

Design & Manufacturing

In this issue:
• A new design optimization method for multi-objective optimization
• A new design optimization method for multi-objective optimization
• A new design optimization method for multi-objective optimization
• A new design optimization method for multi-objective optimization

Data Science, Quality & Reliability

In this issue:
• A new data science method for multi-objective optimization
• A new data science method for multi-objective optimization
• A new data science method for multi-objective optimization
• A new data science method for multi-objective optimization



Monitoring sparse and attributed networks with online Hurdle models

Samaneh Ebrahimi, Mostafa Reisi-Gahrooei, Kamran Paynabar & Shawn Mankad

To cite this article: Samaneh Ebrahimi, Mostafa Reisi-Gahrooei, Kamran Paynabar & Shawn Mankad (2022) Monitoring sparse and attributed networks with online Hurdle models, IISE Transactions, 54:1, 91-104, DOI: [10.1080/24725854.2020.1861390](https://doi.org/10.1080/24725854.2020.1861390)

To link to this article: <https://doi.org/10.1080/24725854.2020.1861390>



Published online: 28 Jan 2021.



Submit your article to this journal [↗](#)



Article views: 223



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 5 View citing articles [↗](#)



Monitoring sparse and attributed networks with online Hurdle models

Samaneh Ebrahimi^a, Mostafa Reisi-Gahrooei^b, Kamran Paynabar^a, and Shawn Mankad^c

^aH. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, GA, USA; ^bDepartment of Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA; ^cDepartment of Operations, Technology, and Information Management, Cornell University, Ithaca, NY, USA

ABSTRACT

In this article we create a novel monitoring system to detect changes within a sequence of networks. Specifically, we consider sparse, weighted, directed, and attributed networks. Our approach uses the Hurdle model to capture sparsity and explain the weights of the edges as a function of the node and edge attributes. Here, the weight of an edge represents the number of interactions between two nodes. We then integrate the Hurdle model with a state-space model to capture temporal dynamics of the edge formation process. Estimation is performed using an extended Kalman Filter. Statistical process control charts are used to monitor the network sequence in real time in order to identify changes in connectivity patterns that are caused by regime shifts. We show that the proposed methodology outperforms alternative approaches on both synthetic and real data. We also perform a detailed case study on the 2007–2009 financial crisis. Demonstrating the promise of the proposed approach as an early warning system, we show that our method applied to financial interbank lending networks would have raised alarms to the public prior to key events and announcements by the European Central Bank.

ARTICLE HISTORY

Received 25 February 2020
Accepted 31 August 2020

KEYWORDS

Online Hurdle models;
networks temporal
monitoring; interbank
network; kalman filter

1. Introduction

Due to advances in information technology and data collection systems, it is becoming increasingly common to collect node or entity-level data over time. When such data includes interactions between nodes, a natural representation that is often utilized is networks (graphs). For instance, social networks composed of individual accounts and their connections are ubiquitous (Goel and Goldstein, 2013), as are supply chain networks (Osadchiy *et al.*, 2016), communication networks (Trier, 2008), financial networks (Brunetti *et al.*, 2019), among others (Fienberg, 2012). Recent scholarly attention has been given to modeling and analysis of a sequence of networks (see Reisi Gahrooei and Paynabar (2018) and references therein). In this setting, detecting changes in the underlying network stream can be extremely important, as it facilitates better management of the overall system being modeled (Woodall *et al.*, 2017). Policy interventions, such as preventive maintenance in mechanical systems or liquidity injections in financial systems, can be taken in an informed and timely manner to mitigate potentially negative outcomes.

Attributed (labeled) binary networks have been widely used for modeling cases in which the connectivity of two nodes is a function of node attributes. For example, in a social network the likelihood of friendship between two people is a function of their age, sex, education, etc. Azarnoush *et al.* (2016) integrated logistic regression with Statistical Process Control (SPC) methods to detect changes in

attributed network streams. One major drawback of their method is that it does not properly model the dynamic evolution of network streams. In order to address this issue, Reisi Gahrooei and Paynabar (2018) proposed a method that integrates Generalized Linear Models (GLMs) with State-Space Models (SSMs). GLMs can help to model the attributed network structure and SSMs are used to capture the temporal dynamic of the network stream. In addition, Xu and Hero (2014) and Zou and Li (2017) employed state-space modeling framework to capture the network dynamics.

In practice, the nodes in attributed networks are often sparsely connected. For example, in financial networks, interactions are rarely seen between banks (nodes), especially in crisis environments, which leads to a sparse attributed network. As an example, consider the e-MID trading platform, the only electronic regulated interbank market in the world, from January 2006 to December 2012. Edges are defined by the number of overnight loans between European banks on this platform, that is, if Bank A lends to Bank B, then an edge is drawn from Bank A to Bank B and weighted by the number of directed loans in the given week. There is around 74% sparsity among all possible bank interactions. This value increases to 84% during a crisis. As GLMs are not capable of modeling zero-inflated distributions, the modeling method by Reisi Gahrooei and Paynabar (2018) would fall short of the required outcome.

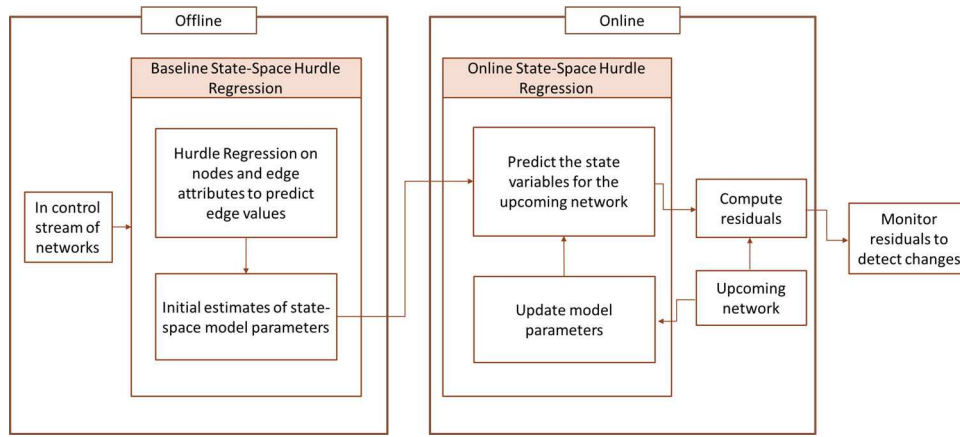


Figure 1. Overview of the proposed network monitoring methodology.

In this article we are primarily concerned with proposing a new modeling and change detection methodology for sparse, weighted, attributed network streams. Here, the weight of an edge represents the number of interactions between two nodes. A novel aspect of our setting is that of sparsely connected networks. Sparsity, a lack of edge connections, is present in a variety of contexts where highly granular data is available, including our case study application of financial networks. Beyond sparsity, edges in networks can in general be weighted and additional covariates (attributes) may be available over time about the nodes or edges. It is in this context that we propose a new modeling, monitoring and change detection methodology.

The main idea behind our approach is to use the Hurdle model (Mullahy, 1986) to understand the edge formation process as a function of the node and edge attributes. The parameters of the Hurdle model evolve based on a state-space model to capture the temporal dynamics in the data. An Extended Kalman Filter is used as an online, recursive inference procedure to estimate and update the parameters over time. Finally, we generate a one-step-ahead prediction of the network and compare it to the realized network to decide whether the observed evolution was smooth or abrupt using an Exponentially Weighted Moving Average (EWMA) control chart. A novel methodological contribution is the combination of these methods: the Hurdle model plus state-space evolution, Extended Kalman Filter estimation and statistical process control charts to monitor sequences of sparse, attributed, and weighted networks. The method is discussed in detail in Section 2.

We validate the proposed approach in two ways. First, in Section 3, we use simulation to establish self-consistency and functional performance of the proposed modeling and estimation framework. We show that the proposed method outperforms several alternative techniques, and that sparsity of the networks is a key data characteristic driving our model's favorable results. This important insight provides guidance into when the proposed model should be strongly considered over alternative methods. Our second validation is done through a detailed case study in Section 4, where we demonstrate how the methodology can be applied for monitoring interbank lending networks during the 2007–2009

financial crises. The application addresses the important societal problem of assessing financial stability. In fact, we find several promising and novel results showing that the proposed model would have raised alarms to the public before official announcements by the European Central Bank. The identified change point dates are also highly interpretable, matching closely with several key real-world events. Overall, our consistently favorable and positive results demonstrate that the proposed approach could be a valuable tool for regulators and financial institutions to utilize when monitoring the banking system as well as in other settings where dynamic networks are encountered.

2. The proposed methodology

2.1. Overview

We propose a new monitoring methodology for sparse attributed network streams with dynamic structures. The methodology is comprised of modeling the network structure and providing a change detection methodology. In our modeling framework, it is assumed that the edge probabilities are functions of the node and edge attributes. For example, in the context of financial networks, the probability of a transaction (or the number of transactions) between two banks could be a function of their country of origins, prevailing interest rates, and so on. Although GLMs have been successfully used to model attributed networks (Azarnoush *et al.*, 2016; Reisi Gahrooei and Paynabar, 2018), network sparsity (extreme lack of node connections) violates the assumption to GLMs that the underlying probability distribution should belong to the exponential family of distributions. To address this issue, we use the Hurdle model, which is capable of handling zero-inflated distributions (Mullahy, 1986).

The Hurdle model has been used previously to account for network sparsity (Heard *et al.*, 2010); however, previous approaches did not utilize edge and node attributes, instead modeling edge probabilities as a function of time. To take the network dynamics coupled with edges and nodes' attributes into account, we integrate the state-space model with the Hurdle model, where it is assumed that the parameters of the Hurdle regression follow a Markovian process, and

develop a sequential estimation scheme using an Extended Kalman Filter (EKF) to update the state space parameters and predict the value of upcoming networks. The overall framework is illustrated in Figure 1. As shown in the figure, in the offline phase, using a stream of in-control (i.e., training) networks, we build a Hurdle model using nodes and edges attributes, and estimate the initial state-space parameters. In the online (i.e., deployment) phase, as new network observations arrive, the estimated Hurdle model is used to predict the edge values for the incoming network snapshot. Additionally, with the upcoming network observations, the parameters of the state-space Hurdle model are updated using EKF. As time progresses, residuals that compare the newly realized network with the predicted one are used to detect a sudden, structural change in the network through EWMA control charts.

