Engineering Applications of Artificial Intelligence 87 (2020) 103285

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

ElHealth: Using Internet of Things and data prediction for elastic management of human resources in smart hospitals☆

Gabriel Souto Fischer [a,*], Rodrigo da Rosa Righi [a], Gabriel de Oliveira Ramos [a], Cristiano André da Costa [a], Joel J.P.C. Rodrigues [b,c]

[a] Applied Computing Graduate Program, Universidade do Vale do Rio dos Sinos, Unisinos, São Leopoldo, Brazil
[b] Federal University of Piauí, Teresina, PI, Brazil
[c] Instituto de Telecomunicações, Portugal

ARTICLE INFO

Keywords:
Internet of Things
Health
Smart hospitals
Data prediction
Human resources
Elasticity

ABSTRACT

Hospitals play an important role towards ensuring proper health treatment to human beings. One of the major challenges faced in this context refers to the increasingly overcrowded patients queues, which contribute to a potential deterioration of patients health conditions. One of the reasons of such an inefficiency is a poor allocation of health professionals. In particular, such allocation process is usually unable to properly adapt to unexpected changes in the patients demand. As a consequence, it is frequently the case where underused rooms have idle professionals whilst overused rooms have less professionals than necessary. Previous works addressed this issue by analyzing the evolution of supply (doctors) and demand (patients) so as to better adjust one to the other, though none of them focused on proposing effective counter-measures to mitigate poor allocations. In this paper, we build upon the concept of smart hospitals and introduce *elastic allocation of human resources in healthcare environments* (ElHealth), an IoT-focused model able to monitor patients usage of hospital rooms and to adapt the allocation of health professionals to these rooms so as to meet patients needs. ElHealth employs data prediction techniques to anticipate when the demand of a given room will exceeds its capacity, and to propose actions to allocate health professionals accordingly. We also introduce the concept of *multi-level predictive elasticity of human resources* (which is an extension of the concept of *resource elasticity*, from cloud computing) to manage the use of human resources at different levels of a healthcare environment. Furthermore, we devise the concept of *proactive human resources elastic speedup* (which is an extension of the *speedup* concept, from parallel computing) to properly measure the gain of healthcare time with dynamic parallel use of human resources within hospital environments. ElHealth was thoroughly evaluated based on simulations of a hospital environment using data from a Brazilian polyclinic, and obtained promising results, decreasing the waiting time by up to 96.71%.

1. Introduction

Internet of Things (IoT) is a concept where physical, digital, and virtual objects (i.e., things) are connected through a network structure and are part of the Internet activities in order to exchange information about themselves and about objects and things around them (Singh and Kapoor, 2017). IoT enables devices to interact not only with each other but also with services and people on a global scale (Akeju et al., 2018). The development of this paradigm is in constant growth due to the continuous efforts of the research community and due to its usefulness to a wide range of domains, such as airports, military, and healthcare (Singh and Kapoor, 2017; Sarhan, 2018).

A particularly relevant scenario for IoT is healthcare (da Costa et al., 2018). According to Pinto et al. (2017), IoT promises to revolutionize healthcare applications by promoting more personalized, preventive, and collaborative ways of caring for patients. In particular, IoT-assisted patients can be supervised uninterruptedly using wearable devices, thus allowing risky situations to be detected and appropriately treated right away (Darshan and Anandakumar, 2015; Srinivas et al., 2018). Moreover, IoT provides a means for health systems to extract and analyze data, which can then be combined with machine learning techniques to early detect health disorders (Singh, 2018; Moreira et al., 2019).

☆ No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.engappai.2019.103285.

* Correspondence to: Universidade do Vale do Rio dos Sinos. Av. Unisinos, 950, CEP 93022-750, São Leopoldo, RS, Brazil.
E-mail addresses: gabriel.souto.fischer@gmail.com (G.S. Fischer), rrrighi@unisinos.br (R.d.R. Righi), gdoramos@unisinos.br (G.d.O. Ramos), cac@unisinos.br (C.A.d. Costa), joeljr@ieee.org (J.J.P.C. Rodrigues).

https://doi.org/10.1016/j.engappai.2019.103285
Received 22 May 2019; Received in revised form 17 August 2019; Accepted 3 October 2019
Available online 11 October 2019

