set of nodes ($v \in V$) representing articles and a set of weighted edges ($e \in E$) representing article similarities between corresponding endnodes. We calculated article similarities (Euclidean distances) from article features using Cosine similarity. The resulting similarity ranges from 0 to 1 for each article element. We used an equally weighted sum of the five element similarities as the final article similarity for the edge weight, ranging from 0 to 5.

Network Sparsification

An article network is almost a complete network due to the existence of non-zero similarities between most article pairs in an SR. Visualizing such a network is meaningless as it is referred to as a "hairball". On average, the article networks of 15 DERP SRs have approximately 1,200,000 edges. Figure 1 illustrates an article network using an SR, Attention Deficit Hyperactivity Disorder (ADHD), which has 851 articles.

To overcome the limitation of visualizing an almost complete network, we implemented three network sparsification methods, AS, AD, and LD, with an expectation to reduce the number of edges, but preserve edges with conceptual or structural importance. The sparsification process consists of three steps: (1) network pre-pruning, (2) edge scoring, and (3) edge sampling.

Step 1. Network pre-pruning. In our experiments, we found that directly calculating AD or LD edge scores using neighborhood information in an almost complete network resulted in indiscriminate edge scores as most nodes share similar neighbors. Therefore, we initiated a relaxed pre-pruning step to keep the top neighbors (edges) based on edge weights for each node. We tested a series of pre-pruning parameters from 50% to 5%, and found that keeping the top 10% edges led to a better balance of avoiding the mass of trivial edges and retaining important edges. At this step, the edge number has been reduced to 176,492 on average. We then used this coarsely pruned 10% network as the *baseline network*. The 10% baseline network, although it sounds satisfying, is still insufficient because the pre-pruning only cuts off very trivial edges. As most SRs contain more than 3,000 articles (some are more than 10,000) and result in a quadratic number of edges, network sparsification is needed to retain the most important edges. Thus in the following steps we further aggressively reduced the number of edges.

Step 2. Edge scoring. With the baseline network, we calculated edge scores to further capture the edge importance from different perspectives using AS, AD, and LD.

Article Similarity (AS) To aggressively retain edges corresponding to strong article similarities, we directly used the established article similarities as edge scores. In other words, the edge weights were used as edge scores.

Algebraic Distance (AD) AD was proposed to preserve strong connections in terms of local structures [13; 15]. It generalizes the idea of estimating the Jaccard coefficient for neighborhoods through lazy random walks to determine the strength of connections of the edges. Specifically, nodes with similar neighbors are considered strongly connected. With the AD approach, these nodes converge to similar values via information propagation within the neighborhood and lead to high AD edge scores between them. Thus, edges with high AD scores represent strong local connections. Algorithm 1 shows the computation of AD. We ran multiple rounds (R=20) to obtain synthesized results.

Algorithm 1 *Computing algebraic distance (AD) [15]*

Input: Parameterω (ω = 0.5), weighted adjacency Matrix (for weights w_{ij} and neighborsN_i), and randomly initialized vector $x^{(0)}$ with |V| elements.
For k = 1, 2, do

$$\forall i \in V\ x_i^{(k)} \leftarrow \omega x_i^{(k-1)} + (1-\omega)\frac{\sum_{j\in N_i} w_{ij} x_j^{(k-1)}}{\sum_{j\in N_i} w_{ij}}$$

$$\forall ij \in E\ s_{ij}^{(k)} = \left| x_i^{(k)} - x_j^{(k)} \right|$$

End for

Local Degree (LD) LD was proposed to emphasize "hub" nodes, which are nodes with relatively high degrees [14]. The hub nodes and the connections to the hubs are important to present a network's global structure. Because LD used the unweighted degree, in our study, we extended it to the weighted degree for our weighted article networks. For each node, we scored an incident (associated) edge based on the weighted degree of the other endnode. The LD approach assigns high scores to edges that lead to the hub nodes, and preserves the network "hub backbone".

Step 3. Edge sampling. After the edge scores were calculated, we sampled the edges based on the edge scores. For each nodev ∈ V, we included the top $\lfloor degree(v)^{e} \rfloor$ edges sorted by edge scores in descending order, where $degree(v)$ is the degree of node v, and e ($0 \leq e \leq 1$) controls the strength of sampling (filtering). We ensured that at least one incident edge was kept for each node. In this study, we used a sparsification parameter $e = 0.5$ to preserve at least $\lfloor degree(v)^{0.5} \rfloor$ edges for each node.