2.2. Hurdle models

As mentioned earlier, in attributed networks, the edge probabilities can be defined as functions of nodes and/or edge attributes and often modeled by using GLMs. However, in practice, networks are often sparse (Faloutsos *et al.*, 2004; Akoglu *et al.*, 2015), where typically each node interacts with only a few other nodes with which it shares some common attributes and characteristics. The sparsity of a network in a sequence of dynamic networks depends on the resolution of the temporal sampling (i.e., the window size). That is the shorter the sampling interval is, the sparser the corresponding network would be. For example, banking networks constructed on the basis of daily interactions would be sparser than the networks constructed based upon weekly data. As a result the number of zeros within the network could be inflated or deflated depend on the sampling resolution. Regular Poisson models are ineffective in modeling these situations, as they cannot capture zero inflation or deflation. A zero-inflated Poisson model is also not appropriate, as it can only capture the inflation in the number of zeros. In addition, We assume that all nodes can have interactions with one another. This assumption aligns with our case study, in which all European banks are in-theory allowed to interact with each other without any restrictions. Therefore, we assume only one type of zero (related to sampling) exists in our model. Therefore, an alternative modeling approach that treats zeros separately from other interaction counts in the network is required. In this article, we consider a Hurdle model and extend it to dynamic settings.

The fundamental idea behind the Hurdle model is to generate count variables (edge weights) in a two-stage process. In the first stage, a binary process specifies whether the edge weight is zero or positive. Given the first stage indicating a non-zero edge weight (i.e., if the hurdle is crossed), another stochastic process generates a positive edge weight. In the first stage of the Hurdle model, we assume that the weight of an edge between nodes i and j at time t is equal to zero with probability of $\pi_{i,j,t}$ and is a positive value with probability $1 - \pi_{i,j,t}$. Our proposed framework can in fact handle both dense and sparse networks. For the dense

networks the proposed framework will estimate a small value for $\pi_{i,j}$ and for sparse networks it estimates larger $\pi_{i,j}$. Given that the edge weight is positive, we assume that its value k follows a probability distribution function $f_1^+(k; \Lambda_{i,j,t})$ with parameters $\Lambda_{i,j,t}$. The associated truncated probability is then calculated as $f_1^+(k; \Lambda_{i,j,t}) / (1 - f_1^+(0; \Lambda_{i,j,t}))$, which ensures that a zero count does not occur in this case. Note that the positive probability needs to be multiplied by $(1 - \pi_{i,j,t})$ to ensure that probabilities sum to one (Cameron and Trivedi, 2013). Different Hurdle models can be introduced based on the choice of f_1^+ . For our Hurdle model, we focus on the ‘‘Poisson-logit’’ specification, where f_1^+ is a Poisson distribution. The probabilities for our hurdle model are shown in Equation (1):

$$P(w_{i,j,t} = k) = \begin{cases} \pi_{i,j,t} & k = 0 \\ (1 - \pi_{i,j,t}) \frac{\exp(-\lambda_{i,j,t}) \lambda_{i,j,t}^k}{k!(1 - \exp(-\lambda_{i,j,t}))} & k \geq 1, \end{cases} \quad (1)$$

where $w_{i,j,t}$ is a random variable for the weight of the edge between node i and j at time t , $\pi_{i,j,t}$ is the probability of having no edge between nodes i and j , and

$$\frac{\exp(-\lambda_{i,j,t}) \lambda_{i,j,t}^k}{k!(1 - \exp(-\lambda_{i,j,t}))}$$

is the probability distribution function of a truncated Poisson process at zero (also known as Positive Poisson; Grogger and Carson (1991)). Equation (1) can clearly be decomposed as the mixture of a Bernoulli distribution with parameter $\pi_{i,j,t}$ and a Positive Poisson distribution with parameter $\lambda_{i,j,t}$.

To extend this model to a regression setting, we assume that model parameters ($\lambda_{i,j,t}$ and $\pi_{i,j,t}$) are functions of covariates (node and edge attributes). To write our model, we first define a Bernoulli random variable $d_{i,j,t}$ as the indicator of positive occurrence (if $d_{i,j,t} = 1$ then $w_{i,j,t} > 0$), and $w_{i,j,t}^+$ as a random variable for positive value of an edge weight. For simplicity, we denote $f_0(\pi_{i,j,t})$ and $f_1^+(\lambda_{i,j,t})$ as distributions of zero and positive counts, respectively. The Hurdle regression, using logit and exponential link functions (Nelder and Baker, 1972) can be written as:

$$d_{i,j,t} \sim f_0(1 - \pi_{i,j,t})$$

$$\pi_{i,j,t}(\beta_{0,t}; \mathbf{z}_{i,j,t}) = 1 - \text{logit}^{-1}(\mathbf{z}_{i,j,t} \beta_{0,t})$$

and

$$w_{i,j,t}^+ \sim f_1^+(\lambda_{i,j,t})$$

$$\lambda_{i,j,t}(\beta_{1,t}; \mathbf{x}_{i,j,t}) = \exp(\mathbf{x}_{i,j,t} \beta_{1,t}),$$

where $f_0 = \pi_{i,j,t}^{(1-d_{i,j,t})} (1 - \pi_{i,j,t})^{(d_{i,j,t})}$ and $\mathbf{x}_{i,j,t}$ and $\mathbf{z}_{i,j,t}$ are the covariates (attributes corresponding to the starting node i , ending node j , and edge ij) used in Bernoulli and Positive Poisson count models, respectively. In general, $\mathbf{x}_{i,j,t}$ ($\mathbf{z}_{i,j,t}$) differs from $\mathbf{x}_{j,i,t}$ ($\mathbf{z}_{j,i,t}$), and so a directed relationship can be explained. Note that in this work, we have two types of variables: One that mainly explains the presence of an edge

denoted by $\mathbf{z}_{i,j,t}$, and one that explains the strength of the relationship represented by $\mathbf{x}_{i,j,t}$. For this reason, we separated their corresponding parameters denoted by $\beta_{0,t}$ and $\beta_{1,t}$. Note that it is assumed that given attributes the interactions between nodes are independent. The set of covariates used in each model can be the same or different.

To estimate the model parameters, $\beta_{0,t}$ and $\beta_{1,t}$, the following likelihood should be maximized

$$l(\beta_0, \beta_1) = f(w_t) = \prod_{i,j \in \mathcal{N}, i \neq j} \pi_{i,j,t}^{(1-d_{i,j,t})} (1 - \pi_{i,j,t})^{d_{i,j,t}} \times \left[\frac{\exp(-\lambda_{i,j,t}) \lambda_{i,j,t}^{w_{i,j,t}}}{w_{i,j,t}! (1 - \exp(-\lambda_{i,j,t}))} \right]^{d_{i,j,t}}$$

Therefore the log likelihood can be written as:

$$\begin{aligned} \mathcal{L}(\beta_0, \beta_1) = \sum_{\{i,j \in \mathcal{N}, j \neq i\}} \left\{ (d_{i,j,t}) \times [\log(1 - \pi_{i,j,t}(\beta_{0,t}, \mathbf{z}_{i,j,t}))] \right. \\ \left. + (1 - d_{i,j,t}) \times [\log(\pi_{i,j,t}(\beta_{0,t}, \mathbf{z}_{i,j,t}))] + (d_{i,j,t}) \right. \\ \left. \times \log \left(\frac{\exp(-\lambda_{i,j,t}(\beta_{1,t}, \mathbf{x}_{i,j,t})) \lambda_{i,j,t}(\beta_{1,t}, \mathbf{x}_{i,j,t})^{w_{i,j,t}}}{w_{i,j,t}! (1 - \exp(-\lambda_{i,j,t}(\beta_{1,t}, \mathbf{x}_{i,j,t})))} \right) \right\} \end{aligned} \quad (2)$$

where $\mathcal{N}$ is the set of total nodes in the network. Assuming independence among $f_0(\beta_0, \mathbf{z})$ and $f_1^+(\beta_1, \mathbf{x})$, the log-likelihood function in Equation (2) can be written as a sum of two separate components, namely, $\sum_{\{i,j \in \mathcal{N}, j \neq i\}} \{ (d_{i,j,t}) \times [\log(1 - \pi_{i,j,t}(\beta_{0,t}, \mathbf{z}_{i,j,t}))] + (1 - d_{i,j,t}) \times [\log(\pi_{i,j,t}(\beta_{0,t}, \mathbf{z}_{i,j,t}))] \}$ and $\sum_{\{i,j \in \mathcal{N}, j \neq i\}} \{ d_{i,j,t} \times \log(f_1^+(w_{i,j,t}^+, \lambda_{i,j,t}(\beta_{1,t}, \mathbf{x}_{i,j,t}))) \}$. Hence, each component can be maximized individually resulting in the following estimates:

$$\hat{\pi}_{i,j,t} = 1 - \text{logit}^{-1}(\mathbf{z}_{i,j,t} \hat{\beta}_{0,t}) \quad (3)$$

$$\hat{\lambda}_{i,j,t} = \exp(\mathbf{x}_{i,j,t} \hat{\beta}_{1,t}), \quad (4)$$

where $\hat{\beta}_{0,t}$ and $\hat{\beta}_{1,t}$ are the estimated regression coefficients at time t . Next we discuss how to incorporate network structural dynamics through a state-space model on the parameters of the Hurdle model, and use the EKF to estimate and update model parameters over time.

Note that, in this article the equations are presented for directed networks; however, this methodology can be applied for estimating the edge values of undirected networks, using symmetric mappings from attributes to edge values. This is because undirected networks are a special case of directed networks when edges take same value in both directions.

2.3. State-space models and the EKF

State-space models provide a flexible framework for modeling dynamic systems. In this approach, although the actual state of the system is unknown, it can be inferred over time using noisy observations. In the context of attributed network streams, we assume that the coefficients of the Bernoulli and truncated Poisson regression (β_0 and β_1) are the state variables, which are driven by a stochastic process.

The observed edge values, $w_{i,j,t}$ for $t = 1, 2, \dots$, are noisy observations. Therefore, the state-space Hurdle model is defined by the following equations:

$$\begin{aligned} \beta_t &= \mathbf{F}\beta_{t-1} + \epsilon_t \\ w_{i,j,t} &= h(\mathbf{x}_{i,j,t}, \mathbf{z}_{i,j,t}, \beta_t), \end{aligned} \quad (5)$$

where $\beta_t = [\beta_{0,t}, \beta_{1,t}]$ is the state vector and $\mathbf{F}$ is the state transition matrix, and $\epsilon_t \sim N(\mathbf{0}, \mathbf{Q})$ is the process noise with mean $\mathbf{0}$ and covariance matrix $\mathbf{Q}$. And h is the nonlinear link function generating a realization of $w_{i,j,t}$ given the state of the system β_t and the vector of covariates ($\mathbf{x}_{i,j,t}$ and $\mathbf{z}_{i,j,t}$). In the case of the Hurdle model, we showed that the likelihood function can be separated into two components and optimized by maximizing each component, separately. This will lead to two decoupled models given in Equations (3) and (4). Therefore, we can take $h(\mathbf{x}_{i,j,t}, \mathbf{z}_{i,j,t}, \beta_t) = 1 - \text{logit}^{-1}(\mathbf{z}_{i,j,t} \beta_{0,t})$ for modeling the zero counts (logistic regression) and $h(\mathbf{x}_{i,j,t}, \mathbf{z}_{i,j,t}, \beta_t) = \exp(\mathbf{x}_{i,j,t} \beta_{1,t})$ for modeling positive Poisson counts.

