IISE Transactions on Healthcare Systems Engineering

ISSN: 2472-5579 (Print) 2472-5587 (Online) Journal homepage: https://www.tandfonline.com/loi/uhse21

Network modeling and Internet of things for smart and connected health systems—a case study for smart heart health monitoring and management

Hui Yang, Chen Kan, Alexander Krall & Daniel Finke

To cite this article: Hui Yang, Chen Kan, Alexander Krall & Daniel Finke (2020) Network modeling and Internet of things for smart and connected health systems—a case study for smart heart health monitoring and management, IISE Transactions on Healthcare Systems Engineering, 10:3, 159-171, DOI: 10.1080/24725579.2020.1741738

To link to this article: https://doi.org/10.1080/24725579.2020.1741738

IISE TRANSACTIONS ON HEALTHCARE SYSTEMS ENGINEERING
2020, VOL. 10, NO. 3, 159–171
https://doi.org/10.1080/24725579.2020.1741738

Network modeling and Internet of things for smart and connected health systems—a case study for smart heart health monitoring and management

Hui Yang[a], Chen Kan[a], Alexander Krall[a], and Daniel Finke[b]

[a]Harold and Inge Marcus Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Park, PA, USA; [b]Applied Research Lab, The Pennsylvania State University, University Park, PA, USA

ABSTRACT

Heart disease is a leading cause of death in the US. Recent advances in the Internet of Things (IoT) provide a great opportunity to realize smart and connected health systems through IoT monitoring and sensor-based data analytics of cardiac disorders. However, big data arising from the large-scale IoT system pose a significant challenge for efficient and effective sensory information processing and decision making. Very little has been done to glean pertinent information about the disease-altered cardiac activity in the context of large-scale IoT network. In this study, we propose a parallel computing framework for multi-level network modeling and monitoring of cardiac dynamics to realize the potential of IoT-enabled smart health management. Specifically, dissimilarities among cardiac signals are firstly characterized among heartbeats for an individual patient, as well as among representative heartbeats for different patients. Then, a stochastic learning approach is developed to optimize the embedding of cardiac signals into a beat-to-beat network model, as well as a patient-to-patient network model. Further, we develop a parallel computing algorithm to improve the computational efficiency. Finally, a statistical process monitoring scheme is designed to harness network features for real-time monitoring and anomaly detection of cardiac activities. Experimental results show the proposed methodology has strong potential to realize a smart and interconnected system for cardiac health management in the context of large-scale IoT network.

KEYWORDS

Cardiac monitoring; electrocardiogram signals; Internet of Things; network modeling; parallel computing; pattern matching

1. Introduction

Heart disease is the No. 1 cause of death in the US (Benjamin et al., 2019). It is estimated that approximately 610,000 deaths each year are attributed to heart disease. Considering the costs of medications, services, and lost productivity, the economic burden amounts to over $200 billion (Mozaffarian et al., 2017) per year. Smart and connected health hinges on real-time monitoring of physiological signals and timely identification of the onset of disease patterns. For example, the effects of ischemia on heart muscle cells are reversible if the incident is detected and treated early (De Luca et al., 2004). When the episode of ischemia is prolonged, cardiac cells will be damaged and become infarcted, thereby triggering heart attacks. It is estimated that a 30-minute delay will increase the risk of one-year mortality by 7.5% (De Luca et al., 2004). There is an urgent need to develop new sensor-based analytical methods and tools for real-time cardiac monitoring and disease pattern recognition, thereby promoting smart health management.

Recent advances in the Internet of Things (IoT) herald a new era of cardiac health management. The IoT connects a multitude of "things" in an Internet-like infrastructure, including wearable sensors, computing units, medical devices, people (e.g., patients, nurses, and cardiologists), and digital infrastructures such as databases and fog and cloud computing paradigms (Kan et al., 2015). Examples include Bluesky from IBM/Pfizer, ApexPro from GE healthcare, and eIAC from Philips, to name a few. There are also more than 325,000 mobile health applications available in Android and iOS platforms, exceeding 3.7 billion downloads in 2017 (Rao et al., 2018).

As opposed to traditional hospital-centered care, the IoT infrastructure and sensing systems provide a great opportunity to realize patient-centered and interconnected health management. As shown in Figure 1, both inpatients and outpatients are continuously monitored and data are transmitted to the cloud server via wireless networks. Analytical algorithms are running on the server to analyze the collected data and identify abnormal cardiac patterns. Cardiologists are able to access patients' data, review analytical results, and communicate with patients and other cardiologists anytime and anywhere. If a patient's condition is identified as high risk, medical intervention can be delivered in a timely manner.

Advanced sensing in the large-scale IoT network leads to the proliferation of large amounts of data, which provides a wealth of opportunity to improve the "smartness" of cardiac

CONTACT: Hui Yang, huy25@psu.edu Harold and Inge Marcus Department of Industrial and Manufacturing Engineering, The Pennsylvania State University, University Park, PA, USA.

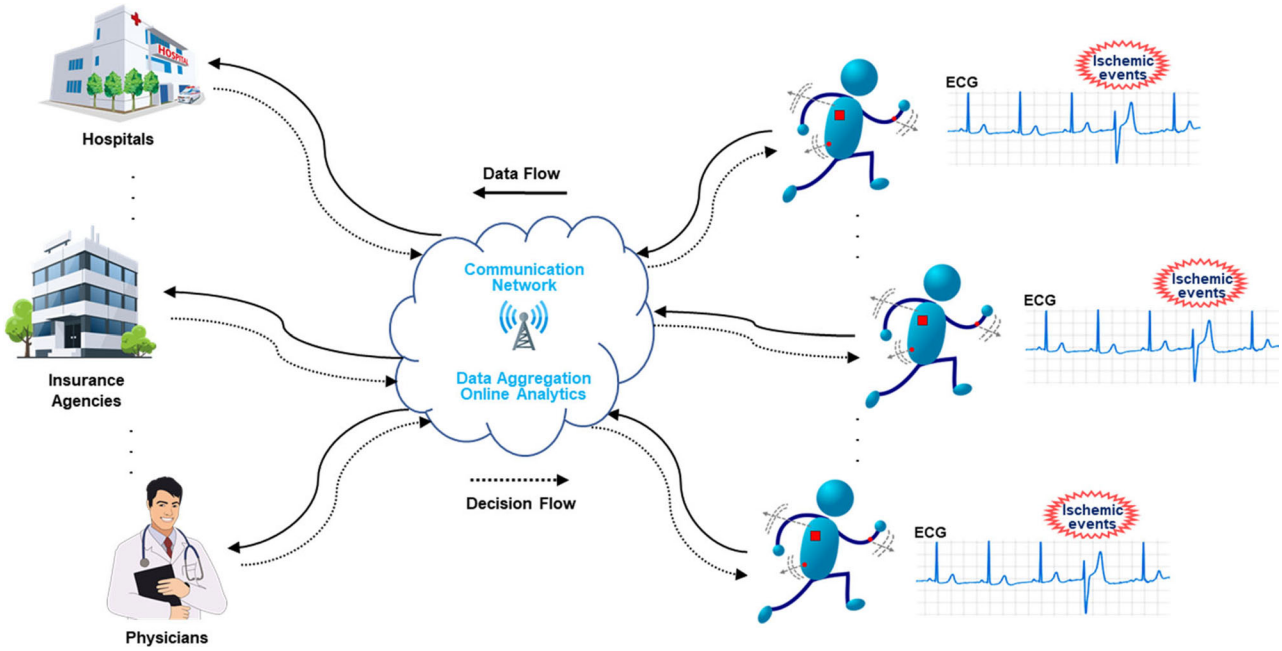

Figure 1. An illustration of the IoT-enabled cardiac health management. Data from patients enter the cloud and are accessible by health institutions. These institutions can then make decisions, which are transmitted back to the patient.

health management on the level of an individual patient and/or population. First, IoT-enabled sensors can collect long-term monitoring data of ECG signals from an individual patient, which contain rich information about the patient's cardiac condition over time. As opposed to episodic care, characterizing the longitudinal cardiac variations of a patient enables the delivery of customized treatment plans and personalized medicine (Verma & Sood, 2018). Second, it is the nature of IoT to embody tens of thousands of patients, which empowers the cross-sectional study on the patient population, rather than focusing on the individual or a small group of patients.

As the IoT system can easily expand by adding new patients or incorporating small-scale sensor networks, the physical network can be quickly scaled up with a variety of one-to-one and one-to-many communications. When a large number of patients are included, the physical network will generate hundreds of thousands of ECG signals as well as other heterogeneous types of health information. This volume of data poses a significant challenge on network modeling and analytics in the cyberspace to support decisions and/or insights that are fed back to the physical world due to the large data volume and high-level network complexity and population heterogeneity.

Also, big data generated in the IoT pose significant challenges for information processing and decision making. As the IoT is relatively new, existing methodologies fall short of addressing the Internet-like structures to extract pertinent information from big data about the disease dynamics in the beat-to-beat (B2B) network for an individual patient, as well as in the patient-to-patient (P2P) network, also called "digital twin." Realizing the full potential of IoT-enabled smart health management calls for the development of new analytical methods for sensor information extraction, large-scale network modeling, and data-driven decision making in the context of large-scale IoT networks.

