CLAWER: Context-aware Cloud-Fog based Workflow Management Framework for Health Emergency Services

Shreya Ghosh*, Jaydeep Das†, Soumya K Ghosh‡, Rajkumar Buyya§
*†Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, India
†Advanced Technology Development Centre, Indian Institute of Technology Kharagpur, India
§Cloud Computing and Distributed Systems (CLOUDS) Lab, School of Computing and Information Systems, The University of Melbourne, Australia
Email: *shreya.ghosh@ieee.org, †jaydeep@iitkgp.ac.in, ‡skg@cse.iitkgp.ac.in, §rbuyya@unimelb.edu.au

Abstract—With the major development of sensor technologies and advancements of communication network infrastructures, there is a growing interest to add more intelligence in the e-health monitoring for facilitating an effective healthcare system. While IoT devices are capable of continuous health-parameter sensing and providing notifications to the user, an effective business process management (BPM) facilitates effective system integration and data processing workflow. This paper proposes an efficient framework for managing emergency situations (specifically, health-related) through the analysis of heterogeneous data sources. The proposed framework, named CLAWER (Cloud-Fog Based Workflow for Emergency Services), aims to bridge the gap between process management and data analytics by providing an automated workflow for personalized health-monitoring and efficient recommendation system. Here, the IoT devices are used for collecting the movement and health data. The smart phone can act as an edge device to acquire data with user movement information. The accumulated data is initially processed inside the fog device, and finally the analysis and recommendations are generated by the cloud. In this paper the indoor health-status of the users are analysed in small cell cloud enhanced eNode B, which is used as fog device. The generated recommendations are stored in the fog device to provide the recommendations to the users with low latency and in timely manner. The experimental analysis of CLAWER yields better precision and recall values than the existing methods.

Keywords—Cloud Computing; Context-aware, Workflow, Mobility, Health management.

I. INTRODUCTION

The pervasive use of GPS-equipped smart-devices, improved sensor and internet technologies have accelerated dramatic revolution in individuals daily living. In the era of Internet of Things (IoT), smart communication among objects such as vehicles, device, buildings and people, has facilitated intelligent human living environment. Smart transportation or mobility services, smart-home, personalized recommendation system, efficient health-monitoring are only a few use-cases of this burgeoning technology. In addition, cloud, fog, edge based [1] infrastructures enhance the functionality of such IoT applications in terms of reduced energy consumption and delay.

The emerging cloud/fog/edge/IoT paradigms have become a key enabler of effective business processes including better resource planning and management, customer satisfaction by providing on-time services. For instance, if the BAN (Body area network) can sense abnormal health-condition of an user, it can trigger alert and may help in taking early preventive measures. Further, other disasters/ emergency situations such as fire, traffic blockage, flight cancellations can be managed efficiently. In brevity, IoT devices facilitate continuous monitoring of events through devices (smartphones, wearable devices, sensors), and cloud/fog/edge/IoT devices provide the computational and storage infrastructures to provision on-demand services. This technology has the potential to change the current business trends, and enterprises are increasingly adapting IoT-based framework in their business workflow. On the other hand, another key challenging aspect is orchestration and manageability of disparate heterogeneous systems, or agencies to work together. For example, while providing support to an ailing person, the ambulance service, healthcare centers and city-traffic need to seamlessly interact to provide an efficient health care service. In such scenarios, the real-time data needs to be analyzed and several business agencies need to cooperate for efficient service and cost-savings.

The IoT healthcare market is growing rapidly for rising demands of improved healthcare with reduced cost/delay, and reliable connectivity. In this paper, we propose an end-to-end workflow management framework, named CLAWER, to provide delay-aware effective personalized health-care services to user. The key contributions of this paper are summarized as follows:

1) A cloud-fog-edge-IoT based context aware framework namely CLAWER is proposed to provision personalized e-health monitoring and healthcare service to user in minimum delay. The hierarchical framework has cloud servers, fog, edge and IoT devices in several layers and it analyzes the sensing data and takes preventive measures in case of emergency. The framework has been implemented and tested over real-life
data samples and promising results have been found.

