18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Temporal and Multi-Source Fusion for Detection of Innovation in Collaboration Networks

Benjamin A. Miller*, Michelle S. Beard†, Manfred D. Laubichler‡ and Nadya T. Bliss§
*MIT Lincoln Laboratory, Lexington, Massachusetts 02420
Email: bamiller@ll.mit.edu
†Charles Stark Draper Laboratory, Cambridge, Massachusetts 02139
Email: mbeard@draper.com
‡School of Life Sciences, Arizona State University, Tempe, Arizona 85287–4501
Santa Fe Institute, Santa Fe, New Mexico 87501
Marine Biological Laboratory, Woods Hole, MA 02543
Email: manfred.laubichler@asu.edu
§School of Computing, Informatics and Decision Systems Engineering
Simon A. Levin Mathematical, Computational and Modeling Sciences Center
Global Security Initiative
Arizona State University, Tempe, Arizona 85287–7205
Email: nadya.bliss@asu.edu

***Abstract*—A common problem in network analysis is detecting small subgraphs of interest within a large background graph. This includes multi-source fusion scenarios where data from several modalities must be integrated to form the network. This paper presents an application of novel techniques leveraging the signal processing for graphs algorithmic framework, to well-studied collaboration networks in the field of evolutionary biology. Our multi-disciplinary approach allows us to leverage case studies of transformative periods in this scientific field as truth. We build on previous work by optimizing the temporal integration filters with respect to truth data using a tensor decomposition method that maximizes the spectral norm of the integrated subgraph's adjacency matrix. We also demonstrate that we can mitigate data corruption via fusion of different data sources, demonstrating the power of this analysis framework for incomplete and corrupted data.**

I. Introduction

In numerous applications, the data of interest are entities and the relationships, connections, and interactions between them. We may be interested in interactions between individuals, communication between computers, or interaction between proteins. Across these diverse application areas, the data of interest are naturally represented as a graph.

One of the application domains where casting the data as a graph is widely used is the analysis of social networks. Analyzing the interactions between people allows for identification of community structure and influential figures. A network of scientific collaborators is a particularly interesting type of social network. Understanding the way innovation manifests itself within the public record via collaborative publications may lead to new insights into the evolution of scientific research.

In this paper, we analyze such a network in the context of a novel anomaly detection framework called signal processing for graphs (SPG) [1]. This framework enables the detection of small, anomalous clusters within large, dynamic background graphs. Within this framework, a filtering technique can be used to emphasize certain patterns of behavior and increase the power of these "signal" components of the graph within the "noise" of the background. This paper considers an optimization technique with respect to a known, rigorously-studied innovation period, and demonstrates that the optimal filter does in fact bring a significant subset of data to a prominent position within the analytical space. This framework can be applied when the graph is derived from many fused sources, which can also improve detection performance by considering multiple "looks" at the data. Since network data are often noisy or incomplete, we also consider observation of corrupted data within this context. While data corruption significantly hinders performance, we can leverage the diversity of multiple measurements and recover the signal by fusing the corrupted observations.

The remainder of this paper is organized as follows. Section II reviews the subgraph detection problem and defines notation. Section III discusses the filtering technique for analyzing dynamic graphs. Our dataset of interest—co-authorship networks of authors who all cite a seminal paper within a large, dynamic collaboration network—is described in Section IV. Section V presents the results of a set of experiments on this dataset, including filter optimization to best emphasize the innovation subnetwork and methods to fuse multiple corrupted observations and still maintain signal power. In Section VI, we summarize and provide possible avenues for further investigation.

II. Subgraph Detection Problem

A graph $G = (V, E)$ consists of a vertex set V, a set of entities, and an edge set E, a set of edges which represent relationships between vertices. The subgraph detection problem is a classical detection problem studied in [2] [3] with a graph as its observation. We cast the problem of subgraph detection as detecting a signal embedded in noise, where our objective is to resolve the binary hypothesis test

H_0 : The observed graph is "noise" G_B
H_1 : The observed graph is "signal+noise" $G_B \cup G_S$.

Let H_0 denote the null hypothesis, an undirected, unweighted graph $G_B = (V_B, E_B)$ generated by some random model. The alternative hypothesis, H_1, has an additional graph $G_S = (V_S, E_S)$ embedded into G_B. The problem is to decide whether or not the null hypothesis is true based on whether the observed graph deviates significantly from normal background behavior.