In this study, we propose a parallel computing framework for multi-level network modeling and monitoring of the B2B and P2P variations of cardiac activities, which leverages big data in the IoT to realize the full potential of IoT-enabled smart health management. Specifically, we first characterize and measure the dissimilarities of cardiac signals among heartbeats for an individual patient, as well as among representative heartbeats for different patients. This, in turn, helps to characterize (1) temporal B2B variations of an individual patient, or (2) cross-sectional variations among a large population of patients. Then, cardiac signals are optimally embedded as nodes in a high-dimensional network, where node-to-node distances preserve the dissimilarity measures between cardiac signals. The distance between network nodes preserves the dissimilarity of cardiac signals. The network structure and topology contain pertinent information about patients' cardiac conditions and thereby provide a new means of cardiac monitoring. Finally, a statistical process monitoring approach is developed based on the attributes of network nodes to monitor disease-altered cardiac dynamics. However, the computation time will be prohibitive for network modeling when a large number of patients are involved. Hence, a parallel computing scheme is further developed to scale up the algorithm and efficiently harness the computing power of multiple processors simultaneously. In the present investigation, our specific contributions are summarized as follows:

(1) We develop a large-scale network model to handle the big data generated in IoT-enabled cardiac monitoring. As opposed to traditional small-scale episodic analysis, this article focuses on the characterization and

modeling of the variations of cardiac signals among a large population of patients, as well as B2B variations over time for an individual patient.

(2) Because large-scale network modeling is computationally expensive due to the population size and long sampling periods, we further develop parallel-computing algorithms of stochastic learning for efficient sensor-based modeling and optimization of large-scale B2B and P2P networks.

(3) In addition, we leverage the optimized network structure and node attributes as a new way to develop a statistical process monitoring scheme for real-time monitoring and change detection of diseased-altered cardiac dynamics in IoT-enabled B2B and P2P networks.

The remainder of this article is organized as follows: Section 2 reviews related works in the literature. Section 3 provides methodological details about the proposed large-scale network model. Section 4 includes materials and experimental design used in this investigation. Section 5 presents the experimental results from both simulation and real-world studies. Section 6 provides discussion and conclusions drawn from this study.

2. Research background

2.1. IoT health monitoring and management

First coined by Ashton in 1999, the IoT has become a game-changer technology with widespread applications in industries (Yang et al., 2019). In particular, the IoT is driving the paradigm shift toward the next generation of smart and interconnected healthcare systems by deploying a large number of networked sensors and computing devices for data collection, information processing, communication, and decision making. This provides a great opportunity to improve the quality of healthcare delivery, increase the availability and accessibility to health care, optimize the management of medical resources, and reduce the costs of healthcare systems. Note that patients are of greater importance vis-à-vis other networked components in the IoT system. Therefore, smart and interconnected IoT health management is the first priority. IoT-enabled solutions are increasingly developed for patient-centered health monitoring.

For example, Al-Taee *et al.* (Al-Taee et al., 2017) proposed a mobile health platform that used humanoid robots for the treatment of diabetes in children. The robots were linked to a web-centric disease management hub via the IoT to analyze patients' data and provide reminders, warning messages, and health advice to patients and caregivers. Yang *et al.* (Yang et al., 2014) developed IoT devices (e.g., wearable sensors and medicine packaging) to improve home-based health monitoring. Alerts would be triggered if abnormities of vital signs as well as medication noncompliance were identified. In addition, Verma and Sood (Verma & Sood, 2018) designed a fog computing layer to improve the efficiency of IoT-based health monitoring. Patients' condition could be quickly determined as safe or unsafe in the fog layer; data were only sent to the cloud for further analysis if an unsafe state was identified.

The IoT also fuels increasing interests to integrate low-power wearable devices and edge/cloud computing to improve the efficiency and effectiveness of cardiac monitoring. For example, a low-power and secure IoT platform was developed to collect electrocardiogram (ECG) signals for the prediction of ventricular arrhythmia (Yasin et al., 2017). Algorithmic enhancement was incorporated in the hardware design to reduce memory and clock speed for energy efficiency. Also, a logic locking technique was implemented to protect the security-critical components in the hardware. An onboard algorithm was designed to assess the quality of collected ECG signals via IoT devices (Satija et al., 2017). Only those signals with acceptable quality were transmitted to the cloud to reduce the consumption of battery power and improve the efficiency of resource utilization. Majumder *et al.* (Majumder et al., 2019) developed a wearable IoT device to collect ECG signals from the patient. Features such as heart rate and RR intervals were extracted from the ECG to detect abnormal cardiac patterns. Islam *et al.* offer a review of IoT-enabled applications in healthcare domains (Islam et al., 2015).

Won *et al.* proposed a method to enhance the predictive performance of network-based machine learning models (Won et al., 2019). A convex semi-infinite programming (SIP) approach is used to support vector machines to deal with the complex requirement of $\updownarrow_0$ norm regularization. The SIP is solved by iteratively finding solutions to a restricted master problem. At each subsequent step, new constraints are added to the master problem, which guarantees convergence to optimality. Tran *et al.* utilized the spectral graph theory to detect transient changes in complex neurological systems (Tran et al., 2019). They formulated a new γ_k statistic based on the spectral content of the inferred graph. When applied to seizure detection methods based on EEG data, the method was able to achieve a true positive rate of 40%. Tucker *et al.* proposed to use low-cost, noninvasive, off-the-shelf hardware sensors (e.g., Microsoft Kinect) to capture human gait data, and then develop data mining methods to enable remote patient-physician assessment and predict the Parkinson's disease conditions (Tucker et al., 2015). Zou *et al.* developed an empirical Bayes transfer learning (ebTL) model that incorporated transfer learning and sparse learning to estimate the parameter posterior distribution and quantify the prediction uncertainty. Further, the ebTL model is evaluated by using features extracted from speech signals to predict the severity level of Parkinson's disease (Zou & Huang, 2018). Cheng *et al.* used the wearable accelerometer sensors to monitor the daily activities and movements of elderly patients in assisted living facilities, and then developed multi-scale network models to characterize the dementia conditions (Cheng & Yang, 2019). In addition, Hill *et al.* studied real-time, remote physiological monitoring of astronauts during field science tasks in the NASA's BASALT Mars Analog field (Hill et al., 2018). This research work is critical to promoting health and safety

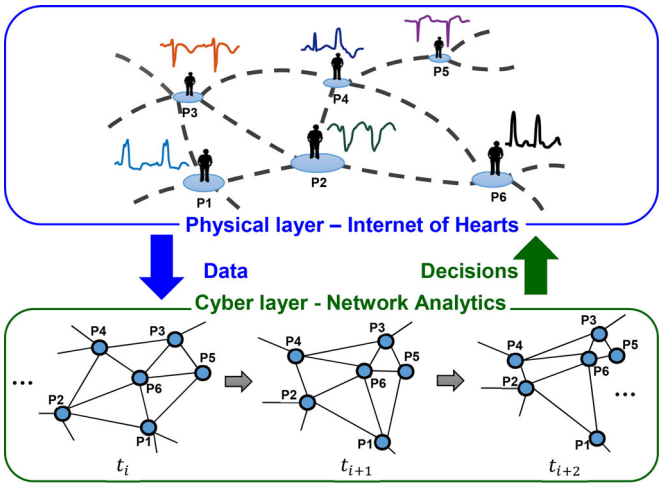

Figure 2. The physical IoT network feeds data to build the network models, and network analytics feed decisions and/or insights back to the physical world.

of workers in an extreme and harsh environment. However, the sampling frequency is about 1 Hz for physiological monitoring, which is lower than cardiac monitoring that typically requires >250 Hz for better delineation of minute details in the ECG waveforms.

Although these prior works demonstrated the promise to integrate wearable sensors and data mining with health monitoring, most of them focus on the physiological monitoring of vital signs in the low frequency (e.g., 1 Hz) or the prediction of neurological conditions with sensor data such as EEG signals, Kinect movement data, speech signals, and accelerometer data. Most previous works are less concerned about efficient and effective data analytics for real-time cardiac monitoring in the large-scale IoT context. Note that our prior studies also proposed a new framework of Internet of Hearts (IoH) and discussed the potential to develop the IoT technology for smart cardiac monitoring and management (Kan et al., 2016; Kan & Yang, 2017b). First, we developed a new Mobile and E-Network Smart Health (MESH) technology specific to the heart to advance the cardiac mHealth with IoT sensing, stochastic modeling, and network analytics (Kan & Yang, 2017b). The MESH development includes wearable ECG sensors, cloud database design, and four analytical models as follows: (1) real-time visualization of ECG time series and feature analysis (Bukkapatnam et al., 2008; Chen & Yang, 2012b); (2) real-time visualization of 3 D VCG trajectory and recurrence analysis (Chen & Yang, 2012a; Yang, 2011; Yang et al., 2012); (3) optimal model-based representation of ECG signals (Liu et al., 2014; Liu & Yang, 2013); (4) spatiotemporal signal processing for disease pattern recognition (Yang et al., 2012; 2013). A mobile application named MESH CARE is also developed in the world's most widely used IOS mobile operating system (i.e., compatible with iPhone, iPad, and iTouch devices). Further, we proposed to develop the MESH into a new generation of IoH system but found that large volumes of data are generated from continuous monitoring of patients (Kan et al., 2016). There is an urgent need to improve the computational efficiency of the IoH for data processing and analytics. Hence, we proposed a map-reduce scheme to implement the parallel computing for ECG analytics, and benchmarked the computational speeds with simulation data (Kan et al., 2016). These preliminary studies laid strong foundations for the present investigation, but did not specifically consider the development of statistical monitoring methods for the IoH systems. In addition, very little has been done to leverage sensing data for network modeling and analysis of disease-altered P2P and B2B cardiac dynamics. Realizing the full potential of IoT-enabled smart health management calls for the development of new methodologies for statistical monitoring of P2P and B2B network dynamics in the large-scale IoT context.

