Communications on Applied Electronics (CAE) – ISSN: 2394-4714
Foundation of Computer Science FCS, New York, USA
Volume 7– No. 26, February 2019 – www.caeaccess.org

A Proposed Architecture for Predicting Breast Cancer using Fog Computing

Laila A. Abd-Elmegid
Faculty of Computers and Information
Helwan University
Cairo, Egypt

ABSTRACT
A new emerged concept, known as fog computing, aims to local processing and storage of large amounts of data at its sources. Such concept differs from the centralized style of processing and storage used with the cloud computing concept. Fog computing technology improves both the performance and the efficiency. The main cause of advantages gained from fog computing is the reduced amount of transferred data to the cloud need to be analyzed, processed, and stored. little research has been dedicated to study how to efficiently apply big data analytics locally where data has been collected and stored. This study proposes a fog computing-based architecture for prediction of the breast cancer prognosis. The proposed architecture uses the BCOAP model for the prediction purpose. It addresses the problem of real time processing of large amounts of data without overwhelming the data center at the cloud. It distinguishes between the processes to be performed on the cloud and the others to be performed on the fog. Initial prototype of the proposed architecture has highlighted its efficiency and how it increases the potentials of the applied BCOAP model.

General Terms
Data mining, Data Science

Keywords
Breast cancer, BCOAP, Classification, Cloud computing, Fog computing.

1. INTRODUCTION
One of the most famous types of cancer is breast cancer that mainly infects women. Early diagnosis of breast cancer results in good opportunities of surviving. Three outcomes are mainly used for the prognosis of breast cancer; Survival rate, Recurrence detection, and Disease-free survival [1]. Ahmed et al [2] proposed an accurate prediction model known as BCOAP for predicting these outcomes efficiently. The classifier of the BCOAP model learns from grouped data clusters rather than the full data set. Additionally, the classifier is optimized by tuning its parameters via specified number of iterations. The experimental study showed how utilizing these techniques resulted in more accurate prediction results. The model has been mainly proposed for centralized processing where the data is generated, stored, and processed at its source.

Emerging technologies characterized by processing large amount of data need special computing platforms. Such platforms should be latency aware to support real-time processing. According to the cloud infrastructure of on-demand services, processing data is done in the cloud. If the volume of data is large, it is impractical to manage such data exclusively in the cloud for latency-sensitive applications like healthcare solutions that require emergency response. A new computing technique known as Fog computing is emerged to allow applications to perform their data processing both on the cloud and the end devices. The Fog architecture is represented in three layers, the first layer contains the IoT devices and sensors generating the data and it is connected to the intermediate layer that contains fog devices. Fog devices are locally located at the end users and it is responsible for the required processing and storage of data. The third layer contains the data center located at the cloud. The architecture allows most computations to be done on the edge and consequently eliminates the full dependency on the cloud resources. This study proposes a fog-based architecture for predicting the prognosis of breast cancer. It works on the centralized BCOAP model. It shows where each phase of the model should be executed and how such architecture is efficiently working [3].

The paper is structured as: Section 2 illustrates in details the preliminaries of breast cancer, the BCOAP model, and the fog computing. A literature review is showed in Section 3. In Section 4, the details of the proposed architecture are presented. Finally, a conclusion of the work of the research is given in Section 5.

2. PRELIMINARES

2.1 Predicting Breast Cancer
The cancer that being formed in the beast cells is known as breast cancer. It occurs when the breast cells are growing in an abnormal manner. This is done by rapid and continuous dividing of cells until a lump has been formed. It is the most prevalent type of cancer diagnosed in women. However, it can rarely infect also men. Early diagnosis of such disease leads to good opportunities of recovery. The main outcomes of prognosis breast cancer are: Survival rate, Disease-Free Survival, and the Recurrence. Survival rate refers to how many breast cancer patients' have survived for a time ranging from 5, 7, and 10 years [1].it is an important indicator about the degree of efficiency of applied treatment. To state if the cancer backed to the same breast or a different one, Recurrence is the outcome for this case. Disease Free Survival (DFS) is a measurement of the time between the start of patient's treatment and patient's recovery of cancer [4].

Ahmed et al [2] proposed a classification model, known as BCOAP, for the prediction of the main prognosis's outcomes of breast cancer. The main objective of the model is to perform the prediction and increase the accuracy of results as possible simultaneously. The BCOAP model works through four phases as follow:

Phase 1 Clustering: The BCOAP model uses the clustering technique to enable the classifier to learn from similar groups of data, according to specified similarity measurements, individually rather than learning from the whole data set. Such technique increases the accuracy of the classifier as proved in the experimental study.