While optimal detection is possible in some scenarios [4], we focus on cases where this would be computationally intractable. We take a spectral approach, which has the benefit of analyzing the data in a space where there are known metrics for power and detectability [5]. Our subgraph detection procedure is based on the spectral analysis of modularity. Modularity is commonly used to detect communities in graphs [6], but in the context of this paper—and in the SPG framework more broadly—we analyze modularity to detect the presence of an anomaly. The modularity matrix B of an unweighted, undirected graph G is given by

$$B = A - \tfrac{kk^T}{2|E|},$$

where A is the adjacency matrix of G (i.e., A_{ij} is 1 if vertices v_i and v_j share an edge and is 0 otherwise) and k is the observed degree vector of G, where k_i is the number of edges connected to v_i. The matrix B can be considered the residuals matrix, a matrix consisting of the difference between the observed edges A and the expected edges $kk^T/(2\,|E|)$ (the expectation under the Chung–Lu model [7], which assumes no community structure).

The algorithms described in [2] [3] analyze the residuals under H_0 and H_1 by studying the eigendecomposition of B (i.e. $B = U\Lambda U^T$) and compute a test statistic to discriminate between the two hypotheses. In [2], we determine the presence of an anomaly by analyzing only the first two eigenvectors of B. To compute the test statistic, matrix B is projected into the space of its 2 principal eigenvectors u_1 and u_2. This is the linear subspace in which the residuals are largest, and, intuitively, a subgraph with particularly large residuals will separate from the rest of the vertices in this space.

III. Detection in Dynamic Networks

Extending the SPG framework to dynamic graphs, our observation is a sequence of time-varying graphs $G(n) = (V, E(n))$ where the vertex sets remain constant and the edge sets vary over time [8], [9]. Dealing with time-series graphs, we consider the residuals integrated over a time window. At each discrete time step n we have a graph $G(n)$ and a modularity matrix $B(n)$. We apply a finite impulse response filter h over the length of a time window ℓ and aggregate the residuals, obtaining

$$\tilde{B}(n) = \sum_{i=0}^{\ell-1} B(n-i)h(i).$$

Let $\tilde{B}(n)$ be the aggregated residuals matrix for the graph at time n filtered by h. Thus $\tilde{B}(n)$ is a matrix where in each vertex entry is the result of a vertex pair having its modularity filtered by h. The sequence of filter coefficients h is designed to effectively emphasize the subgraph and de-emphasize the background. The problem of choosing the appropriate filter coefficients is discussed in further detail later, where h will be computed to maximize the integrated signal power over time for a particular subgraph of interest. We perform the same analysis on $\tilde{B}(n)$ as performed on B for static graphs to discriminate between H_0 and H_1.

IV. Dataset

One of the significant challenges in developing and evaluating subgraph detection techniques is lack of truth for many of the applications of interest. In this work, as in [10], we leverage rigorously studied period of scientific innovation in evolutionary and developmental biology. This multidisciplinary approach allows us to refine our algorithmic techniques while also potentially providing insight into emergence of innovation in scientific literature. In this section, we describe the dataset that is used for our analysis throughout the rest of the paper.

The case study we consider is the emergence of the concept of gene regulatory networks in developmental biology. As discussed in [11], [12], [13], gene regulatory networks are one of the main explanatory concepts in today's evolutionary and developmental biology. The history and emergence of this idea are also well-studied [13]. This includes early conceptual ideas from the beginning of the 20th century and more specifically, the recent developments based on the formulation by Roy Britten and Eric Davidson published in *Science* in 1969 [14]. The Britten-Davidson (BD) model is a clear study of transformative innovation in a scientific field. As discussed in [10], the citations to the 1969 BD paper illustrate its persistent impact. Specifically, the citations rapidly increased throughout the 1970s, dropping somewhat during 1980s and 1990s, and again increasing in the 2000s and 2010s. Furthermore, second order citations, or citations of papers that cite the BD paper, tell a similar story - including a sharp increase in second order citations in 2000s and 2010s. Second order citations are good indicators of broader impact of the idea, especially when combined with first order citation patterns.