Force-directed Graph Drawing

We used a force-directed algorithm to draw sparsified article networks in an aesthetically pleasing way in a two-dimensional space. As a spatial layout, it places nodes and edges by simulating a physical system. When the system comes to a mechanical equilibrium state, the pairwise geometric distance between the drawn nodes matches the graph theoretic pairwise distance. Specifically, similar article nodes tend to aggregate together while dissimilar article nodes are drawn apart. We implemented the algorithm in Gephi with the built-in Force Atlas layout [18].

Article Map

In an article map, articles are represented as a set of points and their similarities are explicitly encoded as spatial positions of article points. We used t-SNE [16] to generate article maps by providing t-SNE the established article feature vectors. t-SNE establishes article similarities (Euclidean distances) in the high-dimensional feature space, and creates an article map by projecting article similarities down to a two-dimensional space.

t-SNE

t-SNE [16] is a technique for visualizing similarity data by embedding high dimensional data into a space of two or three dimensions. The resulting visualization is considered a map with data point distribution or a scatter plot. The t-SNE approach retains data structures by keeping similar data points close together while pushing dissimilar points far apart. It can also reveal important structures such as clustering.

As a non-linear algorithm for dimensionality reduction, t-SNE establishes high-dimensional Euclidean distances (similarities) between data points and converts them into a probability distribution. In the low-dimensional space, map points are placed as counterparts with a similar probability distribution. Gradient descent is used to minimize the divergence between the two distributions with respect to map points' spatial positions. It is worth mentioning that the gradient of the cost function can be interpreted as physical forces between map points just like the force-directed graph drawing for networks. t-SNE has been shown to create higher-quality visualizations than linear methods (i.e. PCA and MDS) and other nonlinear methods (i.e. SNE and Isomap) [16]. We implemented t-SNE in MATLAB.

Evaluation

Network Properties Evaluation

We evaluated article networks based on the network properties in a network structure, including graph diameter, clustering coefficient, communities, and modularity. *Graph diameter* is the length of the shortest path between the most distanced nodes. A smaller diameter indicates a stronger concentration of a graph. *Clustering coefficient* measures the degree to which nodes tend to cluster together. Nodes with a higher clustering coefficient have higher transitivity in the neighborhood. *Communities* are subsets of nodes that are internally densely connected but externally sparsely connected. *Modularity* is designed to measure the strength of division of a network into communities. A high modularity corresponds to a better community structure. We used the Louvain method [19] for community detection which is proven to provide high-quality results. A graph's community structure and modularity also reflect its local structure with respect to intra-community connections. In addition, a graph's global structure can be reflected by the diameter, averaged clustering coefficient, and the number of communities. However, these topological properties were unavailable in article maps.

Clustering Patterns Evaluation

We evaluated the clustering patterns on both article networks and article maps. For article networks, we utilized communities detected by the Louvain method. For article maps, we applied k-means clustering to identify clusters based on the 2-dimensional map computed by t-SNE. For convenience, we used *set* to refer to *community* and *cluster*. We evaluated the clustering patterns of relevant articles that have been identified in the completed DERP SR reports (external criteria). Because of the highly-imbalanced dataset with only 0.55%~27.04% (7.67% on average) of relevant articles, we identified the *set* that contains at least 10% of relevant articles as *dominant set*. We examined the coverage (*recall*) and proportion (*precision*) of relevant articles in all dominant sets, and calculated the balanced F-measure (*F1 score*). This was inspired by the classic measure of clustering quality that evaluates how well the clustering matches the gold standard classes, and interprets the clustering as a series of decisions. Other measures such as the purity, normalized mutual information, and Rand index, are not suitable for the highly-imbalanced dataset.

Results

Network Properties

We reported the network properties of the baseline network and the sparsified networks with the averaged results of the 15 DERP reports (Table 1). Because edge sampling was implemented based on individual nodes, the number of edges after the sampling was not the same when using different sparsification methods. The clustering coefficient was also the average of all article nodes for each SR.