Phase 2 Features Selection: The next phase aims to decrease the number of attributes for more accurate and faster performance. The output of this phase is the most important features critical to the prognosis of breast cancer. The features are sorted descending according to their importance and relevance degrees.

Phase 3 Classification: The classification is done through building decision trees. The goal is achieving a balanced point between accurate prediction and building compact tree with small number of decisions.

Phase 4 Hyper-Parameters Optimization: the model at the last phase uses the Hyper-Parameters Optimization technique to optimize the predictor tree by tuning the important parameters through an iterative process. Figure 1 shows the architecture of the four phases of the BCOAP model.

2.2 Fog Computing

The known cloud computing concept allows centralized data processing and storage. The main advantage of cloud computing is the optimization of resources utilization. However, such advantage is gradually lost when working with big data applications. This is because the moving of large volumes of data over the network that results in severe overhead regarding time, cost, energy consumption, and throughput. To handle this problem, different aspects of data processing could be performed locally at its source [5].

Recently, a new concept has been emerged known as Edge computing. This concept contrasts with the concept of cloud computing. The difference between the two concepts emerges from the communication style with data center. Cloud computing communicates directly only with the centralized data center. However, edge computing communicates in two dimensions; primarily with the edge data center exists at the device and secondary with the central data center at the cloud. Consequently, edge computing can be considered as a computing topology. Table 1 highlights the similarities between the fog and edge computing while Table 2 highlights the differences between them [60.

Table 1: Similarties between Edge and Fog Computing [6]

Factor	Edge Computing	Fog Computing
Architecture	Hierarchal, Decentralized, Distributed	Hierarchal, Decentralized, Distributed
Proximity to end devices	Located in end devices	Near network(s) hop(s)
Latency	Low	Low
Bandwidth costs	Low	Low
Resource	More limited	Limited
Computation and storage capabilities	More limited	Limited
Mobility	Supported	Supported
Scalability	High	High
Service	Virtualization	Virtualization

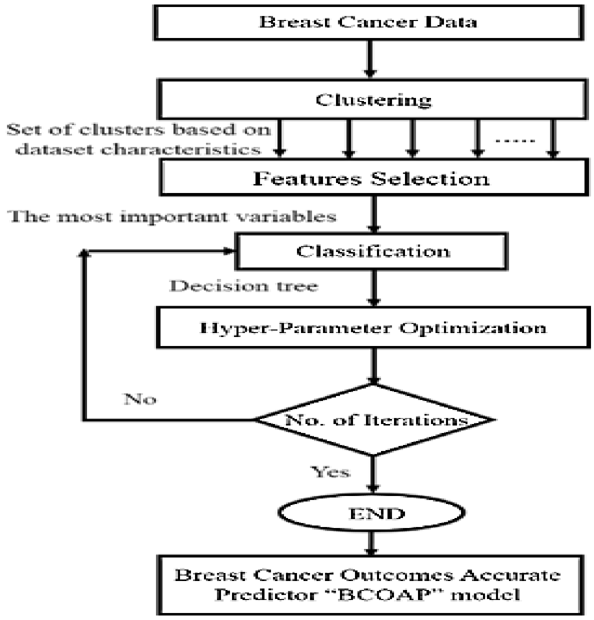

Figure 1: The BCOAP Model [2]

Table 2: Diffeences between Edge and Fog Computing [6]

Factor	Edge Computing	Fog Computing
Location of data collection, processing and storage	Network edge, edge devices	Near edge and core networking network edge and core networking devices
Handling multiple IOT applications	Unsupported	Supported
Resource contention	Serious	Slight
Focus	Things level	Infrastructure level

At January 2014 CISCO has proposed the fog computing. As mentioned in [3] [6] [7] edge computing can be considered as a component of fog computing. Fog computing includes both edge processing and network connections for data transfer. Fog computing considers more specifically the data manipulation processes executed near edge devices.

3. LITERATURE REVIEW

According to the latency-sensitive characteristic of healthcare systems, fog computing is considered the most suitable computing platform for this type of systems. The applications of fog computing on healthcare vary in its purpose from diagnosis, monitoring, detection, and visualization. The following section highlights the famous fog-based healthcare systems [8] [9].

First, General studies of fog computing with healthcare that highlight its potentials, challenges, and limitations exist at [10] and [11]. For detection purpose, Cao et al proposed FAST [12], a fog computing distributed system to monitor fall for stroke mitigation. They have implemented fall detection algorithms and incorporated them into fog-based distributed system. The system distributes the analytics throughout the network by splitting the detection task between the edge devices represented in smart phones attached to the users and the server exists in the cloud.