Study of the BD model and its impact on the field has allowed for an observation of the fact that scientific innovation at least in this case, leads to re-wiring of patterns of collaboration. Based on the analysis of this case study, truth or signal subgraphs were created by co-authorship graphs of the authors that have directly cited the BD paper. The signal subgraphs span 1969 to 2000. All citation graphs were extracted from the Web of Science database and covered a representative sample of the field of developmental biology, specifically, the top 12 journals in the field plus *Science*, *Nature*, and *Proceeding of the National Academy of Sciences*. For each year, the graphs considered were unweighted and undirected (co-authorship is fundamentally undirected) yielding a symmetric adjacency matrices. Total number of unique authors (number of vertices in the graph) was help consistently at 294,700 (representing the number of unique authors in the entire time period). The ordering of authors in the graph, while arbitrary, was preserved (as is necessary) in each year.

V. Experiments

A. Temporal Filter Optimization

Within the SPG framework, the spectral norm is a good power metric for signal and noise power [15]. When an embedded subgraph's spectral norm is large, its vertices are more likely to stand out in the eigenspace. When working with the temporal integration technique described in Section III, this means that it is desirable to choose filter coefficients that maximize the spectral norm of the principal submatrix of the adjacency matrix associated with the subgraph vertices.

As originally discussed in [16], the subgraph's spectral norm can be maximized by forming a 3-way tensor from the subgraph adjacency matrix, and computing a low-rank approximation for this tensor. Let $\mathbf{A_S}$ be an $N_S \times N_S \times \ell$ tensor for the subgraph vertices, where $N_S = |V_S|$. The first two dimensions represent vertices and the third dimension represents time. Much like approximating a matrix with its singular value decomposition, a low-rank tensor decomposition can be used to approximate $\mathbf{A_S}$. For a rank-1 approximation, this is achieved by solving

$$\arg\max_{\lambda,x,y,z} \sum_{i=1}^{N_S}\sum_{j=1}^{N_S}\sum_{t=1}^{\ell} \left(\mathbf{A_S}(i,j,t) - \lambda x_i y_j z_t\right)^2 \qquad (1)$$

subject to $\|x\|_2 = 1, \|y\|_2 = 1, \|z\|_2 = 1.$

Here $x, y \in \mathbb{R}^{N_S}$ and $z \in \mathbb{R}^{\ell}$ are vectors, and $\lambda \in \mathbb{R}$ is a scalar. Our objective is to maximize the spectral norm of the integrated adjacency matrix whose ijth entry is given by

$$a_{ij}^h = \sum_{t=1}^{\ell} \mathbf{A_S}(i,j,\ell+1-t)h(t).$$

It turns out that this quantity is optimized—under the constraint that the squares of the filter weights sum to 1—by setting the filter weights $h(t)$ equal to the time-reversed temporal factor $z_{\ell+1-t}$ from (1). This computation can be done in Matlab using the PARAFAC decomposition [17] in the Tensor Toolbox [18].

The effect of tuning the filter with respect to the vertices of interest has been demonstrated in simulation [16], but here we demonstrate application to the well-studied period of scientific innovation described in Section IV: we optimize the filter applied to the coauthorship graph from 1969 to 1980. As demonstrated in Fig. 1, the impact is extremely significant. Within each plot, there is one curve for each vertex in the subgraph of interest. In each case, the eigenvectors associated with the largest 20 (nonnegative) eigenvalues were computed, with smaller indices corresponding to larger eigenvalues. The values of the plots are the components of the (unit-normalized) eigenvectors that are associated with the subgraph vertices. Without any knowledge of truth, one may assume that simply averaging over time would be a reasonable approach, or that integrating using a ramp filter (where the weight on each successive time step increases in a linear fashion) would detect interesting subgraphs, given that this would emphasize emerging behavior. Using these strategies, as shown in the figure, there is only one vertex that is particularly strong within the eigenvectors with the largest eigenvalues. Using a method that considers the spectral norm of the subgraph at each point in time (i.e., using weights corresponding to the instantaneous power of the foreground) provides some additional benefit, as a few additional vertices stand out more prominently in eigenvector 14. Using a filter that is optimized via the tensor decomposition, on the other hand, brings out several more vertices. When this filter is applied, nine vertices from the subgraph stand out significantly in eigenvector four. Looking back at the data used to optimize the graph (i.e., the authors citing the seminal BD paper), these nine vertices comprise the largest connected component in any given year, and in fact form a clique (a graph with all possible edges) in 1977. Two of the authors in this cluster are also part of a larger clique with nine other authors in the background, who also stand out in the same eigenvector. This is a significant finding: the most interconnected that authors citing the BD paper ever become in a given year, as well as other close collaborators. Without this temporal integration technique, the subgraph would not stand out from the background within this low-dimensional space. We will focus on this subgraph for the remainder of the experiments.