2.2. Disease pattern recognition in cardiac activities

The key in cardiac monitoring is to detect the changes in cardiac activities and identify disease patterns in the early stage. In the literature, a variety of algorithms were designed to extract useful features and patterns from ECG signals for the detection of cardiac diseases. For example, Elmberg *et al.* (Elmberg et al., 2016) measured QRS prolongation in 12-lead ECGs to quantify the severity of ischemia. Meo *et al.* (Meo et al., 2013) characterized the variability of f-wave amplitude from 12-lead ECGs to predict the catheter ablation outcome of atrial fibrillation. Perlman *et al.* (Perlman et al., 2016) extracted features from QRS complex in 12-lead ECGs and developed a classification tree scheme for the identification of supraventricular tachycardia. Our previous works have also investigated nonlinear dynamics algorithms to recognize disease-altered ECG patterns for the detection and identification of myocardial infarction, atrial fibrillation, bundle branch block, and other cardiac diseases (Chen & Yang, 2013; Yang, 2011; Yang et al., 2012). For example, customized wavelet functions were designed to extract fiducial patterns of ECG signals for the detection of atrial fibrillations (Yang et al., 2007). Also, heterogeneous recurrence analysis was proposed to characterize heart rate variability from ECG signals for the identification of dynamic transitions and obstructive sleep apnea (Cheng et al., 2016; Chen & Yang, 2014, 2015). A self-organizing network was developed to characterize pattern dissimilarities among QRS complexes in the ECGs and then recognize abnormal patterns induced by the left bundle branch block (LBBB) (Yang & Leonelli, 2016).

However, most of the existing algorithms are limited in their capacity to scale up and handle large-scale IoH sensing data. When large amounts of data are collected from the IoH system, algorithmic complexity increases due to the inclusion of a large number of individuals, and algorithmic performance is likely to be hampered by the increasing volume of data. This, in turn, will limit the ability of traditional algorithms for cross-sectional study of a large patient population or detect temporal B2B variations in the long-term cardiac monitoring. Realizing the full potential of IoT sensing data calls for the development of new analytical methodologies to efficiently and effectively handle large amounts of patient data and extract pertinent information from the data about patients' cardiac abnormalities.

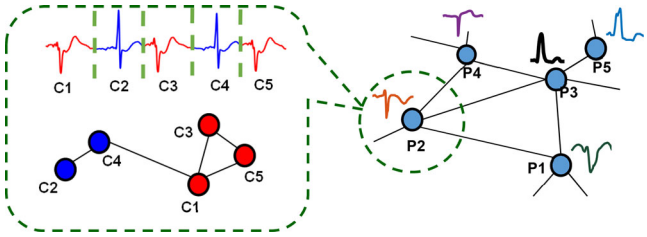

Figure 3. The proposed network modeling: a B2B network (left) and a P2P network (right).

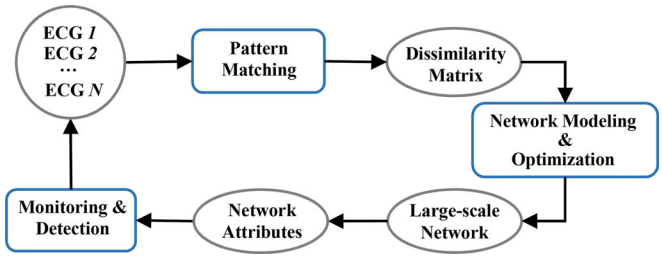

Figure 4. The flowchart for IoH network modeling and analytics. For a P2P network, ECG 1, ECG 2, … , ECG N are ensemble beats from different patients. For a B2B network, these are successive beats of an individual patient.

3. Network modeling and analytics

3.1. Network modeling

In this article, we propose to leverage the Internet-like connection of patients in the physical network to build network models in the cyberspace for data-driven modeling and information processing. As shown in Figure 2, patients in the physical network are represented by nodes in the virtual network and node attributes are patients' ECG signals. As such, variations of patients' conditions can be reflected by the structure and topology of the network model in cyberspace. Notably, the virtual network overcomes practical limitations in the physical network and facilitates the handling of IoT sensing data for information processing and decision making. Analytics in the virtual network (or "digital twin") feed insights or decisions back to the physical world for the timely delivery of medical interventions.

Moreover, there exist two types of stochastic variations in cardiac activities: (1) B2B variations among heartbeats of an individual patient: ECG signals for the same human subject show temporal dynamics, which are reflected as morphological variability of ECG heartbeats over time; (2) P2P variations among different patients: ECG signals vary from one human subject to another in the population. There are random variations even if the subject condition remains unchanged or does not experience cardiac disorders. Cardiac disorders bring more significant changes in the ECG signals. Thus, characterizing and modeling B2B and P2P variations of ECG signals provide a great opportunity for the early detection of cardiac events as well as the differentiation of cardiac disorders.

Therefore, two types of network models are constructed in the IoH systems—a P2P network and a B2B network (see Figure 3). In the P2P network, an ensemble of ECG cycles from a patient is embedded as a network node, which represents the aggregated information of this patient. Nodes in the B2B network represent successive ECG cycles for a specific patient. However, it is not uncommon for the IoT system to include a large population of patients, as well as collect long-term B2B data from an individual patient. The number of patients and data volume pose significant challenges for the construction and optimization of the network model. Realizing the potential of IoT-enabled cardiac health management depends, to a great extent, on network modeling and efficient information-processing capabilities. Existing ECG pattern recognition approaches are limited in their ability to scale up and handle large amounts of data. To address this challenge, parallel computing algorithms and new network modeling schemes in the IoT context are presented.

As shown in Figure 4, pattern matching is first performed among the ECG signals (detailed in Section 3.2). For a B2B network, the ECGs are successive cycles collected from an individual patient. Further, an ensemble cycle can be generated for one patient that represents the typical cardiac pattern for this specific patient. Many ensemble cycles will be collected from different patients and be used to construct a P2P network. After matching the ECG wave patterns, a dissimilarity matrix is obtained and used to build the network model. Each patient is represented as a network node and the distance between nodes will preserve the dissimilarity of corresponding ECGs (Section 3.3). The network structure reveals important information pertinent to patients' cardiac dynamics. To this end, network attributes can be used for the monitoring of patients' conditions and the detection of abnormal cardiac patterns (Section 3.4).

3.2. Pattern matching of sensor signals

As previously mentioned, characterizing and quantifying the morphological variability among ECG signals provide a great opportunity for the identification of disease-altered cardiac electrical activities. It is worth noting that there are various ways to measure the pattern dissimilarity, such as Euclidean distance, correlation, mutual information, and/or wavelet transformation (Liu & Yang, 2017; Chen and Yang, 2016; Zhou et al., 2006). Furthermore, these measures can be computed in the time, frequency, state space, and wavelet domains. Note that the correlation coefficient quantifies linear interdependence between two signals and is limited to effectively capture nonlinear relationships. Mutual information is utilized to determine the similarity of the joint distribution of two signals with respect to the product of marginal distributions. Given two ECG signals $\vec{v}_1$ and $\vec{v}_2$, the mutual information is computed as:

$$MI(\vec{v}_1,\ \vec{v}_2) = \sum_{\vec{v}_1}\sum_{\vec{v}_2} Pr(\vec{v}_1,\ \vec{v}_2)\log\left(\frac{Pr(\vec{v}_1,\ \vec{v}_2)}{Pr(\vec{v}_1)Pr(\vec{v}_2)}\right) \quad (1)$$

Table 1 summarizes the mathematical notations used in this section. In addition, Euclidean distances can be measured from two signals to quantify their morphological dissimilarity. However, signals can be misaligned. Measuring distances without the alignment of signals will produce misleading results. In our previous study, a spatiotemporal warping

Table 1. Mathematical notations used in this section.

Symbol	Meaning
$\vec{v}_k$	*ECG signal k*
D	*The distance matrix*
G	*The search matrix for dynamic programming*
V_k	*The length of signal k*
W	*A warping path, a sequence of index pairs*
$\delta_{k,m}$	*Normalized distance (dissimilarity) between signals k and m*
Δ	*The dissimilarity matrix*
N	*The dimensionality of* Δ
B	*Matrix decomposition of* Δ
H	*Centering matrix*

approach was developed to optimally align the signals (Yang et al., 2013), $\min_W \left[\sum_{(t_i,t_j)\in W} \left|\left|\vec{v}_1(t_i) - \vec{v}_2(t_j)\right|\right|\right]$, where W is the warping path. To find W, a distance matrix D is computed with each element $D(i,j) = \left|\left|\vec{v}_1(i) - \vec{v}_2(j)\right|\right|$, $i = 1,2,\ldots, V_1$ and $j = 1,2,\ldots, V_2$. The warping path W is recursively computed using a dynamic programing approach:

$$G(i,j) = D(i,j) + \min\begin{cases} G(i,\ j\text{–}1) \\ G(i-1,\ j-1) \\ G(i-1,\ j) \end{cases} \tag{2}$$

The warping path W is a sequence of index pairs that cconnect the $(1,1)$ and (V_1,V_2) indices of the G search matrix. The dissimilarity between $\vec{v}_1$ and $\vec{v}_2$ can then be calculated as the normalized distance: $\delta_{1,\ 2} = \frac{G(V_1,V_2)}{V_1+V_2}$. The spatiotemporal warping approach not only enables the measure of pattern dissimilarity among 1-lead ECG, but also multi-lead ECGs such as the gold-standard 12-lead ECGs and 3-lead Frank VCGs (Yang et al., 2013).

