Map Reduce for Optimizing a Large-scale Dynamic Network – the Internet of Hearts

Chen Kan, Fabio M. Leonelli and Hui Yang*, *Member, IEEE*

Abstract—**Rapid advancements of sensing and mobile technology provide an unprecedented opportunity to empower smart and connected healthcare. Realizing the full potential of connected care depends, however, to a great extent on the capability of data analytics. Our previous study proposed a next-generation mobile health system, namely, the Internet of Heart (IoH). The IoH embeds patients into a dynamic network, where the distance between network nodes is determined by the dissimilarity of patients' conditions. Dynamics of the network reveal the change of clinical status of patients. However, it poses a great challenge for real-time recognition of disease patterns when a considerably large number of patients are involved in the IoH. In this present investigation, we develop a novel scheme to optimize the network in a parallel, distributed manner, thereby improving the efficiency of computation. First, a stochastic gradient descent approach is designed to embed patients with similar conditions into a local network. Second, local networks are optimally pieced together to obtain a global network. As opposed to directly embed all patients into one network, the proposed scheme distributes the network optimization into multiple processors for parallel computing. This, in turn, enables the IoH to handle large amount of patients and timely recognize disease patterns in the early stage. Experimental results demonstrated the effectiveness of the proposed scheme, e.g., it achieves 80-fold faster than conventional algorithms for optimizing a network with 20000 patients. The developed scheme is effective and efficient for realizing smart connected healthcare in large-scale IoH contexts.**

I. Introduction

In the past decade, rapid advancements of sensing and mobile technology have fueled increasing interests in mobile health (mHealth). The prevalence of wearable bio-sensors and portable medical devices makes it possible and affordable to remotely monitor patients' conditions and provide timely feedback. Furthermore, mHealth attempts to create connected care defined as a national wide data gathering and exchange, as well as an efficient communication between patients and caregivers. As opposed to traditional isolated care, highly-connected care ecosystem consists of doctors, patients, devices, databases and other entities. Caregivers are able to gain rapid access to patients' complete information for early diagnosis and timely medical intervention. Meanwhile, it motivates individual patients to collect data themselves and become active participants in their own care. The smart connected healthcare provides a data-rich environment for efficient medical decision making, enables personalized patient-centered care, and diminishes care disparities.

This work is supported in part by the National Science Foundation (CMMI-1617148, CMMI-1619648, and IOS-1146882).

Chen Kan and Hui Yang* with the Complex Systems Monitoring, Modeling and Analysis Laboratory, The Pennsylvania State University, University Park, PA 16802 USA (*corresponding author: huy25@psu.edu).

Fabio M. Leonelli a staff cardiologist and practice physician with Cardiac Electrophysiology Lab at the James A. Haley Veterans' Hospital, Tampa, FL 33620 USA.

Realizing the full potential of connected care depends, nevertheless, to a great extent on the capability of data analytics. It may be noted that cardiac diseases are the leading cause of death worldwide. In 2012, nearly 30% of global deaths (17.5 million) were due to cardiac diseases. Optimal management and treatment of cardiac diseases call for early identification and timely delivery of life-saving therapies. However, existing products and services of cardiac mHealth are limited in their ability to automatically detect cardiac diseases in the early stage. Most of them focus on one-lead electrocardiogram (ECG) and simple characteristics, e.g., heart rate, which lacks the diagnostic power to identify complex cardiac disorders. Some of them transmit acquired ECG signals to cardiologists for manually analysis. Despite the overwhelming workload, early signs of cardiac diseases are difficult to be uncovered by only measuring amplitude and duration of the ECG cycles.

Our previous research has proposed a new cardiac mHealth system, namely, the Internet of Hearts (IoH) [1] to assist the diagnosis of cardiac diseases and promote new levels of communication and collaboration between patients and cardiologists. The IoH incorporates wearable sensing, mobile technology and cloud computing for continuous cardiac monitoring and disease pattern recognition. Patients are embedded into a network, where the distance between network nodes preserves the dissimilarity of patients' conditions. As such, the change of a patient's cardiac conditions can be revealed by network dynamics. The IoH enables (*i*) continuous monitoring of multi-channel cardiac signals; (*ii*) real-time management and compact representation of multi-sensor signals; (*iii*) big data analytics in large-scale IoT contexts. These components are integrated to realize patient-centered care and smart management of cardiac health.