In the case of linear state-space models, the Kalman Filter (KF) procedure achieves the optimal estimate of the states (Kalman, 1960). However, as the observation model in Equation (5) is nonlinear, we employ the EKF, which is shown to be effective in incorporating nonlinearity in parameter estimation (Fahrmeir and Kaufmann, 1991; Brown and Hwang, 1997). EKF is mainly designed for normally distributed observations. However, Fahrmeir and Kaufmann (1991) showed that EKF can also be used for the exponential family of distributions. As our proposed Hurdle model is decomposed into Bernoulli and Poisson distributions that are both from the exponential family, we will employ EKF for parameter estimation. Similar to KF, EKF provides a recursive estimation procedure that only uses the current network snapshot (at time t) and the previous parameter estimates (at time $t - 1$) to update the parameter estimates. EKF uses the Taylor expansion to linearize the nonlinear observation function, $h(\mathbf{x}_{i,j,t}, \mathbf{z}_{i,j,t}, \beta_t)$, and then applies the KF estimation equations. Specifically, given $\mathbf{F}$ and $\mathbf{Q}$, the EKF for the state-space Hurdle regression can be summarized as follows. Note that the detailed derivations of the prediction and update equations for EKF can be found in Brown and Hwang (1997).

2.3.1. Prediction step

Let $\beta_{t|t-1}$ and $\mathbf{P}_{t|t-1}$ denote the Kalman predictions of the state β_t and its covariance matrix given observation until time $t - 1$ (w_l ; $l = 1, \dots, t - 1$), and let $\beta_{t|t}$ and $\mathbf{P}_{t|t}$ denote the estimation of the state and its covariance matrix, given observations until time t (w_l ; $l = 1, \dots, t$). Now using the previous estimates, the prediction equations at time t are given by

$$\begin{aligned} \beta_{t|t-1} &= \mathbf{F}\beta_{t-1|t-1} \\ \mathbf{P}_{t|t-1} &= \mathbf{F}\mathbf{P}_{t-1|t-1}\mathbf{F}^T + \mathbf{Q}, t = 1, 2, \dots \end{aligned}$$

where the initial estimates, $\beta_{0|0}$ and $\mathbf{P}_{0|0}$ can be obtained from fitting a Hurdle model to the first network snapshot data.

2.3.2. Update step

Let $\mathbf{w}_t = \text{vec}[w_{i,j,t}]$ to be the vectorized adjacency matrix containing the noisy trades between pairs of nodes. Also let $\mathbf{x}_t$ and $\mathbf{z}_t$ to be covariate matrices whose rows are $\mathbf{x}_{i,j,t}$ and $\mathbf{z}_{i,j,t}$, respectively. At time t , the incoming network data ($\mathbf{w}_t$) are used to update the predicted parameters using the set of equations:

$$\begin{aligned} \mathbf{K}_t &= \mathbf{P}_{t|t-1} \mathbf{H}_t^T (\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^T + \mathbf{R}_t)^{-1} \\ \boldsymbol{\beta}_{t|t} &= \boldsymbol{\beta}_{t|t-1} + \mathbf{K}_t (\mathbf{w}_t - h(\mathbf{x}_t, \mathbf{z}_t, \boldsymbol{\beta}_{t|t-1})) \\ \mathbf{P}_{t|t} &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1}, \end{aligned}$$

where

$$\mathbf{H}_t = \left[\frac{dh}{d\boldsymbol{\beta}} \right]_{\boldsymbol{\beta}=\boldsymbol{\beta}_{t|t}}$$

is the measurement Jacobian matrix used for linearization of the observation function $h(\mathbf{x}_{i,j,t}, \mathbf{z}_{i,j,t}, \boldsymbol{\beta}_t)$, $\mathbf{K}_t$ is known as the Kalman gain, and $\mathbf{R}_t$ is a covariance matrix of observations at time t , which depends on the distribution of observations. Specifically, for Bernoulli observations $R_{i,j,t} = (1 - \hat{\pi}_{i,j,t}) \hat{\pi}_{i,j,t}$ and for Positive Poisson

$$R_{i,j,t} = \frac{\hat{\lambda}_{i,j,t}}{1 - \exp(-\hat{\lambda}_{i,j,t})} \left(1 + \hat{\lambda}_{i,j,t} + \frac{\hat{\lambda}_{i,j,t}}{1 - \exp(-\hat{\lambda}_{i,j,t})} \right).$$

2.3.3. Initializing $\mathbf{F}$ and $\mathbf{Q}$

In practice, the state transition matrix $\mathbf{F}$ and the state covariance matrix $\mathbf{Q}$ are unknown, and thus must be estimated based on an in-control sequence of networks $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_T$; we do so using the following heuristic algorithm. First, for each in-control network observation $\mathbf{w}_i$, we estimate the Hurdle model coefficients by fitting a static logistic regression on edge occurrence and a static positive Poisson regression model on positive edge weights to obtain initial vector parameters $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_T$. This sequence of estimated coefficients can be seen as a multi-dimensional time-series. Using the multi-dimensional time-series representation, a Vector Auto-Regression (VAR) model is fitted to yield initial values $\mathbf{F}_0$ and $\mathbf{Q}_0$ for the Bernoulli model and $\mathbf{F}_1$ and $\mathbf{Q}_1$ for the Positive Poisson model. The VAR model is a common approach for analysis of multivariate time series, which is an extension of the univariate autoregressive model to multivariate time series (Sims, 1980). We fit the VAR on coefficients for each model using R (Pfaff 2008). The overall estimation uses these initializations as inputs to the following procedure, which is repeated until convergence. First, we fit the state-space Hurdle model to the in-control data using the initializations $(\mathbf{F}_0, \mathbf{Q}_0)$ and $(\mathbf{F}_1, \mathbf{Q}_1)$. The estimated state-space Hurdle model yields new coefficient estimates for each network observation, i.e., $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_T$. Once again, the estimated coefficients for each stage of the Hurdle model can be organized into a multidimensional time-series. Another VAR model is fitted on the estimated coefficients to update $(\mathbf{F}_0, \mathbf{Q}_0)$ and $(\mathbf{F}_1, \mathbf{Q}_1)$. This procedure is continued until convergence. In summary, the estimation process for each model is as follows:

- Step 0 Initialize the Hurdle model coefficients $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_T$ by fitting a static logistic regression on edge occurrence and a static positive Poisson regression model on positive edge weights for each in-control network observations $\mathbf{w}_i$.
- Step 1 Estimate $\mathbf{F}$ and $\mathbf{Q}$ using a first-order VAR model to solve Equation (5). Here, for $\mathbf{F}_0$ we have $\boldsymbol{\beta}_{0,t} = \mathbf{F}_0 \boldsymbol{\beta}_{0,t-1} + \epsilon_t$. Similarly, for $\mathbf{F}_1$ we have $\boldsymbol{\beta}_{1,t} = \mathbf{F}_1 \boldsymbol{\beta}_{1,t-1} + \epsilon_t$. After fitting the VAR model, $\mathbf{F}_0$ (and $\mathbf{F}_1$) will be the model coefficients matrix, and $\mathbf{Q}_0$ (and $\mathbf{Q}_1$) will be estimated as the covariance matrix of the model's error.
- Step 2 Using estimated $\mathbf{F}$ and $\mathbf{Q}$ from the step above, use the EKF prediction and update steps explained in Sections 2.3.1 and 2.3.2 to estimate the updated set of coefficients $[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_T]$.
- Step 3 Return to Step 1 and repeat Step 1 and 2 until convergence or until the maximum number of iterations is reached.

2.4. Monitoring of dynamic and sparse network stream

In this section, we propose a monitoring procedure to detect structural changes in sparse attributed networks.

Recall that we fit the data incrementally as it arrives by combining the steps previously outlined. Specifically, the Hurdle model is estimated with its parameters updated at each time point via the EKF. The vector of updated parameters $\boldsymbol{\beta}_{t|t-1}$ are used to predict the upcoming adjacency matrix, i.e., $\hat{w}_{i,j,t} = h(x_{i,j,t-1}, z_{i,j,t-1}, \boldsymbol{\beta}_{t|t-1})$ using Equation (5). Once the network at time t is realized, we compute residuals, defined as $\hat{\epsilon}_{i,j,t} = w_{i,j,t} - \hat{w}_{i,j,t}$, for every possible edge. The residuals reflect network connectivity that cannot be explained by the independent variables in the Hurdle model.

To ensure the residuals have approximately constant unit variance, we use the Pearson residual, denoted by $r_{i,j,t}$, which is computed as,

$$r_{i,j,t} = \frac{\hat{\epsilon}_{i,j,t}}{\sqrt{\text{var}(\hat{w}_{i,j,t})}},$$

where $\text{var}(\hat{w}_{i,j,t})$ is the estimated variance of the observation, which can be calculated using the predicted observation and its probability distribution (Positive Poisson or Bernoulli).

If the process is in-control, Pearson residuals asymptotically follow an independent, standard normal distribution. Hence, we can use SPC charts to monitor the residuals and detect change points. We choose the EWMA control chart for monitoring, which is a control chart used to monitor small shifts in a process by incorporating memories of the previous observations in calculating the monitoring statistic. The EWMA weights observations in a geometrically declining order, such that the newest observations have higher weights while the oldest ones have much smaller weights. At each time, we have a total of m residuals calculated where m

is the total number of edges in a network. To calculate one collective statistic, we define $\bar{r}_t = \frac{\sum_{i,j} r_{i,j,t}}{m}$ and monitor $\bar{r}_t$ over time.

The EWMA statistic corresponding to $\bar{r}_t$ is denoted by ω_t and calculated as

$$\begin{aligned} \omega_t &= \lambda \bar{r}_t + (1 - \lambda) \omega_{t-1} \geq 1 \\ \omega_0 &= 0, \end{aligned} \quad (6)$$

where $\lambda \in [0, 1]$ is a constant that specifies the depth of memory. A higher lambda gives more weight to the current observation, and smaller λ gives more weight to previous observations. The control limits are defined as

$$\begin{aligned} UCL &= \mu_0 + k\sigma_0 \sqrt{\frac{\lambda}{2 - \lambda}} \\ LCL &= \mu_0 - k\sigma_0 \sqrt{\frac{\lambda}{2 - \lambda}}, \end{aligned}$$

where μ_0 and σ_0 are the mean and standard deviation of off-line (training) Pearson errors, and k is a parameter that controls the width the control limits. If $\omega_t > UCL$ or $\omega_t < LCL$, we reject the null hypothesis of a stable network process, indicating a change has occurred in the network stream.