3.3. Network modeling and optimization

After the dissimilarity measure, a $N \times N$ matrix Δ is obtained, where N is the number of cardiac signals. Each element δ_{ij} represents the dissimilarity between ECG signals of two patients i and j. However, the dissimilarity matrix itself is difficult to use in terms of predictor variables or features for predictive modeling and health monitoring. Hence, we propose to exploit the dissimilarity matrix by the network modeling, which optimally represents the signals as network nodes based on the dissimilarities among them. Note that a widely used approach for network embedding and modeling is the classic multidimensional scaling (MDS) (Cheung and So, 2005; Yang, 2008). Let $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ denote the location of i^{th} and j^{th} nodes in the network; the objective function can be defined as:

$$f = \sum_{i,j} \left(\left|\left|\boldsymbol{x}_i - \boldsymbol{x}_j\right|\right| - \delta_{ij}\right)^2 \tag{3}$$

The MDS minimizes Eq. (3) by decomposing a matrix $B = -\frac{1}{2}H\Delta^{(2)}H$, where H $= I - N^{-1}11^T$ is a centering matrix. This is done by the singular value decomposition: $B = V\Lambda V^T = V\sqrt{\Lambda}\sqrt{\Lambda}V^T$, where optimal locations $\boldsymbol{x}_i$'s can be obtained as: $X = V\sqrt{\Lambda}$. Here, V is a matrix of eigenvectors and Λ is a diagonal matrix of eigenvalues. Note that the distance between nodes $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ after network optimization preserves the dissimilarity δ_{ij} between ECG signals of two patients i and j.

However, most network modeling approaches are subject to high levels of computational complexity for the embedding of a large volume of data (Kan et al., 2018). To address this issue, we propose a stochastic network embedding approach in this research. The cost function of Eq. (3) is reformulated as:

$$f = \frac{1}{2}\left[\sum_i \sum_{j\neq i} \left(\left|\left|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right|\right| - \delta_{ij}\right)^2 + \vartheta_t \left|\left|\boldsymbol{x}_j^t - \boldsymbol{x}_j^{t-1}\right|\right|_2^2\right] \times \Psi\left(\left|\left|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right|\right|,\ \lambda\right) \tag{4}$$

where $\Psi(\cdot)$ is a step function:

$$\Psi\left(\left|\left|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right|\right|,\ \lambda\right) = \begin{cases} 1 \ \ if \ \ \left|\left|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right|\right| \leq \ \lambda \\ 0 \ \ if \ \ \left|\left|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right|\right| > \ \lambda \end{cases} \tag{5}$$

and λ is the neighborhood radius. Notably, the cost function (Eq. (4)) consists of two regularization terms: a spatial term $(f_{SP} = \sum_i \sum_{j\neq i} \left(\left|\left|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right|\right| - \delta_{ij}\right)^2)$ to preserve the dissimilarity and a temporal term $(f_{TP} = \vartheta_t \left|\left|\boldsymbol{x}_j^t - \boldsymbol{x}_j^{t-1}\right|\right|_2^2)$ to regularize the update of $\boldsymbol{x}_j$ in successive iterations.

As opposed to learning the network of ECG signals in one batch, the minimization of Eq. (4) can be stochastically updated (Demartines and Herault, 1997). At each iteration t, a node i is randomly picked and its location $\boldsymbol{x}_i$ is fixed. Then, locations of all other nodes $\boldsymbol{x}_j (j \neq i)$ are updated according to the following rules:

$$\begin{cases} \boldsymbol{x}_j^{t+1} = \boldsymbol{x}_j^t - \eta_t\left(\nabla f_{SP} + \vartheta_t\left(\boldsymbol{x}_j^t - \boldsymbol{x}_j^{t-1}\right)\right) \ \ if \ \ \left|\left|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right|\right| \leq \ \lambda_t \\ \boldsymbol{x}_j^{t+1} = \boldsymbol{x}_j^t \ \ if \ \ \left|\left|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right|\right| > \ \lambda_t \end{cases} \tag{6}$$

where

$$\nabla f_{SP} = \frac{\left|\left|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right|\right| - \delta_{ij}}{\left|\left|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right|\right|}\ \left(\boldsymbol{x}_j^t - \boldsymbol{x}_i^t\right). \tag{7}$$

The learning rate η_t, the radius of neighborhood λ_t, and the regularization parameter ϑ_t can be modeled as monotonically decreasing functions (e.g., $\eta_t = \eta_0\left(\frac{\eta_T}{\eta_0}\right)^{\frac{t}{T}}$, where η_0 and η_T specify the initial and terminal values of the learning rate and their ratio controls the decaying speed). This approach iteratively optimizes the locations $\boldsymbol{x}_j$ to minimize Eq. (4) until the maximum number of learning epoch is reached.

Because only one node is considered at each iteration, the computational complexity of stochastic learning is much lower compared to MDS. However, it is still expensive when

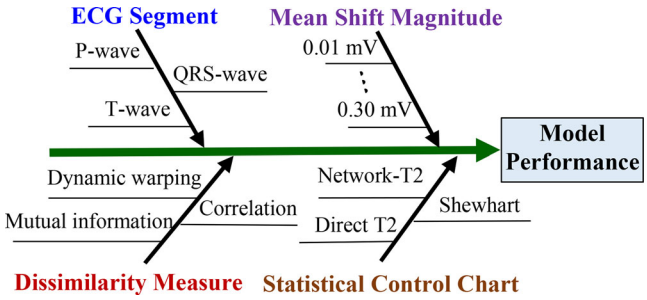

Figure 5. Experimental design for the simulation study.

a large number of nodes are included in the network. As opposed to serial-computing algorithms like the MDS, we have further developed a parallel computing scheme to harness multiple computing resources collaboratively and improve computational efficiency. At each iteration, we consider not only one ECG signal, but multiple signals as a mini-batch Θ. Sub-gradients with respect to signals in Θ are accumulated and the average gradient is computed as:

$$\nabla f_{\Theta}\left(\boldsymbol{x}_j^t\right) = \frac{1}{|\Theta|}\sum_{i\in\Theta}\nabla f_i\left(\boldsymbol{x}_j^t\right) \tag{8}$$

Here, $\nabla f_i\left(\boldsymbol{x}_j^t\right)$ represents fixing $\boldsymbol{x}_j^t$ and calculating the gradient of f with respect to $\boldsymbol{x}_j^t$ and $|\Theta|$ is the size of the mini-batch Θ. As such, we can update $\boldsymbol{x}_j^t$ (if $\left\|\boldsymbol{x}_i^t - \boldsymbol{x}_j^t\right\| \leq \lambda_t$) as:

$$\boldsymbol{x}_j^t = \boldsymbol{x}_j^t - \eta_t\left(\nabla f_{\Theta}\left(\boldsymbol{x}_j^t\right) + \vartheta_t\left(\boldsymbol{x}_j^t - \boldsymbol{x}_j^{t-1}\right)\right) \tag{9}$$

To facilitate the parallel computing, we divide Θ into multiple subsets; i.e., $\Theta = \Theta_1 \cup \Theta_2 \cup, \ldots, \cup \Theta_S$. Thus, Eq. (7) becomes:

$$\nabla f_{\Theta}\left(\boldsymbol{x}_j^t\right) = \frac{1}{|\Theta|}\left[\sum_{i\in\Theta_1}\nabla f_i(\boldsymbol{x}_j^t) + \sum_{i\in\Theta_2}\nabla f_i(\boldsymbol{x}_j^t) + \ldots + \sum_{i\in\Theta_S}\nabla f_i(\boldsymbol{x}_j^t)\right] \tag{10}$$

In this way, parallel computing can be readily implemented: each subset Θ_s is assigned to an individual processor and results from many processors are then combined using Eq. (10) to return the average gradient. Notably, the temporal penalty $\vartheta_t\left\|\boldsymbol{x}_j^t - \boldsymbol{x}_j^{t-1}\right\|_2^2$ in Eq. (4) is similar to the conservative penalty. It was shown that adding such a penalty on mini-batch learning significantly benefits the computational efficiency and performance (Li et al., 2014). The algorithm is summarized as follows.

Algorithm 1 Parallel computing for stochastic network modeling

Input: dissimilarity matrix Δ (δ_{ij}) obtained from warping
Output: coordinates of nodes $\boldsymbol{x}_i, i \in \{1, 2, \ldots, N\}$ in the high-dimensional network

1: initialize $\boldsymbol{x}_i$, learning rate $\{\eta_t\}_{t=1}^T$, neighborhood radius $\{\lambda_t\}_{t=1}^T$, and regularization parameter $\{\vartheta_t\}_{t=1}^T$, T – number of epoch
2: randomly partition data into M mini-batches
3: start a parallel computing pool with S processors
4: **for** t =1, 2, …, T **do**
5: randomly choose a mini-batch Θ_t
6: partition Θ_t into subsets: $\Theta_{t,\ 1}, \Theta_{t,\ 2}, \ldots, \Theta_{t,S}$
7: **for** $n = 1$ to S **do** {**in parallel**}
8: processor n gets partition $\Theta_{t,\ n}$
9: solve the sub-problem $\nabla f_{i\in\Theta_{t,\ n}}\left(\boldsymbol{x}_j^t\right)$ on $\Theta_{t,n}$
10: **end for**
11: average over the batch size $|\Theta_t|$ and update $\boldsymbol{x}_j^t$
12: **end for**
13: close the parallel computing pool