B. Data Corruption

In many applications, the datasets from which the graphs are derived are inherently incomplete or noisy. When forming the graph, these errors can have a significant impact on detection performance. Interference or noise can lead to incorrect or missing edges. Clerical errors can lead to an edge being switched from one vertex to another. And data are frequently sampled from a population, giving an inherently incomplete view of the individual interactions. All of these factors can significantly hinder performance of detection algorithms.

The impact of network uncertainty on subgraph detection has been of interest in recent years [19], [20]. In this paper, we consider the impact of two of the corruption mechanisms from [20] on detection of the subgraph pulled out of the noise in Section V-A. One mechanism is a simple missing data model, in which each edge that exists in the true graph exists in the observed graph with equal probability. As noted in [20], this mechanism reduces the power of random background behavior more slowly than it reduces the power of clusters that stand out in the eigenspace. This is a consequence of Wigner's semicircle law. Considering an Erdős–Rényi random graph—

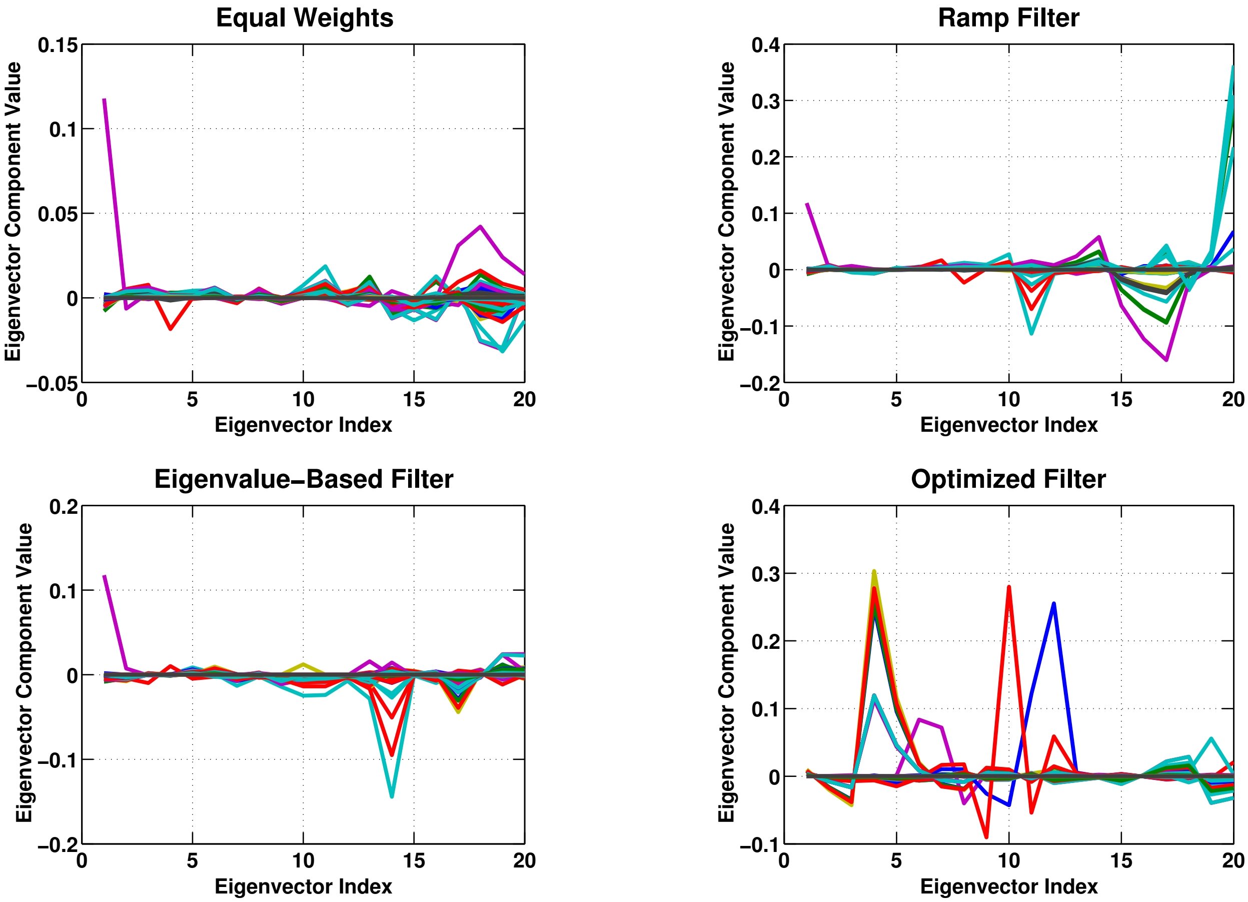

Fig. 1. Projections of subgraph vertices onto principal eigenvectors with various temporal integration techniques. Within the space of the principal eigenvectors, only one vertex is particularly prominent when using equal weights (top left), linearly increasing weights (top right), or weights determines by eigenvalues (bottom left). Only when an optimized filter is applied (bottom right) do a substantial number of subgraph vertices become prominent in the eigenspace.

where all possible edges are equally probable—the range of eigenvalues is proportional to the standard deviation of the edge presence probability, $\sqrt{p(1-p)}$. For sparse graphs, p will be small and thus changing p will change the largest eigenvalues by approximately $\sqrt{p}$. Meanwhile, the subgraph that does not fit the background model will have its spectral norm reduced by a factor of p, reducing the signal-to-noise ratio and making the detection problem more difficult. As shown in [20], this phenomenon also occurs in more realistic models that incorporate arbitrary degree structure.