Note that the values of k and λ effectively set the (false) alarm rate and are adjustable. In our analysis, we set these parameter values by estimating the false alarm rate for given values of k and λ in an in-control state. Specifically, an in-control stream of networks is simulated to match the observed data as closely as possible, and then monitored until an out-of-control alarm is falsely raised. The time until the false alarm is raised is called the run-length. The simulation procedure is repeated several times to compute the average of the run-lengths (ARL), which indicates the average number of observations until a false alarm is raised. In our analysis, we set the values of k and λ so that the control limits for the in-control ARL is equal to 200 ($\alpha \sim 0.005$). Upon the detection of a change based on an aggregated statistics obtained from the prediction errors, one can identify the edges with the highest prediction errors (say top r) and consider those as the root-cause.

3. Performance evaluation using simulation

In this section, we evaluate the performance of our monitoring methodology against benchmark methods through simulation. We start by describing the simulation setup, evaluation criteria, and alternative methods, followed by the results.

3.1. Simulation setup

To closely match our real data setting in terms of size of networks and number of explanatory covariates, each simulated network is composed of 50 nodes, hence there are $50 \times (50 - 1) = 2450$ potential directed edges in each network. We assume that the number of interactions between nodes i and j is a function of five attributes in the model

denoted as $\mathbf{x}_{i,j,t} = \mathbf{z}_{i,j,t} = [x_{i,j,t}^{(1)}, x_{i,j,t}^{(2)}, \dots, x_{i,j,t}^{(5)}]^T$. The attribute values vary for each edge and time, and are generated using normal distribution with mean $\mu = [0.5, 0.5, 0.5, 0.5, 0.5]^T$ and variance $\Sigma = 0.25 \times \mathbf{I}_{5 \times 5}$. The relationship between the attributes and the response value follows a dynamic Hurdle model. Therefore, we assume that the binary outcomes (whether there is a connection between two nodes) have a Bernoulli distribution, with probability $\pi_{i,j,t} = 1 - \text{logit}^{-1}(\mathbf{x}_{i,j,t} \boldsymbol{\beta}_{0,t})$, and the positive edge weight outcomes follow a Positive Poisson distribution with $\lambda_{i,j,t} = \exp(\mathbf{x}_{i,j,t} \boldsymbol{\beta}_{1,t})$. Here, $\boldsymbol{\beta}_{0,t} = [\beta_{0,t}^0, \beta_{0,t}^1, \beta_{0,t}^2, \dots, \beta_{0,t}^5]^T$ are the coefficients of the binary model at time t , where $\beta_{0,t}^0$ is the coefficient corresponding to the intercept. Similarly, $\boldsymbol{\beta}_{1,t} = [\beta_{1,t}^0, \beta_{1,t}^1, \beta_{1,t}^2, \dots, \beta_{1,t}^5]^T$ are the coefficients of the Positive Poisson model at time t .

To simulate a dynamic stream of networks, we assume the underlying state transition model with $\boldsymbol{\beta}_{0,t} = \mathbf{F} \boldsymbol{\beta}_{0,t-1} + \boldsymbol{\epsilon}_{0,t}$ and $\boldsymbol{\beta}_{1,t} = \mathbf{F} \boldsymbol{\beta}_{1,t-1} + \boldsymbol{\epsilon}_{1,t}$. Here, we set $\boldsymbol{\epsilon}_{0,t} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ and $\boldsymbol{\epsilon}_{1,t} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$. In the simulations, we use $\boldsymbol{\beta}_{0,0} = [0.01, 0.01, 0.01, 0.01, 0.01, 0.01]$ and $\boldsymbol{\beta}_{1,0} = [0.2, 0.2, 0.2, 0.2, 0.2, 0.2]$, $\mathbf{F} = 0.8 \mathbf{I}_{6 \times 6}$, and $\mathbf{Q} = 0.25 \mathbf{I}_{6 \times 6}$. We use in-control simulated snapshots of networks to estimate the control chart and calculate the EWMA control limits based on methods described in Section 2.

We investigate three scenarios, each of which induces changes to specific coefficients underlying the network process to create out-of-control situations. Specifically, for each selected coefficient β_i , the shift is $\delta\sigma_i$, where δ is a constant representing the magnitude of the shift and σ_i is the standard deviation of the i th coefficient for the in-control situation, which is equal to

$$\sigma_i = \sqrt{\frac{Q_{ii}}{(1 - F_{ii})^2}} = 2.5.$$

Therefore, for the Bernoulli model the changed coefficient will be $\beta_{0,\tau}^i = F_{ii} \beta_{0,\tau-1}^i + \epsilon_{0,\tau}^i + \delta\sigma_i$ and for Positive for the Poisson model the changed coefficient will be $\beta_{1,\tau}^i = F_{ii} \beta_{1,\tau-1}^i + \epsilon_{1,\tau}^i + \delta\sigma_i$ at time τ , i.e., the coefficients are shifted by $\delta\sigma_i$ at time τ . The three scenarios are as follows:

1. Scenario I represents a case where the change point is affecting the underlying dynamics in two ways. First, it affects the decision of whether two nodes are interacting. Second, it affects the level of interaction (weights) after the first decision is made. In other words, we assume the change has affected both the Bernoulli and Positive Poisson models. In each model, we apply the change in three out of six coefficients so that coefficients $\beta_{0,\tau}^2, \beta_{0,\tau}^4, \beta_{0,\tau}^5$ of the Bernoulli model, and $\beta_{1,\tau}^2, \beta_{1,\tau}^4, \beta_{1,\tau}^5$ of the Positive Poisson model change at time τ .
2. Scenario II represents a case where the change only affects the decision of whether two nodes interact. However, after this decision is made, the coefficients of the Positive Poisson model remain unchanged.

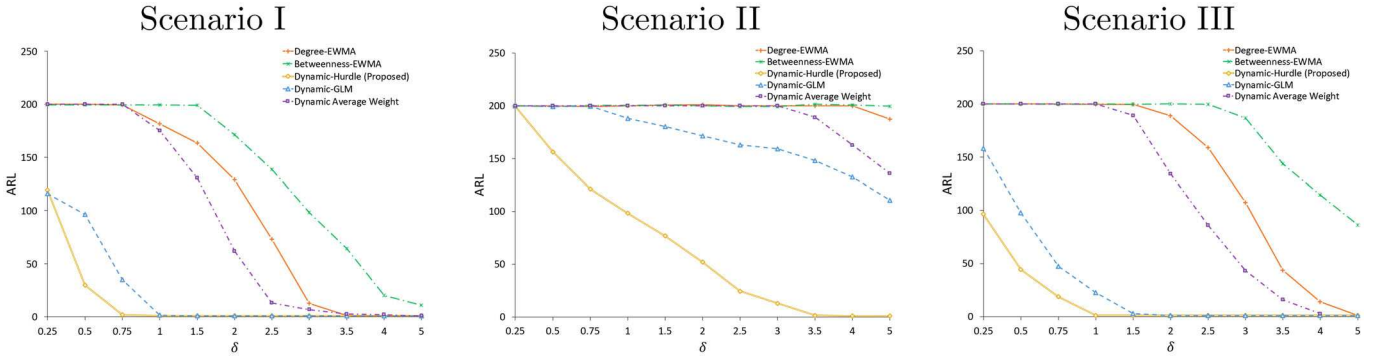


Figure 2. ARL comparison of methods based on simulated data for different magnitude of shifts (δ). Smaller ARL indicates better performance.

Therefore, we suppose that the change has only affected the coefficients of the Bernoulli model. Accordingly, we assume that the change has affected $\beta_{0,\tau}^2$, $\beta_{0,\tau}^4$, $\beta_{0,\tau}^5$ from the Bernoulli model.

3. Scenario III represents a case where the change only affects the amount of interaction (edge weight) between two nodes. The coefficients determining whether two nodes interact remains unchanged, but $\beta_{1,\tau}^2$, $\beta_{1,\tau}^4$, $\beta_{1,\tau}^5$ from the Positive Poisson model have changed at time τ .

Performance of a method is evaluated with the ARL, which measures how quickly the method detects the change induced in each scenario for different values of δ . Specifically, we simulate networks for an out-of-control scenario until an out-of-control alarm is raised. Each time the alarm is raised, we record the run length, which is the number of simulated time points until the change is detected. We iterate this procedure 1000 times and record the ARL over all iterations. A method with the smallest ARL for an out-of-control situation represents superior ability in detecting the change.

We compare the proposed approach against four benchmark methods. The first two benchmark methods monitor network connectivity measures using EWMA control charts; the particular network statistics that we consider, degree and betweenness, have been utilized previously for change point network analysis in financial economics (Adamic *et al.*, 2017) and related areas, such as social networks (McCulloh and Carley, 2011; Hassanzadeh *et al.*, 2012). Specifically, we compute the average of each network statistic among all in-control network observations. Then, we use the average statistics to build the EWMA statistics as explained in Equation (6). The third benchmark treats the weights $\mathbf{w}_{i,j,t}$ as a vector and calculates the mean over the last l_t observed networks as the prediction of the upcoming network (here, we used $l_t = 10$). Afterwards, for each time step t , we calculate errors as the difference between the current $\mathbf{w}_{i,j}$ and the predicted values. Finally, we monitor the mean of the errors using an EWMA control chart similar to our proposed method. For simplicity, we denote this benchmark by dynamic average weight monitoring. The fourth benchmark utilizes the closest extant Statistical method to our knowledge, the “Dynamic GLM” approach proposed by Reisi Gahrooei and Paynabar (2018), which utilizes a Poisson

distribution to model the count data without accounting explicitly for extreme sparsity. Edge weights are assumed to have a Poisson Distribution where the mean of the distribution is a function of network attributes $x_{i,j,t}$. The rest of the Dynamic GLM (using EKF, EWMA, etc.) are similar to our proposed approach.

Finally, note that to ensure that all methods can be compared fairly using the out-of-control scenarios, we specify control limits such that the in-control ARL for all methods is equal to 200 ($\alpha = 0.005$) and subsequently use these tuned control limits for detection in the out-of-control scenarios.