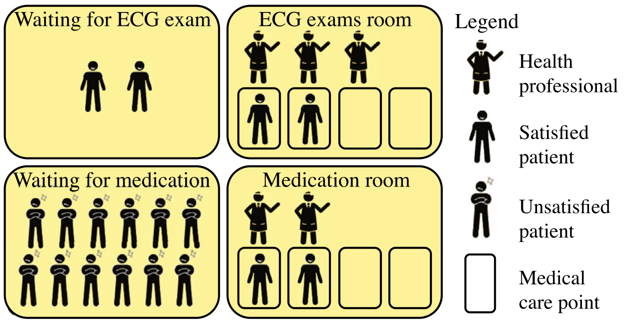

Fig. 1. Example scenario where are more attendants examining than medicating patients, even though the number of patients waiting for exams is considerably smaller than those waiting to receive some medication, generating dissatisfaction for patients awaiting medication.

Hospitals are among the most important service points capable of ensuring appropriate treatment to the population. Considering the importance of such environments, enhancing the efficiency with which a hospital's resources and processes are controlled becomes a central concern. Such a concern is particularly relevant in the context of underdeveloped countries, where the high number of patients associated with the lack of resources leads to overly high waiting times (Graham et al., 2018). In this sense, should be possible to identify *when* health centers would be overloaded, it would allow to establish contingency plans to minimize (or perhaps even eliminate) these bottlenecks.

As stated in Butean et al. (2015), no matter how easy or complicated a situation is, if the medical staff does not react in time, everything regarding patients' health becomes doubtful and unsafe. Hence, health professionals play a major role towards patients' well-being (Nierop-van Baalen et al., 2019). These professionals range from nurses (who carry out triage procedures and small treatments) to doctors (who attend the most diverse medical specialties). In this kind of scenario, a static allocation of health professionals to health sectors may be inefficient, since some professionals may be misallocated to low demanding sectors, while leading to a lack of professionals in high demanding sectors. Such a problem is illustrated in Fig. 1, where the set of available attendants are statically assigned to two service sectors, one for exams and another for medication. In the example, there are more attendants examining than medicating patients, even though the number of patients waiting for exams is considerably smaller than those waiting to receive some medication. In this context, the idle attendants could be moved from the low demanding room to the high demanding one. In fact, the allocation of attendants should always *adapt* to the current conditions of the health sectors.

Data prediction techniques play a role in this kind of scenario. In particular, such techniques make it possible to anticipate patients demand and to prepare the allocation of medical staff accordingly. Several works in the literature have proposed the use of data prediction techniques for optimizing resources usage: to minimize bottlenecks in patients flow (Vieira and Hollmén, 2016), to predict patients arrival at emergency (Graham et al., 2018), and to plan training sessions for doctors based on patients demand (Ishikawa et al., 2017). Although these works do predict demand and the use of health resources, none of them is able to provide counter-measures to mitigate the problems of poor resources allocation. Therefore, to the best of our knowledge, this is the first work to propose an efficient allocation of health professionals based on predictions made out of patients demand as obtained by IoT sensors. We say that our approach is *efficient* in a sense that it is able to properly meet patients demand using as least health professionals as possible.

This work introduces a model of *elastic allocation of human resources in healthcare environments* (ElHealth, for short) as an alternative to the traditional static allocation of medical staff. ElHealth works by adjusting the medical staff allocation of smart hospitals (equipped with IoT sensors) based on predictions of patients demand. In particular, ElHealth uses IoT sensors to keep track of patients demand, which is modeled as a time series and is used to predict future demands. Such predictions allow to identify situations where the staff availability is unlikely to meet the demand. Building upon such predictions, ElHealth proposes an efficient allocation of the medical staff by moving such professionals to the most demanding areas while taking their time constraints into account. The idea is to always offer a reasonable waiting time for patients, regardless of the workload (number of patients in the hospital room). Thus, the main scientific contributions of this paper are twofold:

(i) We devise *Multi-level Predictive Elasticity of Human Resources*, which includes an algorithm for automatic management of human resources in healthcare environments, making use of data prediction techniques, and some evaluation metrics for smart, IoT-enabled hospitals;
(ii) We introduce the concept of *Proactive Human Resources Elastic Speedup* to identify the decrease in medical care time with a dynamic, parallel use of human resources for care in a hospital environment.

The rest of this paper is organized as follows. Section 2 presents the work related to the study. Section 3 presents ElHealth as well as the concepts of Multi-level Predictive Elasticity of Human Resources and of Proactive Human Resources Elastic Speedup. Section 4 details the methodology of evaluation of the model and Section 5 presents an evaluation performed with the developed implementation, as well as the results found. Finally, Section 6 presents the conclusions and future work directions.