2) An automated workflow is presented in CLAWER framework, when different modules of it works together and adaptive measure is taken based on the data analysis and contexts. The automated workflow is useful in taking decisions when emergency situation arises.

3) CLAWER presents a markov-chain based path-prediction model for extracting optimal path. Further, the prediction algorithm is computed on MapReduce paradigm for faster response. This mobility analytics is beneficial when the user has some serious health problem and needs to travel to the health-care center in minimum delay.

In this direction, CLAWER is an well-designed, integrated framework which facilitates effective health-monitoring and decision making in case of emergency situations. The rest of the paper is organized as follows. Section II discusses the related works, CLAWER framework is presented in section III and experimental evaluation is discussed in section IV. Finally, the paper is concluded in section V.

II. RELATED WORK

With this advancement and growth in wireless sensor network, Internet of Things (IoT) has been introduced [2]. In IoT the sensors and actuators collect the status of the environmental objects of the surrounding and the processing takes place usually inside the cloud. However, the data storage and processing in the cloud suffers from delay and energy consumption, which has been dealt with using edge and fog computing [3], [4]. The integration of IoT with fifth generation network has been discussed in [5]. E-health monitoring has also become an emerging area of interest in IoT. Various applications such as Samsung S Health, Apple Healthkit, Google Fit, and Microsoft Health are available today. In Internet of Health Things (IoHT), exchange and processing of the data is performed to monitor health condition of individuals by integrating sensor or IoT devices with advanced mobile technologies. Existing e-health applications use cloud servers for processing of the health data. However, cloud only framework may affect the quality of service in terms of delay and power consumption [6], [7], which has been overcome through the use of fog computing [8], [9].

The context-aware health care system for smart cities has been proposed in [10]. For personalized mobile health care, an IoT based interconnection framework has been designed in [11] for continuous and remote monitoring of vital signs. In the context of business process management (BPM), Xiong et al. [12] presents a framework named SmartCrowd to deliver crowd-sourcing tasks by proposing a novel workflow model. Another work [13] presents an efficient algorithm for for regenerating BPM in multi-cloud environment for cost reduction.

Figure 1: Workflow of CLAWER
Nevertheless, CLAWER is the first work to integrate several business partners/agents (such as hospital, car, IoT devices, cloud data-centers) and analyse location, contextual information, and health information for recommending users’ health status and subsequently taking preventive measures. To the best of our knowledge, no other existing works have considered such workflow-oriented data analysis for building an efficient health monitoring system.

III. CLAWER: ALGORITHMS AND IMPLEMENTATIONS

In this section, we describe the proposed framework named CLAWER that is capable to model and analyse health profiles of individuals and recommend medical assistance periodically. Figure 1 illustrates the overall workflow of the health monitoring and recommendation system. Figure 2 represents the hierarchical placement of cloud/ fog/ edge/ IoT devices in both indoor and outdoor regions. As depicted in the figures, the proposed system architecture consists of users with wearable devices, handheld devices, SCcENB (small cell cloud enhanced eNodeB), RSUs (Road Side Unit) and cloud servers. The handheld and wearable devices are equipped with various sensors and applications helping in health-related and movement data capturing. The handheld devices are connected with the network through SCcENB in indoors. Here, these handheld devices are edge devices, which process data locally. Similarly, when the user is in move, the vehicle (such as ambulance) act as an edge node. These edge devices are connected to fog nodes: RSU (outdoor) and SCcENB (indoor). The fog nodes helps in computing, storing and communicating between edge devices and cloud servers. The cloud storage and computing capacity are utilized to analyse the accumulated huge movement and health information. Besides, the system is also capable to detect abnormal health condition, and notifies to the caregivers accordingly.

In case abnormal health condition is detected or emergency situation occurred, the user needs to reach to the nearby healthcare center in minimum delay to avoid any fatal condition. Along with the ambulance service, another fundamental aspect is to find out the optimal path (less congestion) to traverse the distance in minimal time. This section presents the methods and their implementations to extract the path.