3.2. Simulation results

The ARL results are shown in Figure 2 for all scenarios. As can be readily observed, for all scenarios, monitoring network connectivity measures (degree and betweenness) has the highest ARL (worst performance), showing that static methods should not be preferred for the explicit purpose of detecting changes in networks. We also see from the figure that all methods perform better in the first scenario, where the coefficients are shifted in both the Bernoulli and Positive Poisson models. Yet, in this scenario, we can see that for a change as small as $\delta = 0.75$ our method has $ARL \approx 2$, whereas the Dynamic GLM method has $ARL \approx 35$, and EWMA monitoring of network statistics and dynamic average weight monitoring have $ARL \approx 200$. Hence, for this small shift, our method almost instantly detects the change whereas it takes on average 35 runs for the Dynamic GLM method to detect this shift; the other methods are not capable of detecting this small shift.

For Scenario II, all methods have higher ARL values in comparison with other scenarios, which is due to the shift being solely imposed on the existence of an edge while the edge weights remained intact, hence detecting such a shift more challenging. In this case, approaches based on monitoring network connectivity measures are not able to detect changes with even large δ shifts. Furthermore, we again see that the dynamic GLM method has a significantly higher ARL in comparison with our proposed method. For example, for a shift with $\delta = 3.5$, our method has $ARL \approx 1.5$, whereas the Dynamic GLM method has $ARL \approx 148$, Dynamic average weight monitoring has $ARL \approx 189$, and EWMA monitoring of network statistics has $ARL \approx 200$.

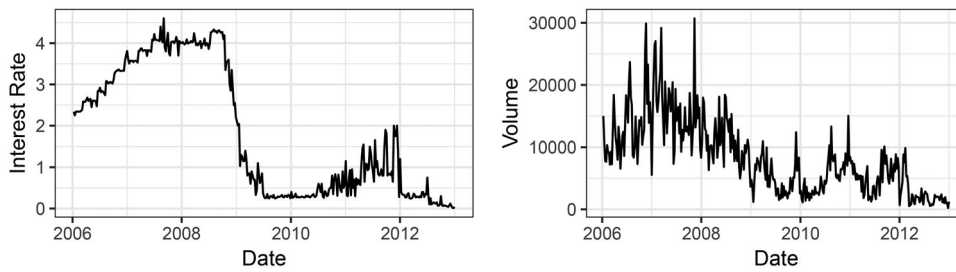


Figure 3. Weekly interest rate and volume in the e-MID interbank market.

For Scenario III, all methods have slightly higher ARL values in comparison with the first scenario, which is because the shift is imposed only on the Positive Poisson model (the edge weights) while the Bernoulli model (decision to connect) remains intact. We can again observe that for a change as small as $\delta = 1$ our method almost instantly detect the change ($ARL \approx 1$), whereas the Dynamic GLM method has $ARL \approx 22$ and EWMA monitoring of network statistic and dynamic average weight monitoring have $ARL \approx 200$.

Thus, by comparing our proposed method with other methods in all three scenarios, we observe that our proposed method has the lowest ARL for different magnitudes of shifts in all the scenarios. Hence, we see evidence that monitoring network statistics or using dynamic average weight for prediction are not well suited for an abrupt change point detection. We also see that our proposed approach outperforms the dynamic GLM model proposed by Reisi Gahrooei and Paynabar (2018) in all cases. The gap between our proposed method and the benchmark methods is particularly pronounced in Scenario II, emphasizing the importance of appropriately modeling the sparse nature of the data.

4. Case study: Monitoring the interbank market during the 2007 financial crisis

As a consequence of the 2007–2009 financial crisis, according to the U.S. Financial Crisis Inquiry Commission’s final report in January 2011, 8 500 000 families lost their homes in foreclosure or were seriously behind on their mortgage payments (Financial Crisis Inquiry Commission, 2011). The unemployment rate peaked at about 10% in October 2009 (U.S. Government Accountability Office, 2013), and the stock market suffered record losses, with the S&P 500 Index losing half of its value between October 2007 and March 2009. Nearly half a trillion dollars of taxpayer money was spent in the United States to stabilize the financial economic system (U.S. Department of the Treasury, 2016). Indeed, the financial crisis induced large societal costs in the form of slower economic growth and direct bailouts, and has thus clearly accentuated the need for more effective monitoring and oversight of financial markets and institutions.

Researchers have responded to this call by analyzing financial networks to capture the interconnectedness among financial institutions. A financial network describes a collection of financial institutions (nodes) and the links between them. Edges in financial networks reveal information about the underlying balance sheets of the connected firms. Thus,

a central idea in financial network analysis is to draw insights about the level of systemic risk from connectivity patterns, e.g., a sparsely connected interbank lending network can indicate that banks have stopped participating in the interbank market due to higher perceived counter-party risk, which has systemic risk implications. Supporting the notion of monitoring financial networks over time for risk management, multiple works have shown empirically that network statistics, such as the average of the network degree distribution, can shift depending on stable or crisis market conditions (Finger *et al.*, 2013; Squartini *et al.*, 2013; Brunetti *et al.*, 2019).

Although the literature has established the importance of characterizing network topology for early warning systems, to our knowledge explicit methodology to systematically identify in real-time whether the network has entered a new state has not yet been developed. We address this gap in the literature by demonstrating our methodology on the e-MID trading platform, the only electronic regulated interbank market in the world, from January 2006 to December 2012. Edges are defined by the number of overnight loans between European banks on this platform, that is, if Bank A lends to Bank B, then an edge is drawn from Bank A to Bank B and is weighted by the number of directed loans in the given week.

4.1. Data

The e-MID market is open to all banks allowed to operate in the European interbank market. As of August 2011, e-MID had 192 members from European Union countries and the U.S., including 29 central banks acting as market observers (Finger *et al.*, 2013). Our data contains all e-MID transactions from January 2006 through December 2012. Each transaction includes the date, lender, and borrower (with their real names anonymized), country of origin for lender and borrower, interest rate, quantity, and an indication of which party initiated the trade. The data includes 40–60 banks that are publicly-traded. Please note that bank identities in the interbank market are confidential and therefore the exact number is not given to protect confidentiality. For these banks, we also observe their weekly returns in the stock market.

Figure 3 shows weekly volume and interest rates in the e-MID market. As the financial crisis progressed, interest rates dropped to near zero and activity in the market decreased markedly. In fact, the changes in these financial variables reflect major real-world events. For example, using

Table 1. Independent variables used in the Hurdle mode for predicting week t .

<i>Node attributes</i>	
<i>Variable</i>	<i>Description</i>
Lender's return	The average stock market return of the lending bank in week $t - 1$
Borrower's return	The average stock market return of the borrowing bank in week $t - 1$
<i>Edge attributes</i>	
<i>Variable</i>	<i>Description</i>
Return Correlation	The correlation between the two banks' returns from the start of the data up to week $t - 1$
Number of Trades	The average number of transactions between two banks in week $t - 1$ (if any transaction occurred)
Rate	The average interest rate of each loan between two banks in week $t - 1$ (if any transaction occurred)
Country Difference	An indicator variable that is one if two banks are from different countries

the same data, Brunetti *et al.* (2019) analyze four sub-periods: (i) a pre-crisis period from January 2, 2006, until August 7, 2007, when the European Central Bank (ECB) noted worldwide liquidity shortages; (ii) the first crisis period from August 8, 2007 until September 12, 2008, when Lehman Brother's collapsed; (iii) the second crisis period from September 16, 2008, through April 1, 2009 when the ECB announced the end of the recession; and (iv) post-recession period, from April 2, 2009 onwards. Note that as opposed to monitoring in real time, previous works performed historical analyses that utilized ex-post information about when critical events occurred. We utilize these four sub-period definitions to validate our monitoring results.

Several works have focused on characterizing the general network structure within the e-MID interbank lending market. For instance, Fricke and Lux (2015a) summarize the degree distribution as heavy-tailed (negative Binomial), but not power-law or scale-free. Fricke and Lux (2015b) find that networks from the e-MID consistently exhibit so-called core-periphery structure, where a temporally stable core of banks that comprise 20–30% of the market is actively engaged in both lending and borrowing. During the financial crisis, the authors find that the reduction of e-MID activity was mainly due to these core banks reducing their trading. Finger *et al.* (2013) find that the level of temporal aggregation is an essential methodological choice, where daily-level network analysis of e-MID interbank lending data looks almost random and uninformative, but meaningful and significant non-random structures appear for longer aggregation periods. As such, in this article, we present a weekly analysis to improve both interpretability and practical utility for regulators and market participants. The networks are created by connecting lender to borrowers, that is, if Bank A lends to Bank B, then an edge is drawn from Bank A to Bank B and weighted by the number of directed loans.

4.2. Model specification

Motivating our use of the Hurdle model, we observe a high level of sparsity (lack of node connections) in the e-MID market networks. During the pre-crisis era, there is 73.68% sparsity among all possible bank interactions. This value increases to 77.12% and 83.78% during Crisis 1 and Crisis 2 sub-periods, respectively. In the post-crisis era, the level of sparsity is 76.87%, nearly returning to pre-crisis levels. Note also that the two-stage process underlying the Hurdle model

mimics the actual decision that a bank would make about whether to participate in the e-MID market with another bank in addition to modeling the number of interactions between banks in a second stage.

We use a number of nodes (bank) and edge attributes listed in Table 1 as independent variables in the Hurdle model, including whether the two banks are originating from different countries, the interest rate in the e-MID market, returns correlation, and so on. Country Difference is motivated by Finger *et al.* (2013), who found that Italian banks tended to trade with other Italian banks a vast majority of the time. Thus, we expect this variable to be significant, especially for modeling whether two banks have any trades with each other. Most of the other variables are based fundamentally on stock market returns. Bank activity in the interbank market can be influenced by stock market performance, especially when the impact on the bank's balance sheet is large (Brunetti *et al.*, 2019). As such, we expect returns-based variables to be important, particularly when the stock market is volatile, i.e., in crisis sub-periods. Note that we are not including Number of Trades and Rate in the logistic regression model, as this information is generated only after a transaction occurs.

4.3. Monitoring results: Uncovering the epochs of the financial crisis

We use a hierarchical analysis approach to simulate how the methodology would have performed if implemented in real time. Specifically, the first 20 weeks of pre-crisis data are used as in-control observations to obtain appropriate values of $\mathbf{F}$ and $\mathbf{Q}$ as well as initial estimates of the regression coefficients. Also based on the Pearson residual errors from the in-control data, we determine the control limits for the in-control ARL to be equal to 200 ($\alpha \sim 0.005$). Starting with week 21, we enter the online monitoring phase (see Figure 1 to review the methodology). Whenever a change point is detected, the entire methodology is restarted, with retraining of all model parameters using the next 20 weeks as in-control data before entering another online monitoring phase.