3.4. Network monitoring of health conditions

Once the network is optimized to use the distance between nodes $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ to preserve the dissimilarity δ_{ij} between ECG signals of two patients i and j, the next step is to exploit the network structure and node attributes (e.g., coordinates) for health monitoring and change detection of the onset of cardiac disorders. In the literature, various modeling approaches (such as particle filtering and neural networks) are available to quantitatively associate network features to patients' cardiac conditions. Practitioners can select the most suitable approach based on the complexity of data and the requirements of processing speed. In this study, we introduce a computationally efficient method—statistical control charts. The most widely used chart is the Shewhart chart, which is designed to monitor the mean or variance of process data. However, the Shewhart chart tends to be limited in the ability to handle multivariate data. In this study, we design and develop the network-based Hotelling T^2 chart to monitor the variations in multivariate data of network features:

$$T^2(i) = (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^T\ \boldsymbol{S}^{-1}(\boldsymbol{x}_i - \bar{\boldsymbol{x}}) \tag{11}$$

Here, $\bar{\boldsymbol{x}}$ is the averaged coordinates of network nodes and $\boldsymbol{S}$ is the covariance matrix. However, $\boldsymbol{S}$ may be singular if the dimensionality of network coordinates is high and if there are redundance and interdependence in the system. As such, the inverse of $\boldsymbol{S}$ cannot be calculated. This issue can be addressed by eigen transformation (see more details in Kan and Yang (2017a)) and the statistic of the network- T^2 chart can be calculated as:

$$T^2(i) = (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^T\ \boldsymbol{S}^{-1}(\boldsymbol{x}_i - \bar{\boldsymbol{x}}) = \sum_{r=1}^{Q}\frac{\mathbf{Z}(i,r)^2}{\lambda_r^2} \tag{12}$$

The upper control limit of the network-T^2 control chart is estimated as:

$$UCL_{T^2} = \frac{q(N+1)(N-1)}{N^2 - Nq}F_{\alpha,q,M-q} \tag{13}$$

where q is the dimensionality of $\boldsymbol{x}_i$, $F_{\alpha,q,M-q}$ is the upper $100\alpha\%$ of critical points of the F distribution with q and $N-q$ degrees of freedoms. The significance level α is set as 0.05.

4. Materials and experimental design

In this article, both simulation experiments and real-world case studies are used to evaluate and validate the developed

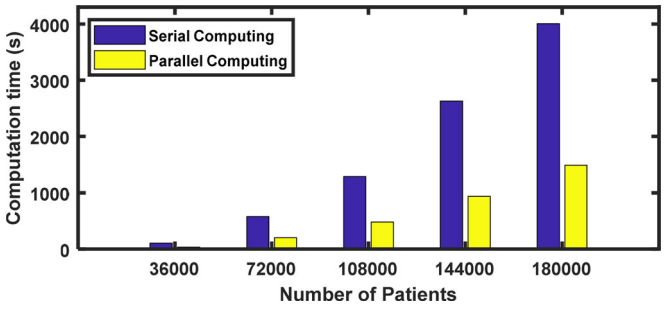

Figure 6. Computational time using parallel and serial computing for stochastic network embedding and modeling.

methodology. In the simulation study, the design of experiments consists of four factor groups, as shown in the cause and effect diagram in Figure 5.

(1) *ECG segments*: Note that cardiac diseases alter the patterns of ECG signals and cause the morphology variations that are reflected in various ECG segments; e.g., P, QRS, T waves (Elmberg et al., 2016; Meo et al., 2013; Perlman et al., 2016). For example, pulmonary hypertension will result in the right atrial enlargement, which is reflected as enlarged P waves. Abnormities in ventricular conductions (e.g., myocardial infarctions) will mainly lead to morphological variations in QRS and T waves. The lengths of P, QRS, T segments are often variable, and the period of ventricular contraction is the longest for most cases. Those segments with smaller lengths pose bigger challenges for change detection. Hence, we will test the proposed network monitoring methods for various ECG segments.

(2) *Shift magnitudes within an ECG segment*: Morphological variability also reveals the disease severity. The onset of heart disease often induces a small shift in ECG morphology, whereas the shift magnitude is larger when the disease has lasted for a longer time and progressed to a late stage (Elmberg et al., 2016; Meo et al., 2013; Perlman et al., 2016). Small magnitude shifts pose greater challenges for the change detection. Hence, there is an urgent need to implement and test the proposed monitoring method on different levels of magnitude shifts in ECG segments, which is conducive to early detection of heart diseases before the progression to a late stage.

(3) *Dissimilarity measures*: Characterizing the dissimilarities among ECG signals is a prerequisite to build the network model and increase the detection power. The signal dissimilarity can be characterized and quantified from various perspectives, such as morphological dissimilarity, signal correlation, and the mutual dependence of the amount of information held in both signals. In the experiments of this study, we will also evaluate and compare the impacts of different dissimilarity measures, namely signal warping, correlation, and mutual information, on the performance of change detection.

(4) *Statistical control charts*: Early detection of the onset of cardiac disorders depends to a great extent on the measure of signal dissimilarities, as well as an effective monitoring scheme. Therefore, we implement the network-T^2 chart and benchmark charts (i.e., direct T^2 chart and Shewhart chart) under different experimental scenarios (e.g., various ECG segments, different levels of magnitude shifts, and three dissimilarity measures) to evaluate the performance of proposed analytical methods and algorithms.

In the real-world case study, ECG signals collected from continuous monitoring (the PhysioNet Sudden Cardiac Death and long-term ST databases (Goldberger et al., 2000; Jager et al., 2003; Moody et al., 2001)) are used to evaluate the performance of the B2B network. The ECGs were digitized at 250 Hz sampling rate with a 12-bit resolution over a range of ± 10 mV. Also, beat-by-beat annotations are provided by cardiologists, labelling each cycle as normal or premature ventricular contraction. We first preprocess the ECG with a Fast Fourier Transform band-pass filter ($1 \sim 120$ Hz), which removes the artifact, baseline wandering, as well as high-frequency noises. ECG cycles (i.e., the P-QRS-T wave for each heartbeat) are then extracted for the next steps to measure the pattern dissimilarity and network modeling.

5. Experimental results

5.1. Simulation experiments

We first evaluate the computational efficiency of the proposed parallel computing scheme over the traditional serial computing algorithm for handling a large volume of ECG cycles from a population of patients. Second, the proposed network monitoring is implemented for the detection of disease-altered cardiac activities. Following the experimental design in Section 5, the performance of the proposed large-scale network model is evaluated by varying the levels of four factors: (1) ECG segments; (2) shift magnitudes within an ECG segment; (3) dissimilarity measures; and (4) control charts.

As shown in Figure 6, the parallel computing algorithm compartmentalizes the updating rules of Eqs. (8–10) to multiple processors with mini-batches, but the serial-computing algorithm only applies the update rule of Eq. (6) using a single processor without the consideration of the mini-batch. When the number of patients increases from 36,000 to 180,000, the computational time of serial algorithms increases much more significantly than the parallel algorithms. The gap between the blue bar (serial) and yellow bar (parallel) becomes significantly larger with the increase of the number of patients in the network. When 36,000 patients are included, the gap is very small. When the size of the network reaches 180,000, the difference between the two approaches increases to >2500 s. Therefore, the parallel computing scheme shows much greater efficiency than the serial scheme and is well suited for use in the large-scale IoT context.

Further, we follow the experimental design in Section 5 to evaluate the performance of the proposed network method for change detection of signal variations in cardiac

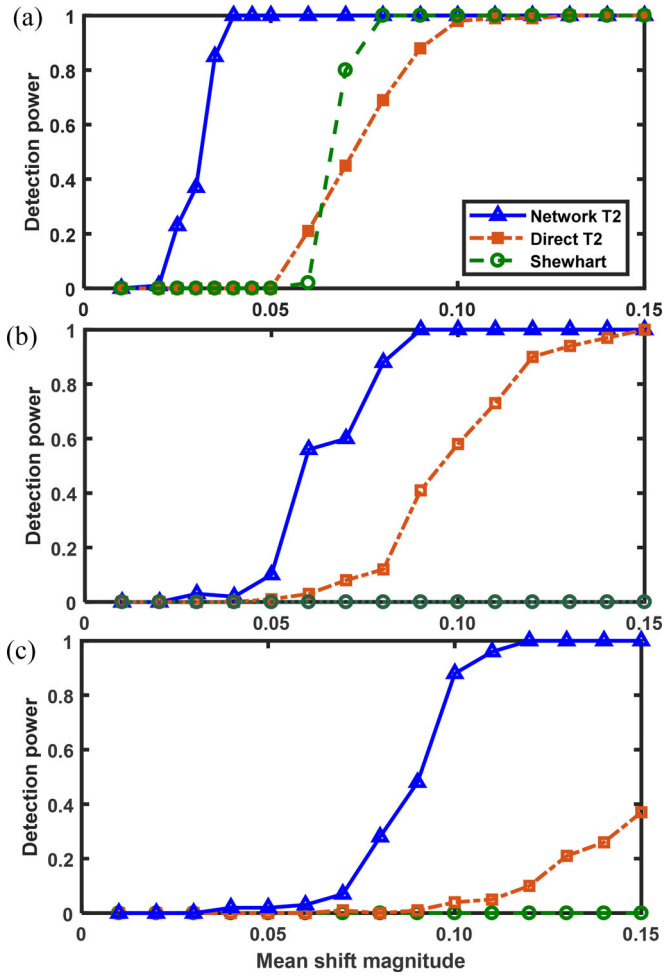

Figure 7. Performance comparison between three types of control charts under different shift magnitudes for (a) shifted ST segments, (b) shifted P segments, and (c) shifted QRS segments.

dynamics. First, we simulate disease-induced morphological variability by shifting three segments of ECG signals (i.e., ST, P, and QRS). Here, an elevated ST segment is used to simulate the acute myocardial infarction. Also, a larger P wave corresponds to the potential atrial enlargement. When the patient is associated with massive pericardial effusion, the voltage of QRS will be lower, which is reflected as a reduced QRS complex. The disease severity is controlled by the level of shift magnitude. In other words, a small shift indicates that the disease is in the early stage, whereas a larger shift is associated with late-stage disease conditions and calls for immediate medical interventions.