Notably, continuous monitoring of an individual patient generates large volumes of data when performed in hours, days, months and years. Such big data not only provides a wealth of opportunities to promote patient-centered, personalized care, but also poses significant challenges for analytics and management. There lacks enabling tools to quickly extract information pertinent to the underlying cardiac disorders and provide timely feedbacks. Furthermore, the number of patients in the IoH is growing. The goal of IoH is to embody patients all across the world into a network to empower smart telehealth and preventive medicine. It is extremely expensive to optimize the network structure when it consists of billions of nodes. Thus, there is an urgent need to

978-1-4577-0220-4/16/$31.00 ©2016 IEEE

increase the computational efficiency of the IoH to handle large volumes of data.

In this study, a novel scheme is introduced to parallel process large volumes of data in the IoH. Specifically, we decompose the large-scale problem of network optimization into local ones and resolved them in a parallel manner. First, a stochastic gradient descent approach is designed to embed patients with similar conditions into a local network. Positions of network nodes within a neighboring region are tuned to achieve the optimal structure of a local network. Second, local networks are optimally pieced together to obtain a global network. Notably, the proposed scheme enables the implementation of parallel computing on a multitude of processors. Our contributions in this present investigation are highlighted as follows:

1) *Reducing computational complexity*: We developed a stochastic gradient descent approach to significantly reduce the computational complexity on a single processor, i.e., from $O(N^3)$ to $O(N)$.

2) *Parallelization*: We proposed a parallel scheme to scale up the network optimization by simultaneously utilizing multiple processors, which achieves 80-fold faster than conventional algorithms for embedding 20000 patients.

The remainder of this paper is organized as follows: Section II introduces the background. Section III presents research methodology. Section IV shows a case study and experimental results. Section V discusses and concludes this investigation.

II. Background of Parallel Computing

Traditionally, computing tasks are executed in a serial manner (see Fig. 1a). Serial computing is efficient when the given data set is small. However, it is impractical to process large volume of data. For example, information is extracted from millions of online documents during a web mining research. Serial computing has been shown to be ineffective for such large and complex problem. Therefore, researchers are seeking a way to scale up the algorithm and utilize more computing resources to collaboratively complete a task.

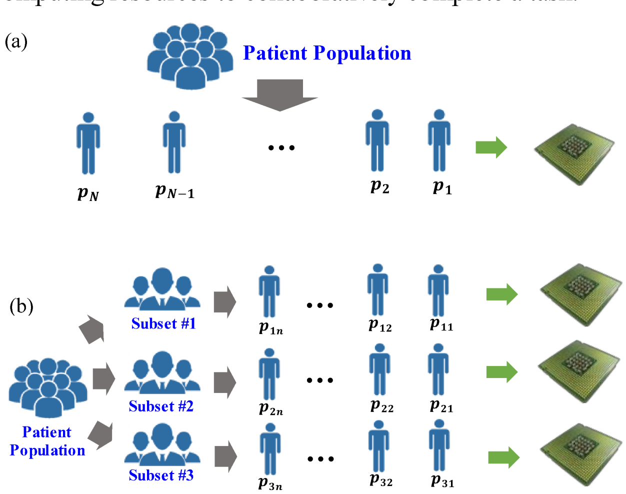

Fig. 1. (a) Serial computing *vs*. (b) Parallel computing.

Rapid advancements of information technology have catalyzed the incredibly fast growth in computing power. In particular, parallel computing has been employed to harness multiple processing units to solve a problem. The basic idea of parallel computing is to break the overall computing task into multiple sub-tasks, which are independent to each other. Each sub-task is then assigned to an individual processor (see Fig. 1b). As such, each processor can execute its part of the algorithm simultaneously with the others, which critically reduces the computing time compared with serial computing. With the availability of multi-core CPUs and cloud computing technology, parallel computing can be readily achieved by deploying strategies such as multi-threaded and single-instruction-multiple-data. Nowadays, parallelism has been widely used in manufacturing, genetics research, search engine, financial modeling and computer vision.

III. Research Methodology

A. Overview of the Internet of Hearts

As shown in Fig. 2, the IoH embodies a number of networked components such as ECG sensors, mobile devices, patients, physicians, emergence centers and hospitals. The wearable ECG device records patients' cardiac activity 24/7 and it is seamlessly connected to the smartphone via Bluetooth. After pre-processing (e.g., denoising), collected data are transmitted to the cloud via 4G network. Analytics models are running on the cloud to uncover hidden cardiac patterns and detect early signs of cardiac events. The IoH enables one-to-one and one-to-many communications between patients and cardiologists. Cardiologists are able to access patients' data, review analytical results and communicate with patients and other cardiologists anytime and anywhere. In addition, emergency centers and hospitals are integral components of the IoH. If a patient's condition is highly risky, emergency centers will be instantly notified.