The other corruption mechanism we consider is an edge-flipping mechanism, where there is a random model for data corruption based on vertex degree. For each vertex in the graph, we assume that the number of errors is proportional to the number of edges the vertex has. Similarly to the Chung–Lu model, we assign a weight w_i to vertex v_i, where

$$w_i = \frac{\alpha}{\sqrt{\sum_{j=1}^{N} k_j}} k_i.$$

The probability of an edge error between vertices v_i and v_j is then $p_{ij}^{\mathrm{corr}} = w_i w_j$. If there is an edge in the true graph between these two vertices, then it will not be observed with probability p_{ij}^{corr}, and if there is no such edge, then this is the probability with which an edge will be incorrectly observed. The scalar α controls the overall number of errors. In this paper, the corruption is focused on those vertices—in the entire graph, including the subgraph of interest—that are most prominent in the principal eigenspace for the true graph. This concentrates the effects of the corruption on the portion of the graph that we analyze, to better demonstrate the impact of this mechanism on eigenvector analysis.

A typical example of the impact of these data corruption mechanisms on subgraph detection ability is illustrated in Fig. 2. Each scatter plot in the figure is created using the two eigenvectors (among the principal 10) that most prominently feature the subgraph vertices. The nine-vertex clique from the data of interest, and the nine other close collaborators, clearly stand out in the fourth eigenvector when the graph is uncorrupted. For missing data, we consider a case where only 15% of the edges are observed. In the plotted instance, the subgraph vertices were most prominent in the eighth and ninth eigenvectors. While some of the vertices still stand out