For each scenario, we simulate 150 ECG cycles based on real-world ECG signals from healthy and diseased subjects. The first 50 cycles are normal (i.e., in-control) signals, whereas the remaining 100 cycles contain the mean shifts with different magnitudes of ST, QRS, and P segments that simulate various disease-altered signals (i.e., out-of-control). The dissimilarities among ECG cycles are measured by the spatiotemporal warping method introduced in Section 4.1. Here, the proposed network-T^2 control chart is implemented on the 150 ECG cycles in each scenario. The detection power is calculated as the proportion of out-of-controls detected in the 100 ECG cycles with disease-induced morphological variability. Furthermore, we

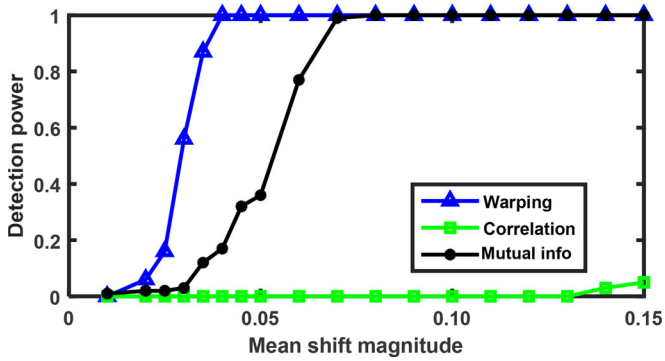

Figure 8. Performance comparison of three dissimilarity measures (warping, correlation, and mutual information) under different shift magnitudes.

evaluate the performance of the proposed network-T^2 chart with two benchmark charts: (1) a T^2 chart based on the original ECG signals (denoted as "Direct T^2" in Figure 7); and (2) a Shewhart chart monitoring the mean change of ECG signals (denoted as "Shewhart" in Figure 7).

The proposed network-T^2 chart performs better than the Shewhart chart and the direct T^2 chart in all of the experimental scenarios. As shown in Figure 7, the detection power of the network-T^2 chart (blue triangles) quickly reaches 100% at very small shift magnitude (0.03 mV) for shifted ST segments (see Figure 7a). When shifted segments are shorter, the network-T^2 chart is also sensitive to small changes and the detection power reaches 100% at a shift magnitude of 0.08 mV and 0.12 mV for the P (Figure 7b) and QRS segments (Figure 7c), respectively. The detection power of the direct T^2 chart (orange squares) increases slower with the increase of shift magnitude. When the length of the shifted segment is short (shifted QRS segments in Figure 7c), the detection power of the direct T^2 chart is only 40% when the mean shift magnitude is 0.15 mV and it reaches 100% when the shift magnitude is over 0.22 mV. Notably, the detection power of the Shewhart chart is comparable to the direct T^2 chart when the shift magnitude is large (e.g., ST segments in Figure 7a). However, few out-of-controls can be detected by the Shewhart chart when shift magnitudes are small (see Figure 7b and c).

In addition, as shown in Figure 8, we have also compared the use of three dissimilarity measures for the matching of ECG patterns. Here, we used the experimental scenario of disease-induced changes in ST segments, and then the dissimilarity matrices obtained from three measures are used to construct the network-T^2 chart and compute the detection power. Notably, the dissimilarity measure with signal warping performs better than the mutual information. The network-T^2 chart with signal warping reaches the detection power of 100% when the shift magnitude is only 0.04. However, the detection power reaches 100% for the dissimilarity measure with mutual information when the shift magnitude is over 0.07. The control chart based on the correlation cannot yield a good performance when the shift magnitude is small from 0.01 to 0.15.

5.2. Real-world case study

In addition to the simulation study, we have collected real ECG signals from a population of 2000 atrial fibrillation

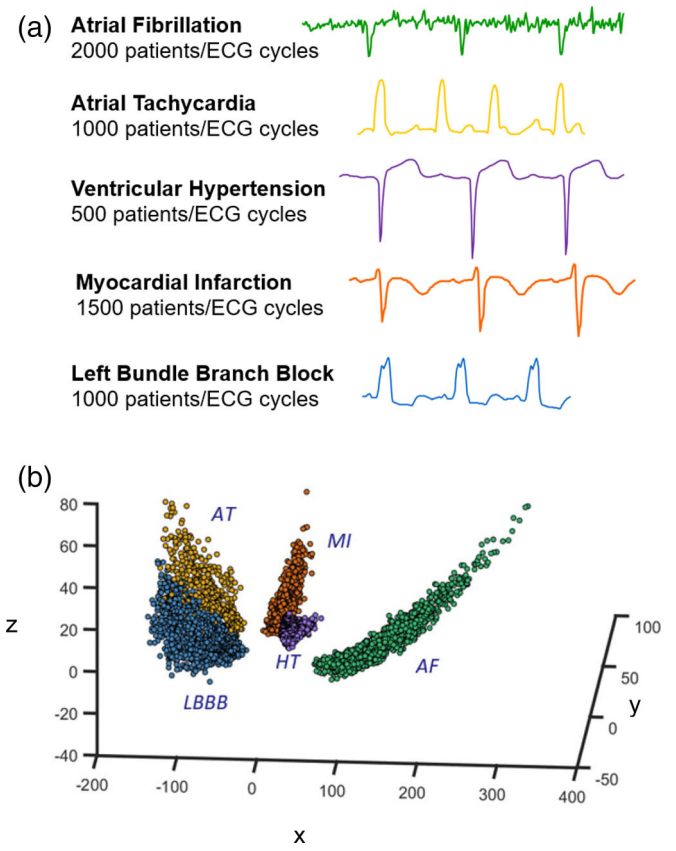

Figure 9. (a) Real-world ECG signals from a population of 2000 atrial fibrillation (AF), 1000 atrial tachycardia (AT), 500 ventricular hypertropy (HT), 1500 myocardial infarction (MI), and 1000 left bundle branch block (LBBB) patients collected from the PhysioNet. (b) The spatial distribution of nodes in the constructed P2P network.

(AF), 1000 atrial tachycardia (AT), 500 ventricular hypertropy (HT), 1500 myocardial infarction (MI), and 1000 left bundle branch block (LBBB) patients from the PhysioNet. Figure 9a shows the representative ECG signals from each group of patients. Then, we built the P2P network model by representing each patient as a network node, and node attributes are patients' ECG signals. Figure 9b shows that variations of patients' conditions are reflected by the network structure and topology. It is worth noting that the variations of AF are significantly higher than other groups, thus forming a cluster of widespread nodes in the network. However, ventricular hypertension has smaller variations in signal patterns and hence forms a small cluster of nodes in the network. The proposed network representation provides an effective approach for the physicians to visualize a variety of patient groups and pinpoint the location of a new patient for clinical decision support. For example, if ECG signals from a new patient are embedded as a node in the cluster of myocardial infarctions, then clinicians can retrieve those patients with similar symptoms in the cluster, make the diagnosis of cardiac conditions for this new patient, and take proactive treatment plans such as Holter monitoring and medications to prevent the events of life-threatening heart attacks.

In addition, we collected 24-hour monitoring data of ECG signals from the long-term ST database, available in PhysioNet (Goldberger et al., 2000; Jager et al., 2003; Moody et al., 2001). As shown in Figure 10, the morphology of ECG cycles varies from time to time for a specific patient: (1) normal cycles (with annotation N) show similar patterns but there are variations among them; and (2) the premature ventricular contraction cycle (with annotation V) appears once in a while with a different morphology from the normal cycle. The premature ventricular contraction is caused by extra heartbeats that begin in one ventricle. As a result, such extra beats will cause disruptions to the normal heart rhythm, making people feel a fluttering heart beat (or a skipped one). The premature ventricular contraction can further lead to heart attack, congestive heart failure, and diseases of heart valves. It is important to detect the onset of the premature ventricular contraction and monitor its temporal variations for the early detection of cardiac disorders.

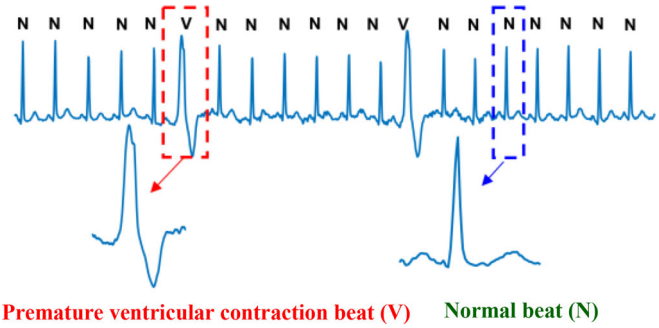

Figure 10. Real-world ECG signals containing normal beats (N) and premature ventricular contraction beats (V).

Hence, we evaluate the proposed methodology to build the beat-to-beat (B2B) network to represent the variations of B2B ECG cycles for a specific patient. As shown in Figure 11, blue nodes represent normal beats and red nodes correspond to beats with premature ventricular contractions. Note that the network structure is dynamically changing when more and more beats are increasingly available over time and embedded in the network. For example, the first 15 beats contain three premature ventricular contractions, which are located apart from the main group of normal beats (see the red dots in Figure 11a). Figure 11b shows the embedded network with 25 beats. Note that five beats associated with premature ventricular contraction are located closer, and their cluster is away from the normal cluster (blue nodes). When there are a total of 50 beats (see Figure 11c), 13 abnormal beats form a distinct group located away from the cluster of normal beats.