A. Health status monitoring

Body area network (BAN) is used to capture the health data of a user, such as blood pressure, body temperature, heart rate etc. The collected health data are sent to the smart phone of the user. The geo-location information of the user, health data and contextual data are accumulated inside the smart phone of the user, which forwards it to the fog device under which the user is registered. The fog device performs preliminary processing on the data before forwarding to the cloud. In case of health data analysis, a functional model is generated to verify whether the user’s health status is normal or not. In the functional model, the user’s geo-location information, contextual information such as humidity, light intensity, temperature of the environment and health data are provided as input. Based on the health and context information, if any abnormality is detected, then the user is notified through an alert. Let the collected health data set is $data_{h}$, context data set is $data_{c}$ and geo-location information is $data_{g}$, then the predicted health status ($S_{h}$)
will be given as,
\[
S_h = f(data_h, data_o, data_g)
\]  
where \( f(data_h, data_o, data_g) \) is the function to be performed on the health data set \( data_h \), context data set \( data_o \) and geolocation information \( data_g \). In this functional model, each of the collected health parameter value is compared with respect to its normal range based on the context data and geolocation information. If the health parameter value falls outside the normal range based on the context data set and geolocation information, then the predicted output is “health status is abnormal based on parameters”. In that case an alert message is sent to the user along with the information of nearby health centre.

B. Probabilistic modeling of road networks and traffic information

A road network can be easily modeled as a Markov chain, where each state of the Markov chain corresponds to junctions (nodes) and a transition edge corresponds to connecting roads (edges) between a pair of junctions. This interpretation of the road networks is defined as primal [14]. On the other hand if we reverse the role of streets and junctions then that representation is the dual representation (streets are states and junctions are transition edges). CLAWER proposes a data-driven Markov model considering quantitative variables which helps in defining the dynamic nature of road traffic.

To handle the dynamic nature of the road network, we decompose the network to avoid the redundant computations. For example, if one of the lane of Kolkata region is blocked due to construction, it will not affect the transportation of Delhi region at any cost. But, blockage of National Highway will surely affect the transport network of Delhi. Hence, hierarchical decomposition or partition of road network based on spatial distribution and connection between two spatially distributed nodes is another important task. We propose to maintain a log of any change occur in road network - like blockage of the road or any inclusion of road, change in parameters. We also maintain a cache to store few computed routes to avoid the duplicate computation. Let us say, cache is updated in time \( t_i \) and road structure change is detected or communicated at time \( t_j \). If \( t_j < t_i \), then updated information need to be reflected in the system, otherwise we ignore the change.

To predict the less congested path, we utilize Markov chain based approach. The probability of the variable is represented as:
\[
P(X_{k+1} = S_{k+1} | X_k = S_k, X_{k-1} = S_{k-1}, \ldots, X_0 = S_0) = P(X_{k+1} = S_{k+1} | X_k = S_k)
\]
(2)
\[
\sum_{j=1}^{N} P_{i,j} = 1, \forall i = 1, 2, \ldots, N
\]
(3)
Markov chains are generally expressed by transition probability matrices. So if a system has \( N \) states, the transition probability matrix \( P \) will be an \( N \times N \) matrix with entries \( P_{i,j} \), where \( P_{i,j} \) is the probability of making a transition from state \( S_i \) to state \( S_j \) and the sum of all entries of every row in \( P \) sum to 1 as shown in Equation 3.

The transition probability matrix \( P \) corresponds to a digraph, where the nodes of the graph are represented by the states of the Markov Chain, and for every \( P_{i,j} \neq 0 \) there is an edge between states \( S_i \) and \( S_j \).

C. Construction of Markov chain of road network

We have shown the procedure to construct the transition matrix of Markov chain for a road network in this section. In order to thoroughly assess the performance of the proposed approach, we have considered a simple road network as a benchmark as shown in Fig. 3. However, since the method is scalable, it can be applied to any real world road network of any size. The network shown in Fig. 3 is of two towns that are separated by a river and connected by bridges.