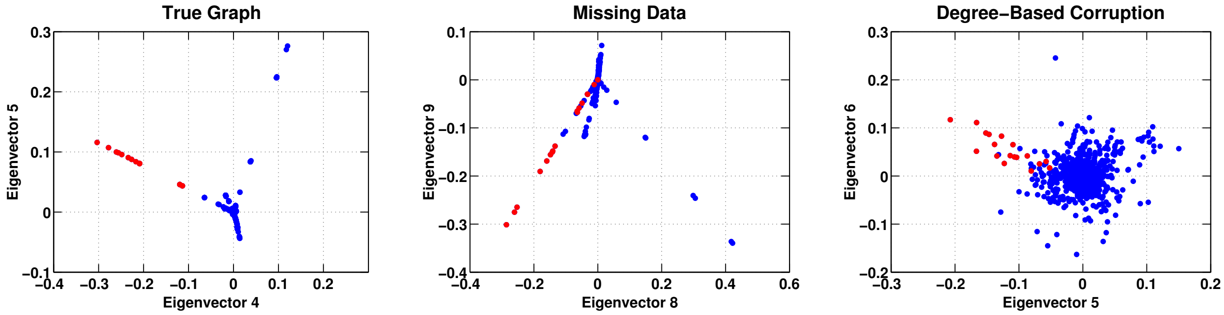

Fig. 2. Scatter plots emphasizing the subgraph from the known period of innovation. Background vertices are in blue, while the nine-vertex clique and its close collaborators are in red. When working with the true graph (left), the vertices all stand out in the fourth eigenvector. When only observing 15% of the edges, the subgraph partially stands out in the eighth and ninth eigenvectors, but many vertices are buried in the background (center). The degree-based corruption method similarly has a few vertices standing out in the 5th and 6th eigenvectors, but many of them are overpowered by background noise.

in this space, most of them are subsumed by other activity, and many vertices are very close to the origin. The degree-based corruption model (where about half of the observed vertices are errors) has a different effect on performance, but the result is similar. In the case plotted in the figure, the subgraph vertices stand out the most in the 5th and 6th eigenvectors. The background is much noisier due to the extra activity, and many of the subgraph vertices are buried within the noise. In both of these cases, the loss in power will reduce detection performance.

C. *Fusion of Corrupted Data*

While the medium through which we observe a network can create artifacts that hinder detection performance, it will sometimes be possible to get multiple "looks" at the data. If the error mechanisms are not correlated, it is possible to use the diversity of the measurement domains to recover performance. As alluded to in [20], this can be done via a Bayesian fusion method or by weighting the individual observations based on the level of trust in the source.

At relatively small scale, a Bayesian fusion method can be quite powerful for performance recovery. With the two corruption mechanisms considered in this paper, we can estimate that an edge exists in the latent graph in the following way. Let p^{prior} be the prior probability of edge existence in the latent graph, and let a_{ij}, a_{ij}^{miss}, and a_{ij}^{corr} be the ijth entry in the adjacency matrix of the true graph, the graph with missing data, and the graph with degree-based corruption, respectively. If a_{ij}^{miss} is 1, then a_{ij} is 1, since edges can only be taken away with the missing data mechanism. If a_{ij}^{miss} is zero, then the probability that the edge exists in the true graph is

$$P\left[a_{ij}=1\right]=\begin{cases}\frac{\left(1-p^{\text{obs}}\right)p_{ij}^{\text{corr}}p^{\text{prior}}}{\left(1-p^{\text{obs}}\right)p_{ij}^{\text{corr}}+\left(1-p_{ij}^{\text{corr}}\right)} & \text{if } a^{\text{corr}}=0\\ \frac{\left(1-p^{\text{obs}}\right)\left(1-p_{ij}^{\text{corr}}\right)p^{\text{prior}}}{\left(1-p^{\text{obs}}\right)\left(1-p_{ij}^{\text{corr}}\right)+p_{ij}^{\text{corr}}} & \text{if } a^{\text{corr}}=1.\end{cases} \quad (2)$$

While fusing in this fashion has the potential to completely recover detection performance—as demonstrated in simulation in [20]—the posterior expected value of A will be dense, and may not have the sort of exploitable structure (e.g., low-rank structure) that enables efficient eigenvector analysis at scale. In practice, it may also be difficult to estimate the model parameters, and there may be mismatch with the true model. We therefore focus on a method for fusing based on a weighted sum.

When given the two observed graphs, they will be fused as follows. For each pair of vertices, a fused observation will be computed as

$$\hat{a}_{ij}=\frac{1}{1+\exp\left(-\beta_0-\beta_1 a_{ij}^{\text{miss}}-\beta_2 a_{ij}^{\text{corr}}\right)}. \quad (3)$$

Here the β parameters are the weights of the corrupted observations. We are operating in the context of logistic regression, where a linear function of the observations is mapped to an expected value via the logistic function. Within this context, values for $\hat{a}_{ij}$ only need to be computed if an edge exists between v_i and v_j in one of the observations. Otherwise, the probability is assumed to be $1/(1+e^{-\beta_0})$, which can be accounted for by adding a rank-1 matrix to the fused observations.