The network topology of 1000 successive ECG beats is shown in Figure 11d. The network structure and spatial locations of network nodes provide an effective means for human experts to visualize and compare the ECG cycles over time for a specific patient and then personalize the intervention plan. When a large number of heartbeats are located closer to the disease cluster (red nodes), this indicates that the patient has a high level of risk and needs immediate medical attention. More importantly, the network topology provides a great opportunity for cardiologists to visualize the temporal change of a patient's cardiac condition. For example, if the network visualization shows $1 \sim 2$ red nodes occasionally, this may be due to random effects in the function of the patient's ventricle. However, immediate

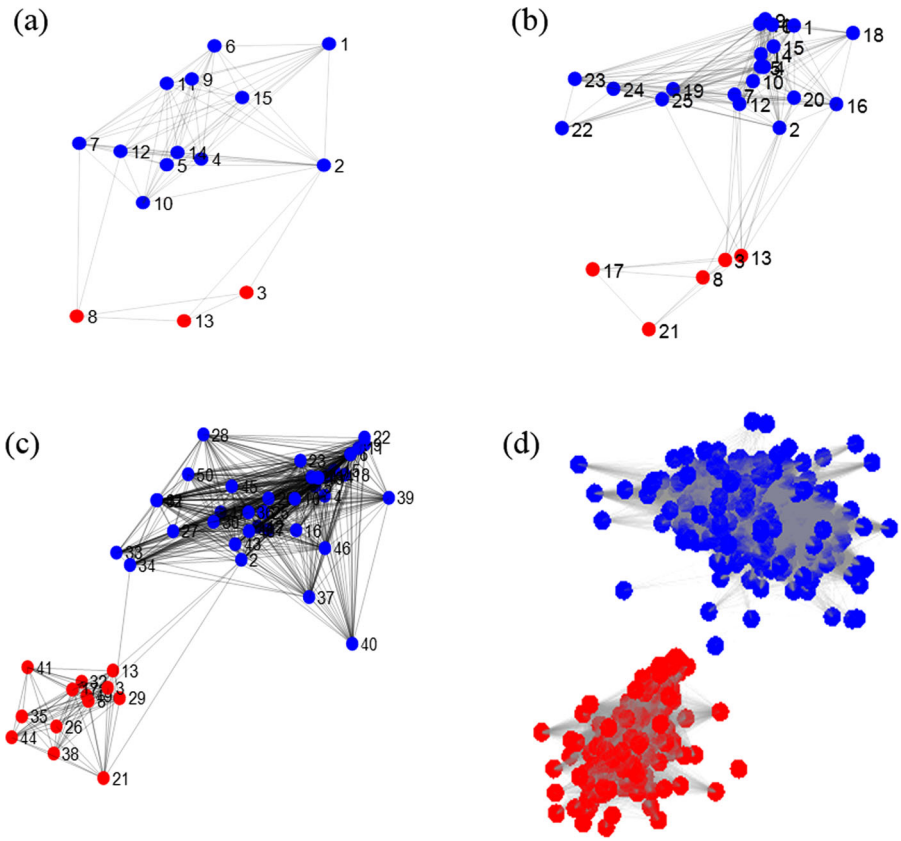

Figure 11. The constructed B2B networks for (a) 15 beats, (b) 25 beats, (c) 50 beats, and (d) 1000 beats from the real-world ECG signals.

medical attention is needed when the number of red nodes is increased rapidly and more frequently. This may be due to a significant deterioration in the ventricle function, which is often associated with fatal cardiac disorders.

6. Discussion and conclusions

With rapid advances of IoT technologies, there are increasingly new opportunities available to implement IoT-enabled information processing and decision support for smart and interconnected health management. In this article, we present parallel computing algorithms for multi-level network modeling and analytics of P2P and B2B variations of ECG signals in big data generated from the IoT system. We boost the computational efficiency by developing new parallel computing approaches that distribute the computing efforts of network modeling and optimization into multiple processors, instead of traditional serial-computing approaches executed in a single processor. Finally, a new statistical process monitoring scheme is developed to realize the full potentials of network models (including the features of network topology, clustering structures, and node attributes) for cardiac monitoring and anomaly detection.

Network modeling and visualization provide an effective means for physicians and cardiologists to identify communities (i.e., node clusters) of disease groups, investigate disease mechanisms in each community, and pinpoint the health condition of a new patient in the communities of the network. The P2P network reveals the cross-sectional information of the patient population. The structure of the network community helps to visualize and differentiate various types of cardiac conditions. The nodes in each community are corresponding to the patients associated with each cardiac condition. The distribution of nodes in each community shows the level of variations that exist among patients in that community. The spatial location of a node facilitates the comparison of the signal patterns of this patient with many other patients in the network and then characterizes the health condition. On the other hand, the personalized B2B network reveals progressive variations of cardiac conditions for a specific patient over time. If nodes in the B2B network are departing away from the cluster or community of normal heartbeats, more attention should be paid to this patient. Our simulation and real-world experiments show that the proposed network modeling and analysis framework has strong potential to help leverage the newly available IoT infrastructure to build the new generation of cardiac monitoring systems for smart health management, as well as gain a better understanding of disease mechanisms in various populations and identify high-risk patients.

Funding

This work is support in part by the NSF I/UCRC Center for Healthcare Organization Transformation (CHOT), NSF I/UCRC award 1624727. Any opinions, findings, or conclusions found in this article are those of the authors and do not necessarily reflect the views of the sponsors.

References

Al-Taee, M., Al-Nuaimy, W., Muhsin, Z., & Al-Ataby, A. (2017). Robot assistant in management of diabetes in children based on the Internet of Things. *IEEE Internet of Things Journal*, *4*(2), 437–445. doi:10.1109/JIOT.2016.2623767

Benjamin, E. J., Muntner, P., Alonso, A., Bittencourt, M. S., Callaway, C. W., Carson, A. P., Chamberlain, A. M., Chang, A. R., Cheng, S., Das, S. R., Delling, F. N., Djousse, L., Elkind, M. S. V., Ferguson, J. F., Fornage, M., Jordan, L. C., Khan, S. S., Kissela, B. M., Knutson, K. L., Kwan, T. W., Lackland, D. T., & Virani, S. S. (2019). Heart Disease and Stroke Statistics—2019 Update: A Report From the American Heart Association. *Circulation*, *139*(10), e56–e528.

Bukkapatnam, S., Komanduri, R., Yang, H., Rao, P., Lih, W.-C., Malshe, M., Raff, L. M., Benjamin, B., & Rockley, M. (2008). Classification of atrial fibrillation (AF) episodes from sparse electrocardiogram (ECG) datasets. *Journal of El*ectrocardiology, *41*(4), 292–299. doi:10.1016/j.jelectrocard.2008.01.004

Chen, Y., & Yang, H. (2013). Wavelet packet analysis of disease-altered recurrence dynamics in the long-term spatiotemporal vectorcardiogram (VCG) signals. In 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 2595–2598).

Chen, Y., & Yang, H. (2015, Aug. 24-28). *Heterogeneous recurrence T2 charts for monitoring and control of nonlinear dynamic processes* [Paper presentation]. Proc. 11th Ann. IEEE Int. Conf. Autom. Sci. Eng. (CASE 2015), Gothenburg, Sweden.

Chen, Y., & Yang, H. (2012a). Multiscale recurrence analysis of long-term nonlinear and nonstationary time series. *Chaos, Solitons & Fractals*, *45*(7), 978–987. doi:10.1016/j.chaos.2012.03.013

Chen, Y., & Yang, H. (2012b). Self-organized neural network for the quality control of 12-lead ECG signals. *Physiological* Measurement, *33*(9), 1399–1418. doi:10.1088/0967-3334/33/9/1399

Chen, Y., & Yang, H. (2014). Heterogeneous recurrence monitoring and control of nonlinear stochastic processes. *Chaos:* An Interdisciplinary Journal of Nonlinear Science, *24*(1), 013138. doi:10.1063/1.4869306

Chen, Y., & Yang, H. (2016). A Novel Information-Theoretic Approach for Variable Clustering and Predictive Modeling Using Dirichlet Process Mixtures. *Scientific Reports*, *6*(1), 38913. doi:10.1038/srep38913

Cheng, C., & Yang, H. (2019). Multi-scale graph modeling and analysis of locomotion dynamics towards sensor-based dementia assessment. *IISE Transactions on Healthcare Systems Engineering*, *9*(1), 95–102. doi:10.1080/24725579.2018.1530315

Cheng, C., Kan, C., & Yang, H. (2016). Heterogeneous recurrence analysis of heartbeat dynamics for the identification of sleep apnea events. *Computers in Biology and Medicine*, *75*, 10–18. doi:10.1016/j.compbiomed.2016.05.006

Cheung, K. W., & So, H. C. (2005). A multidimensional scaling framework for mobile location using time-of-arrival measurements. *Ieee Transactions on Signal Processing*, *53*(2), 460–470. doi:10.1109/TSP.2004.840721

De Luca, G., Suryapranata, H., Ottervanger, J., & Antman, E. (2004). Time delay to treatment and mortality in primary angioplasty for acute myocardial infarction: Every minute of delay counts. *Circulation*, *109*(10), 1223–1225. doi:10.1161/01.CIR.0000121424.76486.20

Demartines, P., & Herault, J. (1997). Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets. *IEEE Transactions on Neural Networks*, *8*(1), 148–154. doi:10.1109/72.554199

Elmberg, V., Almer, J., Pahlm, O., Wagner, G., Engblom, H., & Ringborn, M. (2016). A 12-lead ECG-method for quantifying ischemia-induced QRS prolongation to estimate the severity of the acute myocardial event. *Journal of Electrocardiology*, *49*(3), 272–277. doi:10.1016/j.jelectrocard.2016.02.001

Goldberger, A., Amaral, L., & Glass, L. (2000). PhysioBank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation*, *23*, e215–e220.