Transforming a road network to Markov chain is done by converting the primal to dual. As defined earlier in a dual network the nodes correspond to city roads and edges are junction points. Fig. 4 shows the dual network where each node is labeled as XY which implies that junctions X and Y connected by road XY(X and Y being nodes in the primal). We perform this transformation of primal to dual network because more information is captured by the dual version.

Apart from traffic flow, we need to consider travel times between different junctions. We construct the Markov chain of the dual network, where each node of the dual network corresponds to a state and each edge corresponds to a transition edge in the Markov chain. The time to cover individual roads varies according to the length and width of the road, limitation of speed, road surface condition, time of the day and other dynamic factors. The average travel time between junction points depends not only on the length of the roads but also on other traffic conditions. The cloud server can compute the average travel time between two junctions by mining historical mobility data. After computing the normalized average travel time across all roads, the probability associated with every self loop is given by
\[
P_{ii} = \frac{(att_{i-1})}{att_i}, i = 1, 2, \ldots, n
\]
(4)
where \( att_i \) = average travel time of the \( i^{th} \) road, estimated from the collected data as discussed in the later section.

\[
P_{ij} = (1-P_{ii}), r_{ij}, i \neq j
\]
(5)
where \( r_{ij} \) = turning probability from road \( i \) to \( j \).

It may be noted that we have modeled the road network in this way because the underlying Markov model captures useful information about the real world phenomena and it is more simpler representation for performing analysis.
D. Predicting optimal path

It may be noted that the shortest path may not be the optimal path to reach the destination due to usual high traffic congestion or any other events such as traffic blockage or accidents. While Dijkstra algorithm is the traditional shortest path algorithm, for goal-directed path prediction, we have used a variant of A* Algorithm. The cost function $cost(n_i, n_j)$ is represented by:

$$cost(n_i, n_j) = a(n_i \rightarrow n_j) + b(R_{i,j}, c)$$  \hspace{1cm} (6)

where $cost(n_i, n_j)$ is the cost required to traverse from $n_i$ node to $n_j$ node on the road network ($R$). The components of the cost function are $a(n_i \rightarrow n_j)$ and $b(R_{i,j}, c)$. The first component is $a(n_i \rightarrow n_j)$, which is the cost of the road network, i.e., length of edge, average time and required to traverse. The next component consists of the subgraph ($R_{i,j}$) of the road network consisting the start ($n_i$) and destination node ($n_j$) and context parameter ($c$). CLAWER considers travel patterns ($p_1$) of the subgraph and real-time traffic events ($p_2$) as context while computing the cost function. For instance, the road-segments connecting residential regions and commercial regions usually have high traffic congestion ($p1$) in particular time-slots (0830-0930). Again, real-time traffic events ($p2$) such as accident or road-blockage impact the usual traffic flow - which needs to be incorporated in the cost function. As A* achieves better performance by using heuristics to guide its search, it is a major challenge to define heuristics in practical travel-routing and path-prediction system. The travel pattern is analyzed using auto-regressive integrated moving average [15] to find out probable values of GPS footprints of vehicles in different time-slots from the historical movement traces. The number of vehicles entering and exiting from a road segment ($e_a$) in a particular day $d_1$ is defined by $En_{d_1}^{e_a}$ and $Ex_{d_1}^{e_a}$, respectively. Similarly, the number of moving vehicles are represented by $M_{d_1}^{e_a}$.

$$En_{d_1}^{e_a} = (En_{d_1}^{1a}, En_{d_1}^{2a}, ..., En_{d_1}^{La})$$
$$Ex_{d_1}^{e_a} = (Ex_{d_1}^{1a}, Ex_{d_1}^{2a}, ..., Ex_{d_1}^{La})$$
$$M_{d_1}^{e_a} = (M_{d_1}^{1a}, M_{d_1}^{2a}, ..., M_{d_1}^{La})$$

(7)

Next, for each segment 3 matrices are formed for all $E$ edges of $p$ days.