Fig. 2. The architecture of the Internet of Hearts (IoH).

Analytics models in the IoH were developed in our previous studies [2-4]. First, adaptive basis functions are designed to iteratively represent patients' space-time vectorcardiogram (VCG) signals. Only a limited number of parameters are needed to represent large amount of VCGs, while fully preserving the diagnostic information. Such sparse representation facilitates data compression and extraction of pertinent features. Further, a dynamic network model is developed for real-time recognition of disease patterns. Dissimilarities among functional recordings from N patients are measured by the spatiotemporal warping. To this end, an $N \times N$ warping matrix is obtained, in which each element stores the warping distance between i^{th} and j^{th} patients. Then, the warping matrix is optimally embedded into a high-dimensional network. The distance between network nodes preserves the warping distance of corresponding patients. In this way, the location (i.e., high-dimensional coordinates) of a network node reveals diagnostic information of the patient. Cardiologists can easily monitor the change of a patient's cardiac condition by tracking the trajectory of corresponding node in the dynamic network.

B. Stochastic Gradient Descent for Network Embedding

Traditional algorithms, e.g., classical multidimensional scaling (MDS) and scaling by majorizing of a complicated function (SMACOF) [5] are limited in their capability to handle big data. For example, the computational complexity of classical MDS is $O(N^3)$, due to the double centering operation and eigen-decomposition. The computational complexity of the SMACOF is $O(N^2)$, because the Guttman transform consists of a multiplication between an $N \times N$ matrix and an $N \times L$ matrix (L is the dimensional of the network). As such, it is not cost-effective to implement classical MDS or SMACOF for large-scale network optimization (e.g., in the IoH).

TABLE I. DEVELOPED APPROACH FOR NETWORK EMBEDDING

K – total number of training iterations;
Create index sequence by randomly permuting the set $\Theta = \{1, 2, \ldots, N\}$;
Initialize locations of nodes in the high-dimensional space;
for $k = 1 : K$
  for $n = 1 : N$
    Choose index of the fixed network node $i = \Theta(n)$;
    Update the location of all other nodes $\boldsymbol{x}_j (j \neq i)$ by Eq. (5);
  end
end

In this research, a stochastic gradient descent approach is developed for optimally projecting large numbers of patients into a network, while maintaining a low computational complexity. Table I summarizes the proposed approach. Let δ_{ij} denotes the dissimilarity between i^{th} and j^{th} functional recordings obtained from spatiotemporal warping, and $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ denote locations of i^{th} and j^{th} nodes in the high-dimensional network. The quadratic objective function S can be formulated as:

$$S = \frac{1}{2} \sum_i \sum_{j \neq i} \left(\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - \delta_{ij} \right)^2 \Gamma\left(\|\boldsymbol{x}_i - \boldsymbol{x}_j\|, \lambda \right) \quad (1)$$

where $\Gamma(\cdot)$ is a bounded and monotonically decreasing function to favor the local topology in the network. Here, $\Gamma(\cdot)$ is selected as an exponential function:

$$\Gamma\left(\|\boldsymbol{x}_i - \boldsymbol{x}_j\|, \lambda \right) = e^{(-\|x_i - x_j\| / \lambda)} \quad (2)$$

and λ is a user defined parameter. Minimization of Eq. (1) with respected to $\boldsymbol{x}_i$'s is achieved by a stochastic gradient descent approach. Specifically, the objective function S can be rewritten as a summation of partial costs generated from adjusting individual network node i, i.e., $S = \frac{1}{2} \sum_i S_i$. In each iteration, a node i is randomly chosen and its location in the high-dimensional network (e.g., $\boldsymbol{x}_i$) is fixed. All other nodes $\boldsymbol{x}_j (j \neq i)$ are updated based on the following rule:

$$\boldsymbol{x}_j \leftarrow \boldsymbol{x}_j - \alpha(t) \frac{\partial S_i}{\partial \boldsymbol{x}_j} \quad \forall j \neq i \quad (3)$$

where the learning rate $\alpha(t)$ is a monotonically deceasing function with the form $\alpha(t) = \alpha_0 / (1 + t)$. According to the chain rule, Eq. (3) can be rewritten as:

$$\boldsymbol{x}_j \leftarrow \boldsymbol{x}_j - \alpha(t) \frac{\partial S_i}{\partial \|\boldsymbol{x}_i - \boldsymbol{x}_j\|} \frac{\partial \|\boldsymbol{x}_i - \boldsymbol{x}_j\|}{\partial \boldsymbol{x}_j} \quad \forall j \neq i \quad (4)$$