Fusing the observations in this way improves the representation of our subgraph of interest in the eigenspace, as demonstrated in Fig. 3. Under the same corruption scenarios as in Section V-B, we measured the "power" of the subgraph in the first 10 eigenvectors. Let U be the $N\times 10$ matrix where each column is an eigenvector of the integrated modularity matrix for the observed (or fused) graph, and let $x\in\{0,1\}^N$ be an indicator vector for the subgraph that is emphasized by the optimized filter. We measure the power of the subgraph in this space as $\|U^Tx\|_2^2$, i.e., the L_2 norm squared of the orthogonal projection into the space spanned by the 10 principal eigenvectors. Using the optimized filter, this will be reduced by the corruption mechanisms, but can be recovered by fusing the two observations. Let P_{true} be the power when U is computed from the true graph, and we will compare the power P from other cases to this quantity. Fig. 3 provides cumulative density functions (CDF) demonstrating the probability that a corrupted (or fused) observation will provide the signal power within its principal eigenspace, as determined via a Monte Carlo simulation. As shown in the figure, working with only 15% of

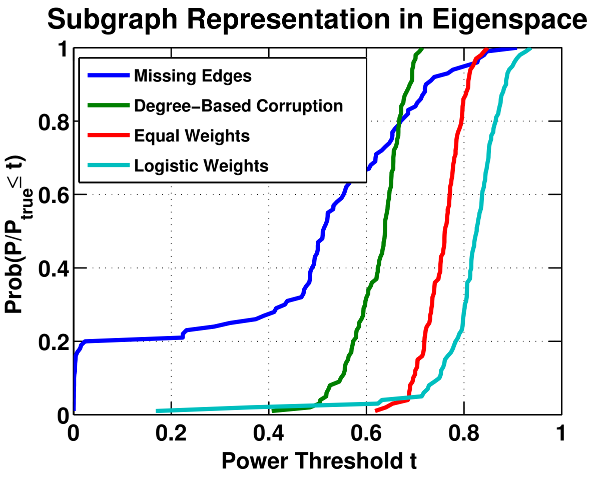

Fig. 3. Cumulative density functions for the signal power maintained in various scenarios. The power level considered is the norm squared of the projection of the indicator vector for the subgraph. The horizontal axis is the ratio of this power level for observed or fused data to the same quantity with the true (uncorrupted) graph.

the edges can significantly reduce the power of the subgraph in the top eigenvectors: about half the time less than 50% of the power remains. The degree-based corruption mechanism in which about half of the observations are incorrect also reduces performance, but not usually as drastically, maintaining, on average, over 62% of the power. By simply averaging the two observations together, we shift the CDF by over 10%. Finally, by using the fusion technique of (3), we improve upon this result, increasing the signal power maintained by an additional 5%.

VI. Conclusion

This paper investigates the use of temporal and multi-source integration to enable detection of known innovation patterns in scientific literature. Dynamic collaboration networks are analyzed with the signal processing for graphs framework, focused principally on eigenspace analysis of graph residuals integrated over time. The temporal weights are optimized with respect to a known innovation period surrounding the Britten-Davidson model for gene regulation, specifically among authors that cite the seminal paper on the model. We demonstrate that this technique boosts the power of the largest connected component of this subset of the data to a point where it can be detected within a low-dimensional projection of the data. Using two simple error models for graph data, we show the negative impact of working with a corrupted graph, with the detected subgraph having its power reduced while being subsumed by other activity in the principal eigenspace. Using a simple weighting procedure, we demonstrate that we can recover the power of the subgraph within this space.

There are numerous potential areas for future development. One interesting area would be determining an approximation to the Bayesian fusion method in (2) that would allow the technique to scale to very large graphs. Another possibility would be to study optimization of filter coefficients when there are missing data in the training set, as in [21]. At a higher level, it would be interesting to determine what other subgraphs can be emphasized by this technique, and to find what subgraphs are detected using the same filters in more recent publication data. A comparative study of which filters detect patterns of innovation in different scientific fields might also contribute to a better understanding of the structure of different scientific practices.

Acknowledgment

The Lincoln Laboratory portion of this work is sponsored by the Assistant Secretary of Defense for Research & Engineering under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the United States Government. M. D. L. is supported by the NSF SES Grant No. 1243575.