$$M = (M_{a}, M_{b}, ..., M_{E})^t = \begin{pmatrix}
M_{d_1}^{e_a} & M_{d_2}^{e_a} & ... & M_{d_1}^{e_a} \\
M_{d_2}^{e_a} & M_{d_2}^{e_a} & ... & M_{d_2}^{e_a} \\
... & ... & ... & ...
\end{pmatrix}^t$$

(8)

Next, we execute the path finding algorithm in a MapReduce platform for a faster response. Algorithm 1 shows the basic steps of the procedure. RGraph is represented in adjacency list format. The Mapper function, generates the key-value pair for all the nodes present in the path.

$$\forall m \in \text{adjacency list} : \text{emit}(m, F_{Cost} + d)$$

(9)

$F_{Cost}$ produces the cost to reach one node to another based on other pre-defined parameters and heuristics given. Another key problem here is to preserve the original graph.
structure. Hence, mapper emits \((n, \text{adjacency list})\) also for tracking the road graph structure. All the reachable nodes are grouped by the Sort or Shuffle function. Reducer selects the path with minimum distance for each reachable node. Also, additional information are recorded to keep track of actual path. Each MapReduce iteration advances to extract the optimal path by one hop. Multiple iterations explore the whole graph. Basically, more reachable nodes are included with subsequent iterations as the search expands.

Algorithm 1: Path Finding using MapReduce platform

```
Function Mapper(NODEID n, NODE N):
1: for NODEID S ∈ AdjacencyList(N) do
2:   d ← ComputeDist(S, n)
3:   EMIT(NODEID n, F_{cost} + d)
4: end for
5: Function Reduce(NODEID M, [d_1, d_2, ...]):
6:   d_{min} ← ∞
7:   N_{neighbour} ← φ
8:   for d ∈ [d_1, d_2, ...] do
9:     if ISANODE(d) then
10:        M ← d
11:     else d < d_{min} then
12:        d_{min} ← d
13:     end if
14:     M.DISTANCE_Computed ← d_{min}
15:     Combine(M, d)
16: end for
```

E. Power consumption of CLAWER

To depict the performance of CLAWER in terms of power consumption and delay, we have theoretically modelled the interconnection of cloud-fog-edge-IoT devices [16] [17]. The delay in health-data transmission using SCceNB in indoor-region is calculated as:

\[
D_{Tran} = (1 + Up_{in}) \times (Damt_{up}/R_{upi}) + (1 + Dw_{in}) \times (Damt_{dw}/R_{dw})
\]

(10)

where \(Damt_{up}\) and \(Damt_{dw}\) are the amount of data (health/ location/ time) transmitted in uplink and downlink respectively in indoor region. The data transmission rates are represented by \(R_{upi}\) and \(R_{dw}\). The failure rates of data transmission are \(Up_{f_{in}}\) and \(Dw_{f_{in}}\) respectively for uplink and downlink. The delay for preliminary processing of health-status in SCceNB is

\[
D_{pn} = D_{p}/Pr_{fog}
\]

(11)

where \(D_{pn}\) and \(Pr_{fog}\) are the amount of health-data and the speed of fog devices (SCceNB). Similarly, the energy consumption of the IoT (mobile device/ sensors) during transmission is represented as:

\[
E_{in} = Ein_{up} \times ((1 + Up_{f_{in}}) \times (Damt_{up}/R_{upi})) + Ein_{dw} \times ((1 + Dw_{f_{in}}) \times (Damt_{dw}/R_{dw}))
\]

\[
E_{p_{in}} = E_{id} \times (D_{p}/Pr_{fog})
\]

(12)

The energy consumption of the IoT device during health-data processing is represented by:

where \(E_{id}\) is the energy consumed by the IoT device (mobile/ sensor) in idle mode. Similarly, the delay and power consumption in the outdoor region is computed following the same formulations.

IV. PERFORMANCE EVALUATIONS

To validate the proposed approach, a test-bed has been built which consists of handheld-devices, wearable sensors (collects health and movement data), fog device and cloud server. We have used the Google Cloud Platform (GCP) for the computation and storage of huge amount of movement and health log.