Table 1- Preservation of network properties

	Base-line	**AS**	**AD**	**LD**
Edge Number	176,492	12,706	13,323	15,722
Diameter	3	6	5	4
Clustering Coefficient	0.4038	0.2581	0.1169	0.3806
Modularity	0.4103	0.6231	0.4336	0.3860
Community Number	5	10	7	5

As shown in Table 1, the number of edges was significantly reduced from 176,492 to approximately 13,000 after sparsification. All sparsification methods resulted in increased diameters and decreased clustering coefficients because of the removal of most edges. However, they also showed differences. AS brought the highest modularity with a larger number of communities, but altered the baseline graph diameter and clustering coefficient to a greater extent. Similarly, AD altered the baseline diameter and clustering coefficient, but provided a slight gain in modularity. LD retained a similar graph diameter, clustering coefficient, modularity, and community number compared to the baseline. In summary, AS resulted in a better community structure; AD tended to retain the baseline local structure (with slightly higher modularity); LD performed the best in preserving the baseline global structure.

Clustering Patterns

We examined the clustering patterns of relevant articles in article networks and article maps. The optimized number of communities in an article network was determined by the *Louvain method* (default resolution setting). For the *k-means* clustering in an article map, we applied the *knee (elbow)* *method* to identify a proper value range for the number of clusters, *k*. We found the resulting range approximately aligned to the number of communities detected by the Louvain method in AS networks. Thus for each SR, we had *k* equal to the number of communities in the corresponding AS network.

In Table 2, we reported the total number of *sets* (*communities* or *clusters*), the number of *relevant sets* that contain at least one relevant article, and the number of *dominant sets* that contain at least 10% of relevant articles. We also reported the overall *size* of all dominant sets by calculating the ratio of articles contained by the dominant sets. We calculated the corresponding *recall*, *precision*, and *F1 score* regarding the relevant articles in all dominant sets. Again, all the results were averaged from the 15 DERP SR reports.

Table 2- Article distribution and clustering of relevant articles

	Article Networks				Article Map
	Base-line	AS	AD	LD	
Total Set #	5	10	7	5	10
Relevant Set #	4	6	5	4	8
Dominant Set #	3	3	3	3	3
Dominant Size	63.10%	35.13%	56.35%	65.59%	31.36%
Recall	94.87%	83.16%	90.14%	92.99%	76.45%
Precision	10.84%	17.53%	12.40%	10.08%	17.36%
F1 Score	0.1823	0.2623	0.1983	0.1713	0.2618

As shown in Table 2, the AS network and article map generated by t-SNE had the largest number of sets (10) and relevant sets (6 and 8), but the number of dominant sets was only 3. Their dominant sets covered 83.16% and 76.45% of relevant articles with a size of 35.13% and 31.36% of entire articles. Both of them had lower recalls but the highest precisions and F1 scores (0.2623 and 0.2618). They achieved a good quality of clustering relevant articles by decomposing articles into finely separated sets. The LD network behaved similarly to the baseline network, with the relevant articles spreading into coarsely divided sets. With the highest recalls, 3 out of 5 sets acted as dominant sets and covered over 90% of relevant articles. However, their precisions and F1 scores (0.1713 and 0.1823) were lower than others because the conservative discrimination. The AD network brought moderate performance in recall, precision, and F1 (0.1983). Specifically, 90.14% of relevant articles were covered by dominant sets, with a size of 56.35%.

In summary, network sparsification led to a more recognizable network structure by distributing articles into distinct sets. According to the F1 score, the quality of clustering relevant articles was improved with AS and AD. In addition, the AS network and article map had the highest precisions and F1 scores, but the AS network had a better recall than the article map. Overall, the sparsified article networks and the article map provided a more interpretable distribution and clustering patterns.

Illustration of Visualization

We illustrated the above mentioned results with visualization using an SR report, ADHD, as an example. The ADHD report has a total of 851 articles: 20 were included at the full-text level (relevant articles), 64 were only included at the title/abstract level, and 767 were excluded at the title/abstract level.

Figure 2 shows article networks before and after sparsification. The baseline network had clutter presents that limited human perception without explicit structures. All sparsified networks provided more interpretable structures and revealed the clustering of relevant articles (green nodes). Specifically, AS provided the most manifest community structure with meticulous separations, where relevant articles were highly concentrated. AD retained local connections and led to densely connected neighborhoods. LD preserved the hub backbone structure by concentrating nodes towards hubs and forming bridges among hubs. Figure 3 shows an article map generated by t-SNE. We observed the clustering patterns of relevant articles (green points).