References

[1] B. A. Miller, N. T. Bliss, P. J. Wolfe, and M. S. Beard, “Detection theory for graphs,” *Lincoln Laboratory J.*, vol. 20, no. 1, 2013.

[2] B. A. Miller, N. T. Bliss, and P. J. Wolfe, “Toward signal processing theory for graphs and non-Euclidean data,” in *ICASSP*, 2010, pp. 5414–5417.

[3] ——, “Subgraph detection using eigenvector L1 norms,” in *Advances in Neural Inform. Process. Syst. 23*, J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, Eds., 2010, pp. 1633–1641.

[4] T. Mifflin, “Detection theory on random graphs,” in *Proc. Int. Conf. Inform. Fusion*, 2009, pp. 954–959.

[5] R. R. Nadakuditi, “On hard limits of eigen-analysis based planted clique detection,” in *Proc. IEEE Statistical Signal Process. Workshop*, 2012, pp. 129–132.

[6] M. E. J. Newman, “Finding community structure in networks using the eigenvectors of matrices,” *Phys. Rev. E*, vol. 74, no. 3, 2006.

[7] F. Chung, L. Lu, and V. Vu, “The spectra of random graphs with given expected degrees,” *Proc. of National Academy of Sciences of the USA*, vol. 100, no. 11, pp. 6313–6318, 2003.

[8] B. A. Miller, M. S. Beard, and N. T. Bliss, “Matched filtering for subgraph detection in dynamic networks,” in *Proc. IEEE Statistical Signal Process. Workshop*, 2011, pp. 509–512.

[9] ——, “Eigenspace analysis for threat detection in social networks,” in *Int. Conf. Inform. Fusion*, 2011, pp. 1–7.

[10] N. T. Bliss, B. R. E. Peirson, D. Painter, and M. D. Laubichler, “Anomalous subgraph detection in publication networks: Leveraging truth,” in *Proc. 48th Asilomar Conf. Signals, Syst. and Comput.*, 2014.

[11] E. H. Davidson, *The regulatory genome: Gene regulatory networks in development and evolution*. Academic Press, 2010.

[12] D. C. Krakauer, J. P. Collins, D. Erwin, J. C. Flack, W. Fontana, M. D. Laubichler, S. J. Prohaska, G. B. West, and P. F. Stadler, “The challenges and scope of theoretical biology,” *Journal of Theoretical Biology*, vol. 276, no. 1, pp. 269–276, 2011.

[13] M. D. Laubichler, J. Maienschein, and J. Renn, “Computational perspectives in the history of science: To the memory of Peter Damerow,” *Isis*, vol. 104, no. 1, pp. 119–130, 2013.

[14] R. J. Britten and E. H. Davidson, “Gene regulation for higher cells: A theory,” *Science*, vol. 165, no. 891, pp. 349–357, 1969.

[15] B. A. Miller, M. S. Beard, P. J. Wolfe, and N. T. Bliss, “A spectral framework for anomalous subgraph detection,” 2014, preprint: arXiv:1401.7702.

[16] B. A. Miller and N. T. Bliss, “Toward matched filter optimization for subgraph detection in dynamic networks,” in *Proc. IEEE Statistical Signal Process. Workshop*, 2012, pp. 113–116.

[17] E. Acar, D. M. Dunlavy, and T. G. Kolda, “A scalable optimization approach for fitting canonical tensor decompositions,” *Journal of Chemometrics*, vol. 25, no. 2, pp. 67–86, February 2011.

[18] B. W. Bader, T. G. Kolda *et al.*, "Matlab tensor toolbox version 2.5," Available online, January 2012. [Online]. Available: http://www.sandia.gov/∼tgkolda/TensorToolbox/

[19] J. B. Collins and S. T. Smith, "Network discovery for uncertain graphs," in *Proc. Int. Conf. Inform. Fusion*, 2014.

[20] B. A. Miller and N. Arcolano, "Spectral subgraph detection with corrupt observations," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process.*, 2014, pp. 3449–3453.

[21] E. Acar, D. M. Dunlavy, T. G. Kolda, and M. Mørup, "Scalable tensor factorizations for incomplete data," *Chemometrics and Intelligent Laboratory Systems*, vol. 106, no. 1, pp. 41–56, March 2011.