Figure 4 shows the EWMA control charts for Logistic and Positive Poisson regressions, where we can see that both aspects of the Hurdle model “raise the alarm” before the official announcement that marks the beginning of the first crisis sub-period. Specifically, the Positive Poisson model would have alerted regulators and financial institutions the

week starting on May 28, 2007, and the Bernoulli model would have raised an alarm the week of June 18, 2007.

To compare the performance of our method with network statistics commonly used in financial economics (Finger *et al.*, 2013; Adamic *et al.*, 2017; Basu *et al.*, 2017; Brunetti *et al.*, 2019), we provide the EWMA control charts for monitoring network's average degree and betweenness in Figure 5. Network degree centrality is defined by the number of connections attached to each node. We can calculate the degree centrality of node i by the sum of the weights of edges starting from node i . Hence $d_i^t = \sum_{j \neq i} w_{i,j,t}$. The network overall degree at time t is the average of degrees over all nodes, hence $d^t = \bar{d}_i^t$. Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. For a node i the betweenness centrality is calculated as $c_i^t = \sum_{j,k} \frac{\sigma_{jk(i)}}{\sigma_{jk}}$ where σ_{jk} is the number of shortest paths from node j to k and $\sigma_{jk(i)}$ is the number of those paths that pass through node i . Similarly, the network overall betweenness is the average of betweennesses over all nodes, hence $c^t = \bar{c}_i^t$.

As the results show, these control charts fail to raise the alarm before an official announcement of the first crisis sub-

period, providing evidence that network statistics may serve as a weak foundation for an early-warning system.

When monitoring using our approach, since the EWMA statistic is representing the difference between the actual and estimated value, observing its trend during the first crisis can help us interpret the changes in comparison with the Pre-Crisis era. For the Bernoulli model, the EWMA statistic is sharply negative at the onset of the first crisis sub-period, which means that the model over-estimates the existence of edges (loans between banks). In other words, for two banks, the probability that they have any transactions sharply decreases at the start of the crisis. Adding evidence that activity in the interbank market dropped precipitously, as shown in Figure 6, the regression coefficient for Country Difference is consistently negative, indicating that banks generally prefer to trade with other banks based in the same country. The U shape shows that as the crisis unfolded, banks became even less willing to have transactions with banks from other countries, but as the crisis concluded and in the Post-Recession sub-period trading activity (specifically counter-party trust) was returning to normal. Similarly, for the Positive Poisson model, we can see, in contrast with activity during the crisis, an apparent increasing trend within the Post-Recession sub-period with coefficients ending close to zero. Thus, by the end of 2012, the number of transactions among two connected banks is not affected by country differences. The estimated coefficients for the Number of Trades and Rate variables in Figure 7 also show meaningful patterns. There is a clear increasing trend in the coefficient for Number of Trades, denoting that post-recession, banks were able to obtain more funding in comparison with before this era. The coefficient for Rate was positive during the crisis, but negative in the post-recessionary period. One potential explanation is that banks that wanted funding had to pay higher rates during the crisis (i.e., it was a lender's market), but after the crisis, interbank funding was more readily available. Estimated coefficients for other independent variables are not shown, since they were centered on zero without meaningful trends. Overall, in addition to detecting the onset of the crisis in real-time, the

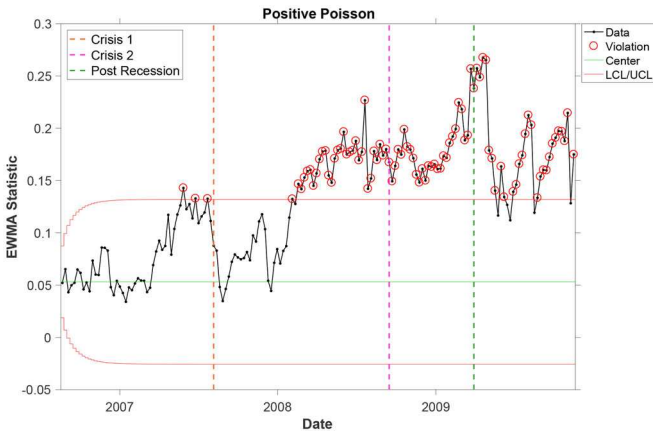


Figure 4. EWMA charts for Pearson Residuals from the proposed model to detect the onset of Crisis 1.

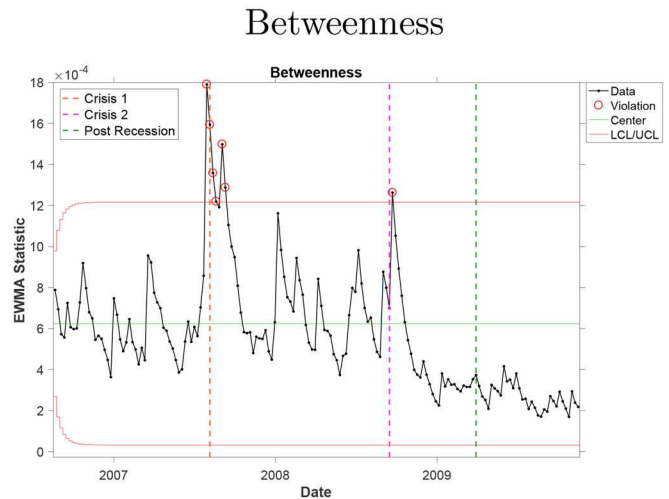
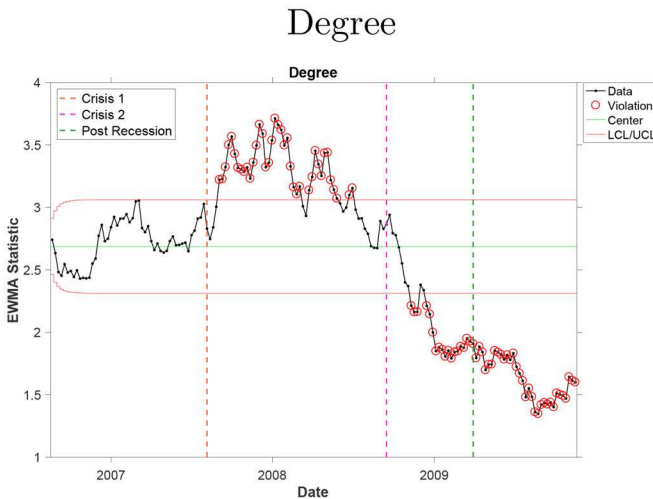


Figure 5. EWMA charts for networks statistic (degree, and betweenness) to detect the onset of Crisis 1.

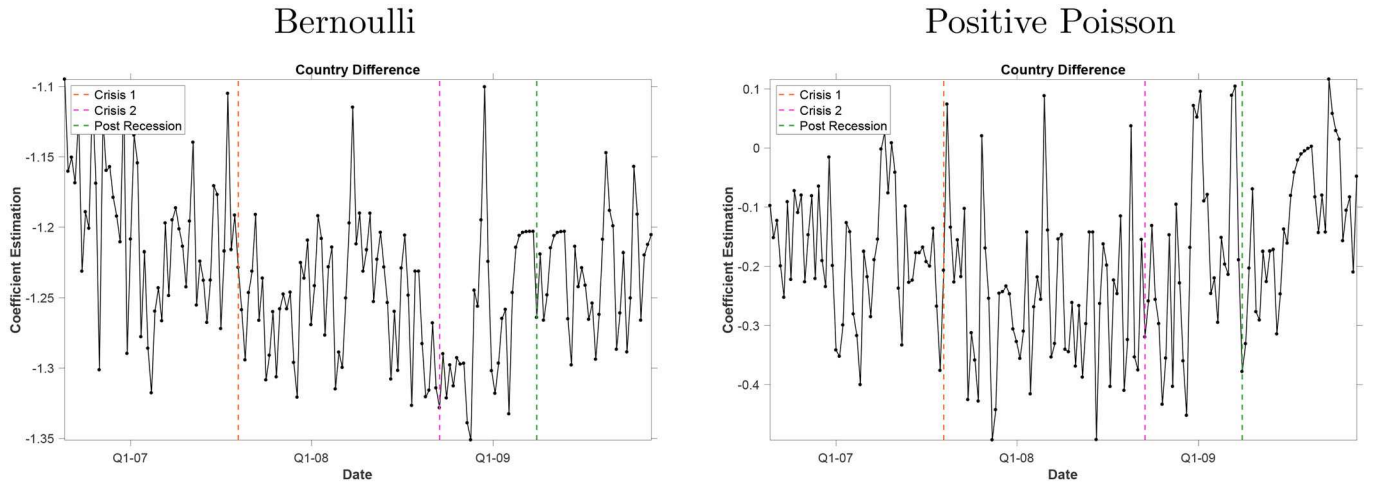


Figure 6. Estimated coefficients for Country Difference in the Bernoulli and Positive Poisson models starting from Pre-Crisis data.

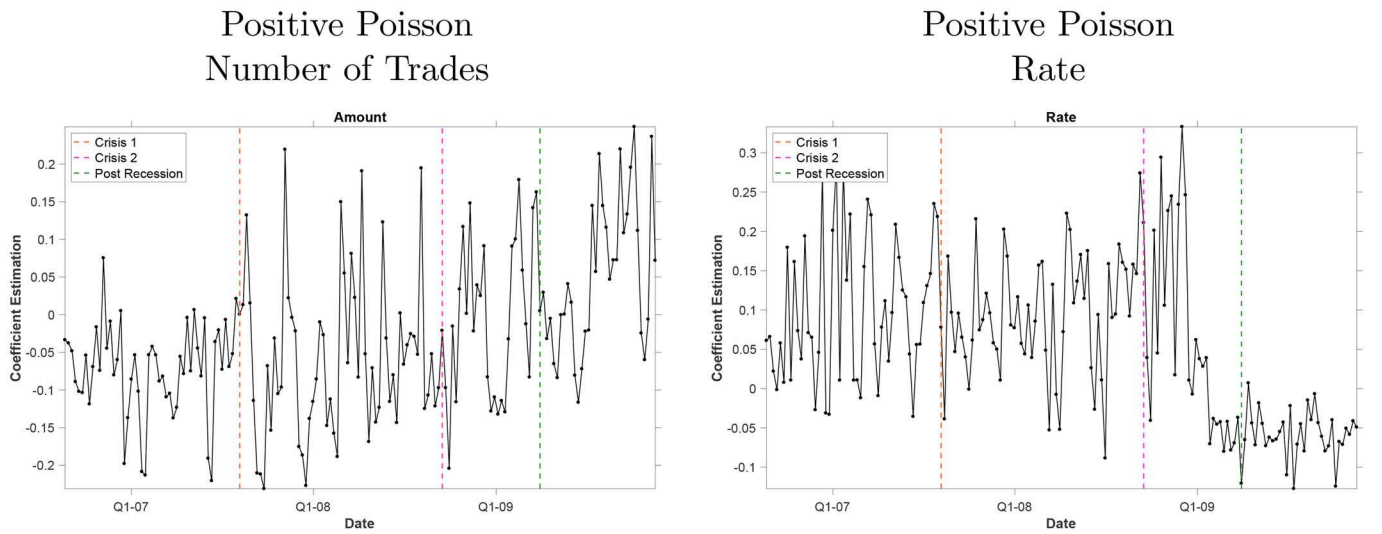


Figure 7. Estimated coefficients for Number of Trades and Rate in the Positive Poisson model starting from Pre-Crisis data.

detailed results are consistent with accepted narratives about the crisis (Finger *et al.*, 2013; Brunetti *et al.*, 2019), where trust in the interbank market dropped markedly causing banks to stop participating entirely in the e-MID market, followed by a return to pre-crisis conditions.