Kyriazakos et al [13] presented eWALL as an intelligent home environment offering personalized context-ware applications based on advanced sensing and fog computing on the front and cloud solutions on the back. A health Fog framework is developed in [14] where fog computing is used as an intermediary layer between the end users and the cloud. The design of Health Fog successfully reduces the extra communication cost that is usually high in similar systems. Another system known as fHealth is proposed as an open source framework used as a use case of fog-oriented health care applications. A thorough study of fog computing regarding its architecture, key technologies, applications and open issues is presented in [6] in addition to a comparison between the pre-illustrated fog-based health care frameworks.

Table 3: Comparison between Different Fog-based Healthcare Frameworks [6]

Framework	Diseases	Techniques	Devices	Open source?
FAST	Brain attack	Detection	Smart phones, Cloud servers	No
eWALL	CODD, Mild Dementia, Aging related diseases	Monitoring	Sensors, Actuators, eWALL cloud, Cloud middleware	No
HealthFog	Multi purposes	Recognition	Smart phones, Smart home devices, Wearable sensors	No
fHealth	Fitness	Tracking	Smart phones, Cloud servers	Yes

4. A PROPOSED FOG-BASED ARCHITECTURE FOR PREDICTING BREAST CANCER PROGNOSIS

As mentioned in Section 2.1, the BCOAP model has been designed to be implemented in a centralized style. Such implementation limits the potentials that could be gained from the model. In this section we propose a new architecture for the implementation of BCOAP model based on fog computing. Consequently, the new proposed architecture inherits all the advantages of fog computing described earlier and increase the potentials of BCOAP model. The general proposed architecture relies on a continuous interplay between the layer of the cloud and the layer of the fog as shown in Figure 2.

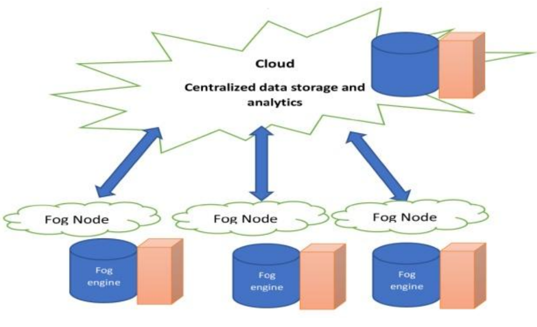

Figure 2: The Inteaction between the Cloud and Fog Layer of the Proposed Architectuer

The proposed architecture consists of four layers as follows. The general architecture is shown in Figure 3 whereas the detailed flow of the architecture is shown in Figure 4.

- The base layer represents the system entities, Hospitals and Clinics, where the medical data is generated.
- The next layer, fog computing, firstly collects the patients' data and starts with its preprocessing. The main activities of data preprocessing are data cleaning and transformation. Next step is data mining. At the fog, mining the data is just an implementation of the optimized tree predictor, developed at the cloud, on the collected data. Therefore, there is no need for special storage and processing requirements at any fog node. Ideally, the fog node is responsible of local data collection of its patients and light processing represented in direct implementation of a given predictor tree. For security issues, the fog layer implements the Cloud Access Security Brokers (CASB). The main roles of CASB is to monitor and control the data privacy on the cloud. Accordingly, it prevents unauthorized entities from joining the architecture from anywhere. A detailed technical information about CASB is described in [6].
- The cloud layer, firstly, collects data generated from all fog nodes and integrate it in a central data repository. The central data mining engine at the cloud is responsible of performing the phases of clustering, features selection, classification, and hyper-parameters tuning. The main output of this step is an optimized tree predictor that will be sent to the fog node.
- The end users' layer is represented in medical institutes, hospitals, clinical, and research institutes. This layer can interact with the cloud layer for central knowledge sharing of the complete data set collected from all medical centers. As noticed from the architecture, hospitals and clinics are common entities between the system entities layer and end users layer.

5. CONCLUSIONS

Recently, a new computing platform, Fog computing, has been emerged to locally manipulates the data at its sources. Consequently, it reduces the latency and cost payed as a result of transferring the data to a data center at the cloud. Such platform is desirable for latency sensitive and mission intensive services. In this study, we propose an architecture based on fog computing for predicting the prognosis of breast cancer. The architecture works on the BCOAP model by extending its phases to be executed on the fog devices and the cloud simultaneously. The proposed architecture efficiently overcomes the problems of centralized processing and allows efficient real time processing of large amount of data