A. Experimental testbed

The datasets have been collected from individuals of different ages at Kharagpur, India region voluntarily. A total of 65 subjects participated in the survey for six months. The subjects were requested to install the Android application and carry the wearable devices for the study period. Fitbit is used for collecting health data. We have created an Android application which is capable to communicate with the wearable device, on-board sensors of the handheld Android device and bluetooth signals from beacons. In our experiment, the Raspberry Pi 3 is used as Bluetooth beacon, where we install the Eddystone Bluetooth Beacon for sending data periodically. In this regard it has to be mentioned that the data comes in the form of .csv file. In our lab we have used a Raspberry Pi as fog device. The Android application installed inside the smart phone accumulates the movement and health data, and forwards to the fog device. The Raspberry Pi acting as the fog device sends the data to the cloud. Google IoT core instance is used as cloud server in our experiment. The cloud after processing of the health and movement data, predicts the health status and generate health recommendations for the user. The health-related recommendations are sent to the Raspberry Pi, which stores them and sends to the user according to his/her requirement. In case of abnormal health status and emergency, an alert is provided to the user and subsequent procedures are initiated.

B. Experimental observations

The experimental results are illustrated in two broad aspects: (i) power consumption and delay and (ii) accuracy of path prediction.

1https://developers.google.com/beacons/eddystone
For measuring delay and power consumption of the framework, we have taken different samples of data from the experiments. The results are shown in Figures 5 - 8. The performance of CLAWER is compared with existing Health-Fog system [8]. It is observed that our proposed framework has outperformed the existing approach in a large margin. Figures 5 and 6 illustrate the delay and energy consumption indoor region. The fog devices (SCceNB) are used to analyse preliminary health-status and communicate with cloud in case abnormal health condition is detected. SCceNB stores the health-profile along with medicine recommendations of patients and thus reduces the delay. It is observed that CLAWER has around 28% better response time and the energy consumption is reduced by 26%. Figures 7 and figure 8 represents the delay and energy consumption in outdoor region, where RSUs are used as fog devices. CLAWER utilizes markov-predictor for selecting the best route to reach the destination in minimum delay. The results show promising reduction of delay and energy consumption compared to the existing (cloud-only solution) approach.

The performance of CLAWER framework for path prediction from source to destination avoiding traffic congestion is illustrated using Precision, Recall and Accuracy metrics [16]. To demonstrate the effectiveness, we have segregated the data into different time-bins such as 10mins, 20mins upto 60mins. For each of the time-bins, we have computed the metrics. Figure 9 shows the precision and recall values...
for several time-bins required to reach the destination from source. It is observed that CLAWER has achieved high precision and recall values in the range of $0.83 - 0.93$ and $0.84 - 0.956$ respectively. Figure 10 illustrates the accuracy measure of the path prediction model and we observe greater than $80\%$ accuracy with all time-bins. While these values represent high accuracy measures, along with increasing time-bins, our CLAWER method maintains steady performance. These metrics show encouraging results and depict the overall effectiveness of CLAWER framework.

V. CONCLUSIONS

This paper has proposed a context-aware workflow framework named CLAWER for health management. The environmental context, health data, user movement data are collected using IoT devices, which are assimilating data inside the smart phone (edge), and sent to the cloud through a fog device. For the indoor users SCceNB is used as fog device, which performs preliminary processing on the data received from the smart phone. Then the SCceNB forwards the data to the cloud. The data analysis is performed by the cloud and health recommendations are generated. For abnormal health status, an alert is sent to user’s mobile device for emergency care. CLAWER implements an end-to-end framework where several agencies/partners work together sequentially to provide better service. The experimental analysis show that the proposed framework has better precision, recall and accuracy as well as power consumption than the existing schemes. While CLAWER is a generic framework, in future, we will extend it for other domains such as defence application or agricultural-IoT. The proposed CLAWER framework will act as a foundation of bridging the gap between process management and IoT data analytics.

ACKNOWLEDGEMENT

The work is partially supported by TCS PhD research fellowship.

REFERENCES