In Figure 4, we illustrated communities in the AS network and clusters in the article map. Article nodes (points) were colored by communities (clusters). Dominant sets of relevant articles were marked by green rectangles, which further demonstrated the effective clustering of relevant articles.

Discussion

Sparsification Schemes: AD and LD were applied to sparser networks (i.e. social networks and citation networks) in early works [13; 14]. To our knowledge, we were the first to apply AD and LD to article network sparsification. Due to the densely-connected nature of article networks, a relaxed pre-pruning step was used. In this study, we found that AD retained the local network structure, LD preserved the global network structure (also supported by other works [13; 14]), and AS performed the best in revealing the community structure. In addition, considering the clustering of relevant articles, AD and LD had lower precisions, but higher recalls, resulting from their integration of the network structure; while AS had a lower recall, but a higher precision because it aggressively concentrated relevant articles. Another encouraging finding was that by keeping only 1-3% of edges (13,000 edges on average) from the original networks (1,200,000 edges on average), we can reveal meaningful network structures and important clustering patterns.

Article Network vs. Article Map: Article networks sparsified by AS and articles maps generated by t-SNE achieved similar results in revealing article distribution and clustering patterns. Both of them aggregated the majority of relevant articles into finely separated set(s). While the AS network had a slightly lower precision than the article map, it had a higher recall which is important for SRs. For article networks, a sparsification process is needed to eliminate clutter presents; but we were able to explore graph topology and apply graph-based algorithms for advanced visual analytics, such as community detection and graph traversal. Article maps are created by t-SNE or other dimensionality reduction algorithms that can handle the crowded article feature space; but topological analysis is not available. The same article feature generation step was applied to the network visualization and the map projection process. Future investigation could include a user study to gather feedback regarding the selection of visualization approaches for SRs.

Significance for SRs and Future Work: Effective visualization of article similarities has practical and significant implications for SRs. By revealing the structure and distribution of articles, SR researchers can conduct SRs in a more effective and efficient manner. Consider a scenario of screening hundreds of articles in the ADHD report, researchers can rapidly identify relevant articles by screening only a small percent of articles once the dominant sets are located. In our future work, we will further investigate the identification of the dominant set. We also plan to provide visual feedback on the selection of article features, and integrate advanced semantic features to improve the similarity calculation.

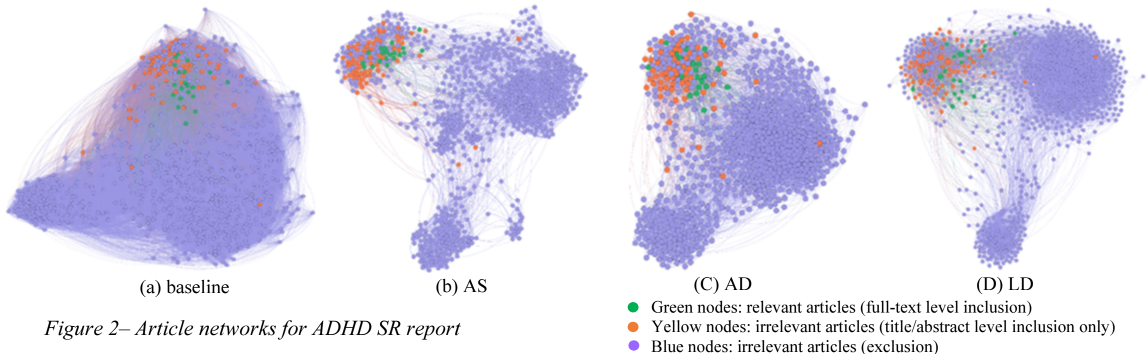

Figure 2– Article networks for ADHD SR report

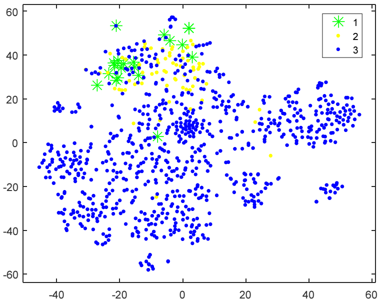

Figure 3– Article map for ADHD SR report

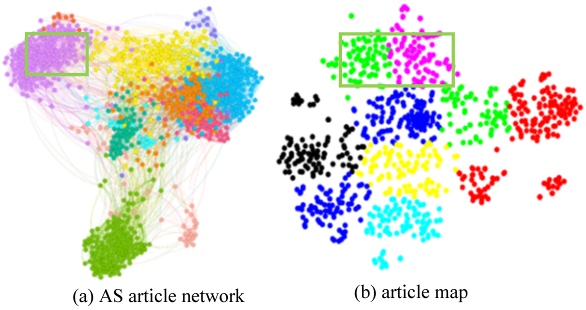

Figure 4– Dominant sets of relevant articles (ADHD)