In practice, it can be important to identify possible root causes for the alarms, which can be accomplished by estimating the actual date of the change point. Doing so through network analysis is an important novelty and innovation of this work to the financial network analysis literature. Nishina (1992) proposed a method to estimate the EWMA change point after receiving an out-of-control signal at time T . In this method, if the out-of-control signal is raised when the monitoring statistic is above the Upper Control Limit (UCL), the estimated change point is $\hat{\tau} = \max[i : z_i \leq \mu_0, i \leq T]$, i.e., the estimated change point $\hat{\tau}$ is the first point before the alarm time when the EWMA statistic is below the center line μ_0 . Similarly, if the out-of-control signal is raised when the monitoring statistic is below the Lower Control Limit (LCL), the estimated change point is $\hat{\tau} = \max[i : z_i \geq \mu_0, i \leq T]$, i.e., the estimated change point $\hat{\tau}$ is the first point before the alarm time when the

EWMA statistic is above the center line μ_0 . Using this heuristic, we find a change to crisis conditions in the Bernoulli model dated to the week of May 28, 2007, and for the Positive Poisson model dated to the week of March 12-16, 2007.

Moving to the detection of the onset of Crisis 2 and the Post-Crisis eras, we use the first 20 weeks of Crisis 1 era for offline training. The EWMA control charts for Bernoulli and Positive Poisson regression are shown in Figure 8, where we see that both parts of the Hurdle model do not detect any change in the interbank market around the failure of Lehman Brothers (the onset of Crisis 2). Both models do successfully capture the change before the onset of Post of the Recession announcement. The Bernoulli model raises the alarm the week of January 5, 2009, with the change point dated to October 6-10, 2008. The Positive Poisson model raises the alarm the week of December 15, 2008, with its change point dated back on November 3, 2008.

Inspecting the trend in the EWMA statistics, we see increases for the Bernoulli and Positive Poisson models before the onset of the Post-Recession period that continue to the end of the data, demonstrating that banks returned to

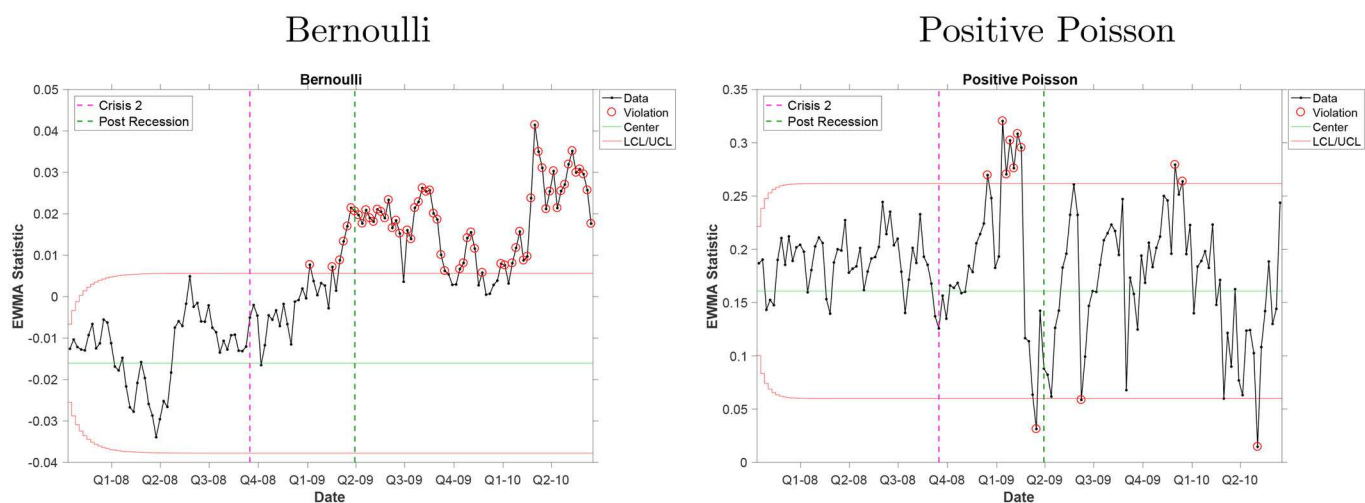


Figure 8. EWMA charts for Pearson Residuals from the proposed model to detect the end of the financial crisis and start of the post-recessionary period.

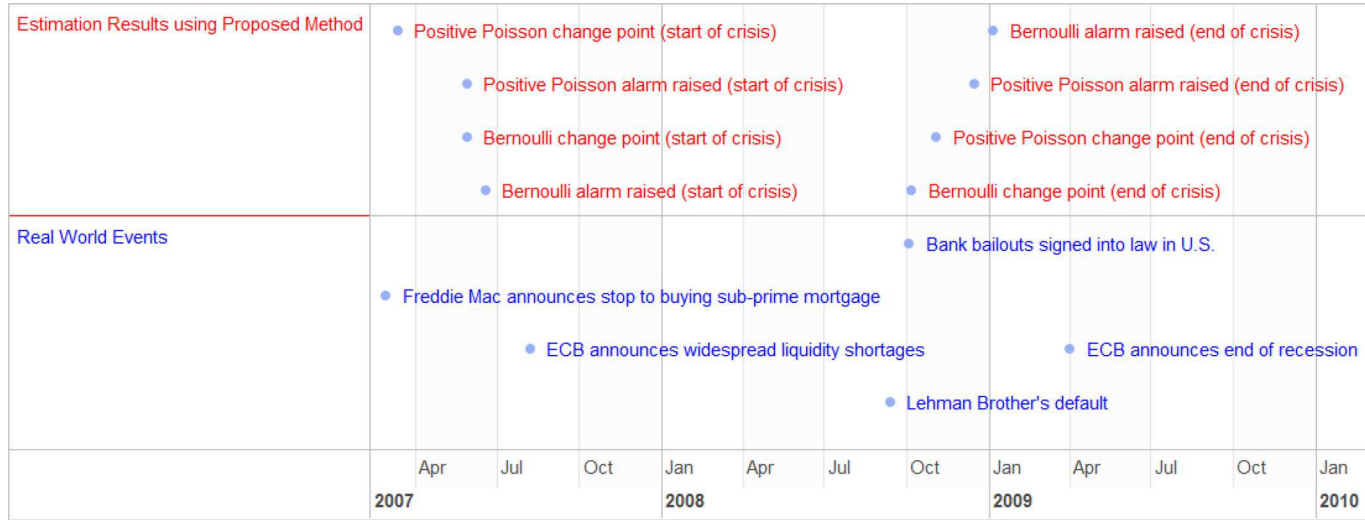


Figure 9. Timeline of estimation results and main events in the financial crisis. The proposed monitoring framework would have raised alarms in real time to regulators and financial institutions about changes in interbank market conditions that coincide with the onset and end of the crisis.

the market as overall conditions and trust levels improved. In unreported results, a very similar pattern emerges as previously reported when inspecting the estimated coefficients for independent variables.

4.4. Validation of results with ECB activity

As shown in Figure 9, our monitoring approach applied to the e-MID data results in practically meaningful results. In fact, a regulator could have raised the alarm in real time about the change from calm to crisis conditions during the week of June 18, 2007 using our methodology. Note that this is well before August 8-9, 2007, which is widely considered the official recognition of the crisis when central banks around the world announced significant liquidity shortages (Brunetti *et al.*, 2019). This is a notable finding given that it is difficult to correctly identify the onset of the crisis from typical financial variables in our data (see Figure 3) and that previous research using network analysis on the same data had difficulty correctly identifying this moment as the

beginning of the crisis. Indeed, Finger *et al.* (2013, p.205) conduct a detailed network analysis of data from the same market and remark that “The start of the GFC [global finance crisis] is not easy to determine [from the data] ...”

Although we do not know precisely when governments realized internally there was a crisis, we find evidence that the model results were ahead of official policy. For instance, both aspects of the Hurdle model raised alarms by June 18, 2007. On June 6, the ECB raised interest rates, followed by the publication of the Financial Stability Review (European Central Bank, 2007) on June 15, which struck a cautiously optimistic tone. The Financial Stability Review included positive outlooks, with statements such as:

Looking forward, with the euro area financial system in a generally healthy condition and the economic outlook remaining favorable, the most likely prospect is that financial system stability will be maintained in the period ahead.

Such forecasts were coupled with warnings about how the financial system was growing particularly vulnerable “to an abrupt and unexpected sharp decline in market liquidity”,

which underscores the importance of our results given that we are studying an interbank market – a key source of liquidity for banks.

Similarly, the alarms signifying the end of the crisis using the proposed methodology is raised between December 15, 2008, to January 5, 2009. Coincidentally, the ECB published another Financial Stability Review on December 15 (European Central Bank, 2008) stating that:

The extraordinary remedial actions taken by central banks and governments, which are aimed at addressing liquidity stresses and strengthening capital positions, thus contributing to restoring confidence in, and improving the resilience of financial systems, were successful in stabilising the euro area banking system.

This period also coincides with an intense activity by the U.S. Treasury Department as a result of the new law giving it broad new powers (see discussion of the Emergency Economic Stabilization Act below) to strengthen the financial and auto sectors of the U.S. economy.

To further validate our results, we consider whether the identified change point dates match previously reported results or known events. The earliest detected change point for Crisis 1, the week of March 12, 2007, for the Positive Poisson regression, closely follows the root-cause event as identified by the Federal Reserve Bank of St. Louis (2018) of a Freddie Mac announcement on February 27, 2007, that they would no longer buy the riskiest type of mortgages (sub-prime). Similarly, the earliest identified change point signaling the end of the crisis was October 6-10, 2008 for the Bernoulli model, which coincides with the “bank bailouts” (the Emergency Economic Stabilization Act of 2008) being signed into law by then President Bush on October 3, 2008. To our knowledge, this is a new finding within the network analysis literature that shows evidence that U.S. fiscal and monetary policy directly impacted European financial activity and markets, particularly during the 2007–2009 crisis which was U.S.-based.