Similar to [6], let's ignore the derivative of the exponential function in Eq. (2). Since $\frac{\partial \|x_i - x_j\|}{\partial x_j} = \frac{x_j - x_i}{\|x_i - x_j\|}$, the update rule of Eq. (3) can be written as:

$$\boldsymbol{x}_j \leftarrow \boldsymbol{x}_j - \alpha(t) \frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\| - \delta_{ij}}{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|} e^{(-\|x_i - x_j\| / \lambda)} (\boldsymbol{x}_j - \boldsymbol{x}_i) \quad (5)$$

The approach iterates until the maximum number of learning iterations is reached. Notably, the computational complexity for each iteration is only $O(N)$.

C. Parallel Computing for Optimizing Large-scale Network

As mentioned in Section II, it is impractical to optimize the large-scale IoH network on a single processor. It is worth mentioning that we construct the network by considering the warping distance between each pair of patients. This is similar to the localization problem in large-scale wireless sensor network, where exact locations are known for a limited number of sensors. For the majority of sensors, only local information (distance between a sensor to its neighbors) is available. In the literature, distributed localization is developed to address this problem, which first configures local networks and then merges them together. It has been shown that distributed localization provides the best solution in terms of efficiency and accuracy for locating large amount of sensors.

It is noteworthy that the distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ in the high-dimensional network preserves the warping distance. As shown in Fig. 3, patients with similar conditions in the IoH are tightly connected and form small local networks. On the contrary, connections among patients with different cardiac conditions (e.g., myocardial infarction and hypertension) are indeed loose. Thus, local networks are approximately independent to each other. This enables the decomposition of a global network into local ones for parallel computing. In this study, patients with similar conditions (i.e., small warping distances) are embedded into a local network by the stochastic gradient descent approach. Notably, local network embedding is simultaneously computed using multiple processors, which significantly reduces the computation time.

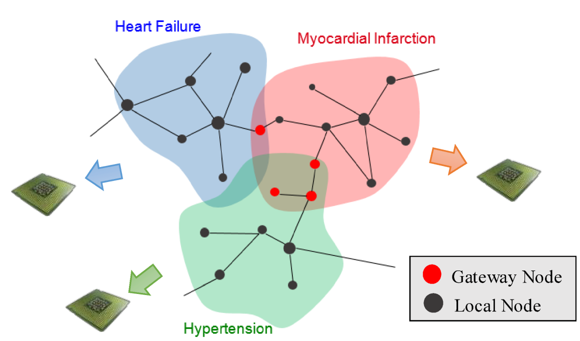

Fig. 3. Illustration of parallel computing for a large scale network.

Further, local networks are merged sequentially to form a global network. First, we select a local network with largest number of patients as the core map. Then the core map is extended by absorbing networks with shared nodes (i.e., gateway nodes). Each time, a local network with the maximum number of gateway nodes is selected and merged into the core map. Eventually the core map covers all patients. Notably, a gateway node has both intra and inter-group connections and it has coordinates in ≥2 local networks (See Fig. 3). By scaling, rotating and translating, coordinates of gateway nodes between the local network and the core map are first matched. Then, the same transformation is applied to the rest of the local network and merge it into the core map.

IV. Experiments and Results

Experiments are designed to evaluate performances of the developed methodology. As shown in Fig. 4, stochastic gradient descent is more efficient than traditional eigen-decomposition for network embedding on a single processor. Here, computation time is calculated for an individual iteration of stochastic approach. Notably, the eigen-decomposition is not suited for large data set due to its $O(N^3)$ complexity. However, stochastic approach reduces the computational complexity to $O(N)$ in each iteration, which occupies less CPU resource and memory.