Hill, J. R., Caldwell, B. S., Downs, M., Miller, M. J., & Lim, D. S. S. (2018). Remote physiological monitoring in a Mars Analog field setting. *IISE Transactions on Healthcare Systems Engineering*, *8*(3), 227–236. doi:10.1080/24725579.2018.1501624

Islam, S., Kwak, D., Kabir, H., Hossain, M., & Kwak, K. (2015). The internet of things for health care: A comprehensive survey. *IEEE Access*, *3*, 678–708. doi:10.1109/ACCESS.2015.2437951

Jager, F., Taddei, A., Moody, G., Emdin, M., Antolic, G., & Dorn, R. (2003). Long-term ST database: A reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia. *Medical & Biological Engineering & Computing*, *41*(2), 172–183. doi:10.1007/BF02344885

Kan, C., & Yang, H. (2017a). Dynamic network monitoring and control of in-situ image profiles from ultraprecision machining and biomanufacturing processes. *Quality and Reliability Engineering International*, *33*(8), 2003–2022. doi:10.1002/qre.2163

Kan, C., & Yang, H. (2017b). Internet of hearts - large-scale stochastic network modeling and analysis of cardiac electrical signals. In J. Li, N. Kong and X. Xie, (Eds.), *Stochastic Modeling and Analytics in Healthcare Systems* (pp. 211–251). World Scientific.

Kan, C., Chen, Y., Leonelli, F., & Yang, H. (2015). Mobile sensing and network analytics for realizing smart automated systems towards health internet of things. In Automation Science and Engineering (CASE), 2015 IEEE International Conference On, Gothenburg, Aug. 24-28 (pp. 1072–1077).

Kan, C., Yang, H., & Kumara, S. (2018). Parallel computing and network analytics for fast Industrial Internet-of-Things (IIoT) machine information processing and condition monitoring. *Journal of Manufacturing Systems*, *46*, 282–293. doi:10.1016/j.jmsy.2018.01.010

Kan, C., Leonelli, F. M., & Yang, H. (2016). Map reduce for optimizing a large-scale dynamic network – internet of hearts. In Proceedings of 2016 IEEE Engineering in Medicine and Biology Society Conference (EMBC), Orlando, FL (pp. 2962–2965).

Li, M., Zhang, T., Chen, Y., & Smola, A. (2014). Efficient mini-batch training for stochastic optimization. In KDD '14 Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York (pp. 661–670). doi:10.1145/2623330.2623612

Liu, G., & Yang, H. (2013). Multiscale Adaptive Basis Function Modeling of Spatiotemporal Vectorcardiogram Signals. *Biomedical and Health Informatics, IEEE Journal Of*, *17*(2), 484–492.

Liu, G., & Yang, H. (2017). Self-organizing network for group variable selection and predictive modeling. *Annals of Operations Research*, *33*(9), 1399–1418. doi:10.1007/s10479-017-2442-2

Liu, G., Kan, C., Chen, Y., & Yang, H. (2014). Model-driven parametric monitoring of high-dimensional nonlinear functional profiles. In Automation Science and Engineering (CASE), 2014 IEEE International Conference On, Taipei (pp. 722–727).

Majumder, A., ElSaadany, Y., Young, R., & Ucci, D. (2019). An Energy Efficient Wearable Smart IoT System to Predict Cardiac Arrest. *Advances in Human-Computer Interaction*, *2019*, 1–21. doi:10.1155/2019/1507465

Meo, M., Zarzoso, V., Meste, O., Latcu, D., & Saoudi, N. (2013). Spatial variability of the 12-Lead surface ECG as a tool for noninvasive prediction of catheter ablation outcome in persistent atrial fibrillation. *IEEE Transactions on Biomedical Engineering*, *60*(1), 20–27. doi:10.1109/TBME.2012.2220639

Moody, G. B., Mark, R. G., & Goldberger, A. L. (2001). PhysioNet: A web-based resource for the study of physiologic signals. *IEEE Engineering in Medicine and Biology Magazine*, *20*(3), 70–75. doi:10.1109/51.932728

Mozaffarian, D., Benjamin, E., & Go, A. (2017). Heart disease and stroke statistics - 2016 update: A report from the American Heart Association. *Circulation*, *136*(18), e38–e360

Perlman, O., Katz, A., Amit, G., & Zigel, Y. (2016). Supraventricular tachycardia classification in the 12-Lead ECG using atrial waves detection and a clinically based tree scheme. *IEEE Journal of*

Biomedical and Health Informatics, *20*(6), 1513–1520. doi:10.1109/JBHI.2015.2478076

Rao, G. K. L., Mokhtar, N., & Iskandar, Y. H. P. (2018). Managing orthodontic needs through mobile apps. *Journal of Hospital Management and Health Policy*, *2*, 12–12. doi:10.21037/jhmhp.2018.01.07

Satija, U., Ramkumar, B., & Manikandan, M. (2017). Real-Time Signal Quality-Aware ECG Telemetry System for IoT-Based Health Care Monitoring. *IEEE Internet of Things Journal*, *4*(3), 815–823. doi:10.1109/JIOT.2017.2670022

Tran, H. M., Bukkapatnam, S. T. S., & Garg, M. (2019). Detecting changes in transient complex systems via dynamic network inference. *IISE Transactions*, *51*(3), 337–353. doi:10.1080/24725854.2018.1491075

Tucker, C., Han, Y., Black Nembhard, H., Lee, W.-C., Lewis, M., Sterling, N., & Huang, X. (2015). A data mining methodology for predicting early stage Parkinson's disease using non-invasive, high-dimensional gait sensor data. *IIE Transactions on Healthcare Systems Engineering*, *5*(4), 238–254. No10/02, doi:10.1080/19488300.2015.1095256

Verma, P., & Sood, S. (2018). Fog Assisted-IoT Enabled Patient HealthMonitoring in Smart Homes. *IEEE Internet of Things Journal*, *5*(3), 1789–1796. doi:10.1109/JIOT.2018.2803201

Won, D., Manzour, H., & Chaovalitwongse, W. (2019). Convex Optimization for Group Feature Selection in Networked Data. *INFORMS Journal on Computing*, *32*(1), 1–198. doi:10.1287/ijoc.2018.0868

Yang, H., & Leonelli, F. (2016). Self-organizing visualization and pattern matching of vectorcardiographic QRS waveforms. *Computers* in Biology and Medicine., *79*, 1–9. doi:10.1016/j.compbiomed.2016.09.020

Yang, H. (2011). Multiscale Recurrence Quantification Analysis of Spatial Cardiac Vectorcardiogram Signals. *Biomedical Engineering, IEEE Transactions On*, *58*(2), 339–347.

Yang, L. (2008). Alignment of overlapping locally scaled patches for multidimensional scaling and dimensionality reduction. *Pattern Analysis and Machine Intelligence, IEEE Transactions On*, *30*(3), 438–450.

Yang, H., Bukkapatnam, S. T. S., & Komanduri, R. (2012). Spatio-temporal representation of cardiac vectorcardiogram (VCG) signals. *Biomedical Engineering Online*, *11*(1), 16. doi:10.1186/1475-925X-11-16

Yang, H., Bukkapatnam, S., & Komanduri, R. (2007). Nonlinear adaptive wavelet analysis of electrocardiogram signals. *Physical Review E*, *76*(2), 026214. doi:10.1103/PhysRevE.76.026214

Yang, H., Bukkapatnam, S. T. S., Le, T., & Komanduri, R. (2012). Identification of myocardial infarction (MI) using spatio-temporal heart dynamics. *Medical* Engineering & Physics., *34*(4), 485–497. doi:10.1016/j.medengphy.2011.08.009

Yang, H., Kan, C., Liu, G., & Chen, Y. (2013). Spatiotemporal Differentiation of Myocardial Infarctions. *IEEE Transactions on Automation Science and Engineering*, *10*(4), 938–947. doi:10.1109/TASE.2013.2263497

Yang, H., Kumara, S., Bukkapatnam, S., & Tsung, F. (2019). The Internet of Things for Smart Manufacturing: A Review. *IISE Transactions*, *51*(11), 1190–1216. doi:10.1080/24725854.2018.1555383

Yang, G., Xie, L., Mantysalo, M., Zhou, X., Pang, Z., Xu, L. D., Kao-Walter, S., Chen, Q., & Zheng, L.-R. (2014). A health-IoT platform based on the integration of intelligent packaging, unobtrusive bio-sensor, and intelligent medicine box. *IEEE Transactions on Industrial Informatics*, *10*(4), 2180–2191. doi:10.1109/TII.2014.2307795

Yasin, M., Tekeste, T., Saleh, H., Mohammad, B., Sinanoglu, O., & Ismail, M. (2017). Ultra-low power, secure IoT platform for predicting cardiovascular diseases. *IEEE Transactions on Circuits and Systems I: Regular Papers*, *64*(9), 2624–2637. doi:10.1109/TCSI.2017.2694968

Zhou, S., Sun, B., & Shi, J. (2006). An SPC monitoring system for cycle-based waveform signals using Haar transform. *Automation Science and Engineering, IEEE Transactions On*, *3*(1), 60–72.

Zou, N., & Huang, X. (2018). Empirical Bayes Transfer Learning for Uncertainty Characterization in Predicting Parkinson's Disease Severity. *IISE Transactions on Healthcare Systems Engineering*, *8*(3), 209–219. doi:10.1080/24725579.2018.1496495