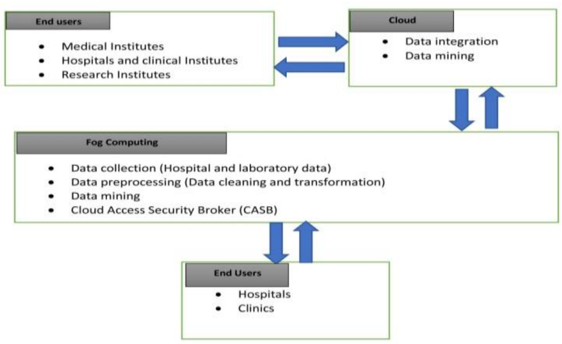

Figure 3: The General Architectuer of the Proposed Architectuer

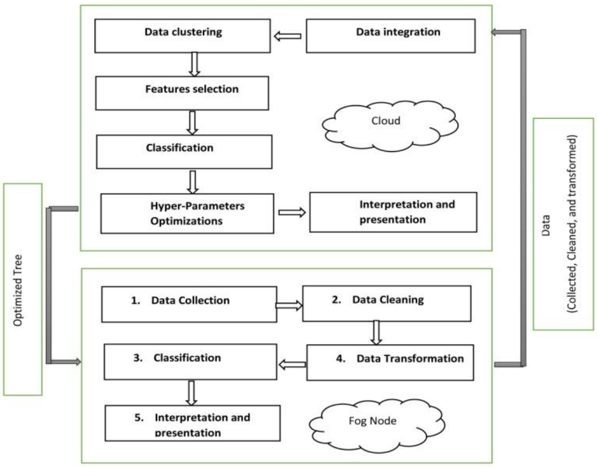

Figure 4: The Detailed Architectuer of the Proposed Architectuer

6. REFERENCES

[1] Haffty B., Buchholz T., Perez C. 2008. Early stage breast cancer. In: Halperin EC, Perez CA, Brady LW, editors. Principles and practice of radiation oncology. 5th ed. Philadelphia: Lippincott Company

[2] Ahmed, A., Laila A., Sherif K., Ayman G.,2018. Classification Based on Clustering Model for Predicting Main Outcomes of Breast Cancer Using Hyper-Parameters Optimization. International Journal of Computer Science and Applications 9 (12), 268-273

[3] Zhu, J., Chan, D.S., Prabhu, M.S., Natarajan, P., Hu, H., Bonomi, F., 2013. Improving websites performance using edge servers in fog computing architecture. In: Proceedings of the IEEE Seventh International Symposium on Service-Oriented System Engineering, pp. 320–323.

[4] Houriyeh, E., Mitra M., Reza K., Raziyeh H., Ali N., Vahid M. 2017. Prognosis and Early Diagnosis of Ductal and Lobular Type in Breast Cancer Patient" ,Iran J Public Health,. 46 (11), pp.1563-1571.

[5] Botta, A., Donato, W.D., Persico, V., Pescap, A., 2015. Integration of cloud computing and internet of things: a survey. Future Gener. Comput. Syst. 60 (5), 23–30.

[6] P. Hu, S. Dhelim, H. Ning, T. Qiu, Survey on Fog Computing: Architecture, Key Technologies, Applications and Open Issues, Journal of Network and Computer Applications 98 (2017) 27-42

[7] Ahmed, E., Rehmani, M.H., 2017. Mobile edge computing: opportunities, solutions, and challenges. Future Gener. Comput. Syst. 70, 59–63.

[8] Stojmenovic, I., Wen, S., 2014. The fog computing paradigm: Scenarios and security issues. In: Proceedings of the Federated Conference on Computer Science and Information Systems, pp. 1–8.

[9] Aazam, M., Huh, E.N., 2016. Fog computing: the cloud-iot/ioe middleware paradigm. IEEE Potentials 35 (3), 40–44.

[10] Stantchev, V., Barnawi, A., Ghulam, S., Schubert, J., Tamm, G., 2015. Smart items, fog and cloud computing as enablers of servitization in healthcare. Sens. Transducers 185 (2), 121–128.

[11] Shi, Y., Ding, G., Wang, H., Roman, H.E., 2015. The fog computing service for healthcare. In: Proceedings of the International Symposium on Future Information and Communication Technologies for Ubiquitous Healthcare, pp. 70–74.

[12] Cao, Y., Chen, S., Hou, P., Brown, D., 2015. Fast: a fog computing assisted distributed analytics system to monitor fall for stroke mitigation. In: Proceedings of the IEEE International Conference on Networking, Architecture and Storage, pp. 2–11.

[13] Kyriazakos, S., Mihaylov, M., Anggorojati, B., Mihovska, A., Craciunescu, R., Fratu, O., Prasad, R., 2016. eWALL: an intelligent caring home environment offering personalized context-aware applications based on advanced sensing. Wirel. Personal. Commun. 87 (3), 1093–1111.

[14] Ahmad, M., Amin, M.B., Hussain, S., Kang, B.H., Cheong, T., Lee, S., 2016. Health fog: a novel framework for health and wellness applications. J. Supercomput. 72 (10), 36773695.