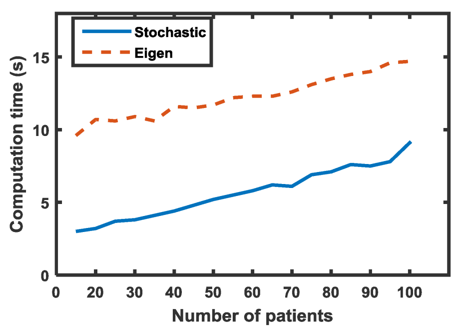

Fig. 4. Computation time of stochastic gradient descent approach and Eigen-decomposition approach on a single processor.

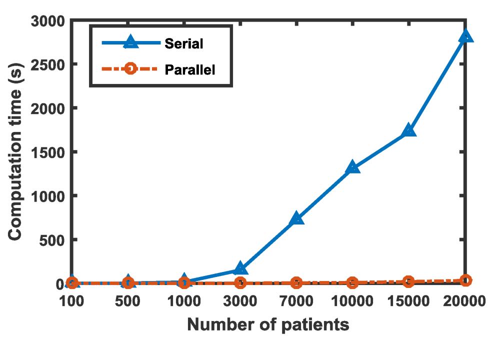

Fig. 5. Computation time of the developed parallel computing scheme and traditional serial computing.

Further, the performance of proposed scheme is compared with traditional serial computing. Here, stochastic gradient descent is deployed in both schemes. The entire set of patients is divided and assigned to 12 CPU cores. As shown in Fig. 5, two schemes achieve similar performances when the number of patients are small (e.g., N <1000). However, the computation time of serial computing increases significantly when large number of patients are involved. On the contrary, the computation time of the proposed scheme does not increase much. It may be noted that the gap between two curves are enlarged dramatically with the increasing number of patients. When 3000 patients are embedded, serial computing is 150s slower than parallel computing. When the number of patients reaches 20000, serial computing is >2700s slower than parallel computing. Performances of the parallel computing scheme with respect to even larger data sets are evaluated and shown in Table II.

Table II. Computation Time of Proposed Scheme With Respect to The Increasing Number of Pateints

Number of Patients	Computation Time (s)
30000	68.5
40000	119.6
50000	184.8
60000	262.9
70000	350.9
80000	460.4
90000	588.9
100000	718.9

V. Discussion and Conclusions

Traditional approaches are limited in their ability to optimize large-scale network of patients for the early diagnosis of diseases and timely medical intervention. In this present paper, we developed a new scheme to parallel process large volumes of data in the IoH, which has been granted an international PCT patent [7]. First, we designed a stochastic gradient descent approach to efficiently embed cardiac patients into local networks. Then, local networks are optimally pieced together to obtain the global network. Experimental results show that the developed scheme significantly increases the capacity for optimizing large-scale network of patients. When the number of patients reaches 100000, the developed scheme can process them in around 10 minutes, whereas traditional serial computing has to take hours. By parallelizing the computation into multiple processors, the IoH shows strong potentials to handle millions and billions of patients to improve connected care worldwide.

References

[1] C. Kan, Y. Chen, F. Leonelli and H. Yang, "Mobile sensing and network analytics for realizing smart automated systems towards health internet of things," in *Proc. 11th Ann. IEEE Int. Conf. Autom. Sci. Eng. (CASE 2015)* Gothenburg, Sweden, Aug. 24-28, pp. 1072-1077.

[2] H. Yang, C. Kan, G. Liu and Y. Chen, "Spatiotemporal Differentiation of Myocardial Infarctions," *Automation Science and Engineering, IEEE Transactions on*, vol. 10, No.4, pp. 938-47, 2013.

[3] G. Liu and H. Yang, "Multiscale Adaptive Basis Function Modeling of Spatiotemporal Vectorcardiogram Signals," *Biomedical and Health Informatics, IEEE Journal of*, vol. 17, No.2, pp. 484-492, 2013.

[4] H. Yang, S. T. S. Bukkapatnam and R. Komanduri, "Spatiotemporal representation of cardiac vectorcardiogram (VCG) signals," *Biomedical Engineering Online*, vol. 11, No.12, 2012.

[5] S. Bae, "Scalable high performance multidimensional scaling," *Ph.D. Dissertation, Indiana University,* 2012

[6] P. Demartines and J. Herault, "Curvilinear component analysis: a self-organizing neural network for nonlinear mapping of data sets," *Neural Networks, IEEE Transactions on,* vol. 8, *No.1,* pp. 148-154, 1997.

[7] H. Yang, "Systems and Methods for Diagnosing Cardiovascular Conditions", *International PCT Patent*, No. WO2014043216 A1, priority date: September 13, 2012, publication date: Mar 20, 2014.