Conclusions

We visualized article similarities with sparsified article networks and article maps. We demonstrated the effectiveness in revealing meaningful articles structures, and exhibiting clustering of relevant articles in an intuitive and human interpretable manner. Effective visualization of article similarities has practical meanings to facilitate article screening for SR researchers.

Acknowledgements

This work was supported by the Agency for Healthcare Research and Quality (AHRQ), R03HS025047-01.

References

[1] Oxford Centre for Evidence-based Medicine - Levels of Evidence. http://www.webcitation.org/6AZnaE7iG
[2] P.P. Glasziou, Information overload: what's behind it, what's beyond it, *Medical Journal of Australia* **189** (2008).
[3] S. Majid, S. Foo, B. Luyt, X. Zhang, Y.-L. Theng, Y.-K. Chang, and I.A. Mokhtar, Adopting evidence-based practice in clinical decision making: nurses' perceptions, knowledge, and barriers, *Journal of the Medical Library Association (2011)*.
[4] A.M. Cohen, W.R. Hersh, K. Peterson, and P.-Y. Yen, Reducing workload in systematic review preparation using automated citation classification, *Journal of the American Medical Informatics Association* **13** (2006), 206-219.
[5] S. Matwin, A. Kouznetsov, D. Inkpen, O. Frunza, and P. O'Blenis, A new algorithm for reducing the workload of experts in performing systematic reviews, *Journal of the American Medical Informatics Association* **17** (2010), 446-453.
[6] X. Ji and P.-Y. Yen, Using MEDLINE Elemental Similarity to Assist Article Screening for Systematic Reviews, *JMIR* (2015).
[7] X. Ji, R. Machiraju, A. Ritter, and P.-Y. Yen, Examining the Distribution, Modularity, and Community Structure in Article Networks for Systematic Reviews, in: *AMIA*, 2015.
[8] R. Lengler and M.J. Eppler, Towards a periodic table of visualization methods for management, in: *IASTED Proceedings of the Conference on Graphics and Visualization in Engineering (GVE 2007), Clearwater, Florida, USA*, 2007.
[9] A. Kerren, J. Stasko, J.-D. Fekete, and C. North, *Information Visualization: Human-Centered Issues and Perspectives*, Springer, 2008.
[10] C. Ware, *Information visualization: perception for design*, Elsevier, 2012.
[11] H. Ding, C. Wang, K. Huang, and R. Machiraju, GRAPHIE: graph based histology image explorer, *BMC Bioinformatics* **16** (2015), 1.
[12] P. Hu and W.C. Lau, A Survey and Taxonomy of Graph Sampling, *arXiv preprint arXiv:1308.5865* (2013).
[13] E. John and I. Safro, Single- and Multi-level Network Sparsification by Algebraic Distance, *arXiv:1601.05527 [cs.SI]* (2016).
[14] G. Lindner, C.L. Staudt, M. Hamann, H. Meyerhenke, and D. Wagner, Structure-Preserving Sparsification of Social Networks, *arXiv:1505.00564v1 [cs.SI]* (2015).
[15] J. Chen and I. Safro, Algebraic distance on graphs, *SIAM Journal on Scientific Computing* **33** (2011), 3468-3490.
[16] L. Van der Maaten and G. Hinton, Visualizing data using t-SNE, *Journal of Machine Learning Research* **9** (2008), 85.
[17] Drug Effectiveness Review Project (DERP) Systematic Drug Class Review Gold Standard Data. skynet.ohsu.edu/~cohenaa/systematic-drug-class-review-data
[18] M. Jacomy, T. Venturini, S. Heymann, and M. Bastian, ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software, *PLoS One* **9** (2014), e98679.
[19] V.D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, Fast unfolding of communities in large networks, *Journal of Statistical Mechanics: Theory and Experiment* **2008** (2008), P10008.

Address for correspondence

Xiaonan Ji: ji.62@osu.edu
1800 Canon Drive, Columbus, OH 43210, USA