5. Conclusion

In this article we proposed a new network monitoring system to detect changes within a sequence of sparse and attributed networks. We started by modeling a single network with the Hurdle model, which captures sparsity appropriately while allowing edges to be modeled as a function of node and edge attributes. Then, to capture temporal dynamics of the edge formation process, we allowed the parameters of the Hurdle model to evolve using a state-space model. A sequential estimation scheme relying on the EKF is used to update the state-space parameters and predict the value of upcoming networks. Finally, SPC charts are used to monitor the network sequence in real time in order to identify changes in connectivity patterns that are caused by regime shifts. The proposed approach was validated in two ways. First, with simulation, we established self-consistency and functional performance of the proposed modeling and estimation framework. Second, with a detailed case study, we demonstrated how the methodology can be applied to

monitor interbank lending networks. We found several promising and novel results showing that the proposed model would have raised alarms to the public before official announcements by the ECB. The identified change point dates were also highly interpretable, matching closely with several key real-world events. This article focused on detecting global changes in a network as they are the most crucial ones in our case study of financial networks. Therefore, we take the average over all prediction residuals as the monitoring statistics. Nevertheless, this monitoring approach can be extended for detection of local changes and performing diagnosis as a future line of work.

Notes on contributors

Samaneh Ebrahimi earned her PhD from H. Milton Stewart School of Industrial and Systems Engineering, at Georgia Tech in 2018. She also holds a MSc in statistics from Georgia Tech. Samaneh received her B.Sc. and M.Sc in industrial engineering from Sharif University of Technology in 2011 and 2013, respectively. Her research interests lie on developing effective methodologies for monitoring, diagnostics and prediction of high dimensional complex data including multi-stream data, network data, and multi-media data using statistical data mining and machine learning tools. She is currently a Senior Data Scientist at Pandora Media, LLC.

Mostafa Reisi-Gahrooei received the master's degree in computational science and engineering and the PhD degree in industrial and systems engineering from Georgia Institute of Technology, Atlanta, GA, USA, and the MSc degrees in transportation engineering and applied mathematics from the Southern Illinois University Edwardsville, Edwardsville, IL, USA. He is currently an assistant professor with the Department of Industrial and Systems Engineering, the University of Florida, Gainesville, FL, USA. His research focuses on modeling, monitoring, and control of complex systems with high-dimensional data. Dr. Reisi Gahrooei is a member of the Institute for Operations Research and the Management Sciences (INFORMS) and the Institute of Industrial and Systems Engineers (IISE).

Shawn Mankad is an Assistant Professor in the S.C. Johnson Graduate School of Management at Cornell University. He received his PhD in statistics from The University of Michigan and his BS in mathematics from Carnegie Mellon University. His research develops and applies mathematical and statistical methodology for addressing business, economic, and policy problems. His work has been featured across several domains, including in the *Journal of Financial Economics*, *Manufacturing & Service Operations Management*, and the *Annals of Applied Statistics*, as well as over a dozen popular media outlets, including the *Wall Street Journal* and the *Chicago Tribune*.

Kamran Paynabar is the Fouts Family Early Career Professor and Associate Professor in the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Tech. He received his PhD in IOE and M.A. in statistics from The University of Michigan. His research interests comprise both applied and methodological aspects of machine-learning and statistical modeling integrated with engineering principles. He is a recipient of the Data Mining Best Student Paper Award, the Best Application Paper Award from IIE Transactions, the Best QSR refereed paper from INFORMS, and the Best Paper Award from POMS. He has been recognized with the GT campus level 2014 Junior Faculty Teaching Excellence Award and Provost Teaching and Learning Fellowship. He served as the chair of QSR of INFORMS, and the president of QCRE of IISE. He is an Associate Editor for *Technometrics* and *IEEE-TASE*, a Department Editor for *IISE-Transactions* and a member of editorial board for *Journal of Quality Technology*.

References

- Adamic, L., Brunetti, C., Harris, J.H. and Kirilenko, A. (2017) Trading networks. *The Econometrics Journal*, **20**(3), S126–S149.
- Akoglu, L., Tong, H. and Koutra, D. (2015) Graph based anomaly detection and description: A survey. *Data Mining and Knowledge Discovery*, **29**(3), 626–688.
- Azarnoush, B., Paynabar, K., Bekki, J. and Runger, G. (2016) Monitoring temporal homogeneity in attributed network streams. *Journal of Quality Technology*, **48**(1), 28–43.
- Basu, S., Das, S., Michailidis, G. and Purnanandam, A.K. (2017) A system-wide approach to measure connectivity in the financial sector. Available at <https://ssrn.com/abstract=2816137>.
- Brown, R.G. and Hwang, P.Y.C. (1997) *Introduction to Random Signals and Applied Kalman Filtering: with MATLAB Exercises and Solutions*, Wiley, New York, NY.
- Brunetti, C., Harris, J.H., Mankad, S. and Michailidis, G. (2019) Interconnectedness in the interbank market. *Journal of Financial Economics*, **133**(2), 520–538.
- Cameron, A.C. and Trivedi, P.K. (2013) *Regression Analysis of Count Data*, Cambridge University Press, New York, NY.
- European Central Bank. (2007) ECB financial stability review. <https://www.ecb.europa.eu/press/pr/date/2007/html/pr070615.en.html>. Accessed January 1, 2019.
- European Central Bank. (2008) ECB financial stability review. <https://www.ecb.europa.eu/press/pr/date/2008/html/pr081215.en.html>. Accessed February 20, 2019.
- Fahrmeir, L. and Kaufmann, H. (1991) On Kalman filtering, posterior mode estimation and Fisher scoring in dynamic exponential family regression. *Metrika*, **38**(1), 37–60.
- Faloutsos, C., McCurley, K.S. and Tomkins, A. (2004) Connection subgraphs in social networks, in *SIAM International Conference on Data Mining, Workshop on Link Analysis, Counterterrorism and Security*, SIAM, Philadelphia PA.
- Federal Reserve Bank of St. Louis. (2018) Full timeline. <https://www.stlouisfed.org/financial-crisis/full-timeline>. Accessed January 1, 2019.
- Fienberg, S.E. (2012) A brief history of statistical models for network analysis and open challenges. *Journal of Computational and Graphical Statistics*, **21**(4), 825–839.
- Financial Crisis Inquiry Commission. (2011) Financial crisis inquiry report. http://fcic-static.law.stanford.edu/cdn_media/fcic-reports/fcic_final_report_full.pdf. Accessed February 20, 2019.
- Finger, K., Fricke, D. and Lux T. (2013) Network analysis of the e-MID overnight money market: The informational value of different aggregation levels for intrinsic dynamic processes. *Computational Management Science*, **10**(2-3), 187–211.
- Fricke, D. and Lux, T. (2015a) On the distribution of links in the interbank network: Evidence from the e-MID overnight money market. *Empirical Economics*, **49**(4), 1463–1495.
- Fricke, D. and Lux, T. (2015b) Core-periphery structure in the overnight money market: Evidence from the e-MID trading platform. *Computational Economics*, **45**(3), 359–395.
- Goel, S. and Goldstein, D.G. (2013) Predicting individual behavior with social networks. *Marketing Science*, **33**(1), 82–93.
- Grogger, J.T. and Carson, R.T. (1991) Models for truncated counts. *Journal of Applied Econometrics*, **6**(3), 225–238.
- Hassanzadeh, R., Nayak, R. and Stebila, D. (2012) Analyzing the effectiveness of graph metrics for anomaly detection in online social networks, in *International Conference on Web Information Systems Engineering*, pp. 624–630. Springer, Berlin, Heidelberg.
- Heard, N.A., Weston, D.J., Platanioti, K., Hand, D.J. (2010) Bayesian anomaly detection methods for social networks. *The Annals of Applied Statistics*, **4**(2), 645–662.
- Kalman, R.E. (1960) A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, **82**(1), 35–45.
- McCulloh, I. and Carley, K.M. (2011) Detecting change in longitudinal social networks. Technical report, Military Academy West Point NY Network Science Center, West Point, NY.
- Mullahy, J. (1986) Specification and testing of some modified count data models. *Journal of Econometrics*, **33**(3), 341–365.
- Nelder, J.A. and Baker, R.J. (1972) *Generalized Linear Models*. New York, NY.
- Nishina, K. (1992) A comparison of control charts from the viewpoint of change-point estimation. *Quality and Reliability Engineering International*, **8**(6), 537–541.
- Osadchiy, O., Gaur, V. and Seshadri, S. (2016) Systematic risk in supply chain networks. *Management Science*, **62**(6), 1755–1777.
- Pfaff, B. (2008) VAR, SVAR and SVEC models: Implementation within R package vars. *Journal of Statistical Software*, **27**(4), 1–32.
- Reisi Gahrooei, M. and Paynabar, K. (2018) Change detection in a dynamic stream of attributed networks. *Journal of Quality Technology*, **50**(4), 418–430.
- Sims, C.A. (1980) Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, **48**, 1–48.
- Squartini, T., van Lelyveld, I. and Garlaschelli, D. (2013) Early-warning signals of topological collapse in interbank networks. *Scientific Reports*, **3**, 3357.
- Trier, M. (2008) Research note: Towards dynamic visualization for understanding evolution of digital communication networks. *Information Systems Research*, **19**(3), 335–350.
- U.S. Department of the Treasury. (2016) Citizens' report. https://www.treasury.gov/initiatives/financial-stability/reports/Documents/FY2016/%20OFS/%20CitizensReport_FINAL/%20-%202011.22.16.pdf. Accessed February 20, 2019.
- U.S. Government Accountability Office. (2013) GAO-13-180, financial regulatory reform: Financial crisis losses and potential impacts of the Dodd-Frank act. <http://www.gao.gov/assets/660/651322.pdf>. Accessed February 20, 2019.
- Woodall, W.H., Zhao, M.J., Paynabar, K., Sparks, R. and Wilson, J.D. (2017) An overview and perspective on social network monitoring. *IIE Transactions*, **49**(3), 354–365.
- Xu, K.S. and Hero, A.O. (2014) Dynamic stochastic block models for time-evolving social networks. *IEEE Journal of Selected Topics in Signal Processing*, **8**(4), 552–562.
- Zou, N. and Li, J. (2017) Modeling and change detection of dynamic network data by a network state space model. *IIE Transactions*, **49**(1), 45–57.