2. Related work

A number of approaches focused on optimizing the flow of patients to properly allocate health resources. Vieira and Hollmén (2016) investigated ways of minimizing bottlenecks in the flow of patients due to appointments, visits, usage of resources, etc. The objective was to improve patients' satisfaction and maximize hospital's profit. To this end, using data from a Finnish hospital, the authors used k-Nearest Neighbors (Fix and Hodges, 1951; Cover and Hart, 1967) and Random Forests (Ho, 1995) to predict such a flow. In the same line, Graham et al. (2018) aimed at predicting the arrival of patients in the emergency department of a hospital to properly prepare allocate medical staff. To accomplish such a task, the authors used logistic regression (Cox, 1958), decision trees (Breiman et al., 1984), and gradient boosted machines (Friedman, 2001) with data from a British hospital. In both cases, however, the objective was exclusively on identifying specific patterns on data, but not on proposing counter-measures to improve the allocation of health resources.

In an attempt to increase health coverage, some studies proposed forecasting models to understand the evolution of doctors supply and patients demand so as to better adjust one to the other. Ishikawa et al. (2017) concentrated on training enough physicians to meed the patients demand in Japan until 2030. Liu et al. (2017) focused on a similar problem, but from a global perspective. In contrast to our work, nonetheless, the adaptation of hospital's resources to the patients flow was left aside these works.

In order to recommend drugs and diets to patients, Ali et al. (2018) developed a recommender system based on Fuzzy ontologies (Cross, 2014) and Type-2 Fuzzy logic (Zadeh, 1975). The proposed system keeps track of patients' health conditions using wearable sensors, and then suggests drugs and diets prescriptions so as to specifically treat diabetes. However, their approach had a treatment-oriented perspective, as opposed to our resources-optimization-based perspective.

Several works proposed the use of IoT sensors to keep track of patients flow in hospital environments. Orimaye et al. (2015) focused on identifying humans presence by analyzing their interference on sensors readings. To this end, a variation of *support vector machines* (Cortes and Vapnik, 1995) was used. Zamanifar et al. (2017), on the other hand, investigated human movement to allow sensors to be put for sleep, thus making a more efficient use of the available resources. To this end, they employed the concept of distributed movement prediction, and developed a customized second-order hidden Markov model (Kundu et al., 1989). In both cases, however, the focus was mainly on improving the efficiency of the IoT sensors rather than of health resources themselves.

2.1. Comparison and research opportunities

Table 1 presents a comparison of the collected papers, presenting some of their main characteristics, and pointing out some of their gaps. Based on the selected articles, we can identify that several works concentrate on predicting issues related to the healthcare area using the Internet of Things and Data Prediction. In particular, it was possible to see that it is not only possible to use the technology for this, as it is already being used in several approaches in the scientific community. Most of the IoT systems with Data Prediction applied to healthcare researched focus on the monitoring the patient's health conditions in order to generate alerts if any risk situations are identified. These systems are able to predict when the patient's vital signs will be at risk, identify heart problems, treatment efficacy, and environmental risk situations for patients. When we have Data Prediction with the problem of the lack of resources in hospital environments, the articles found just focus on predicting the future demand of patients or the future quantity of available doctors, not proposing solutions to the problem, leaving others in charge of decision-making. In this context, we can enumerate some of the main gaps of the area as follows:

- Even though these models integrate data captured by IoT sensors with Data Prediction techniques, they neither analyze the *use of resources* in hospital environments nor the *overload of patients* in certain places.
- Although several models are capable of identifying future demand in a hospital environment, these models lack *concrete solutions* to help solving the problem of deficiency of hospital resources.
- The lack of ability to *optimize processes* in an automated way.

The lack of sufficient human resources in healthcare environment is not new and, based on studied works, we can see that this problem will remain at least in the near future. Hence, finding ways of optimizing the use of existing resources and adjust hospitals' capacity to meet patients demand are challenges that can make all the difference. The use of Data Prediction and Internet of Things contributes towards future solutions or automation of processes in the health area. But the potential of the technologies is being underused since it is possible to propose solutions such as optimization and better use of existing human resources.

3. ElHealth model

Based on the current state-of-the-art and the gaps discussed in the related works, we can state not only that it is possible to control the health status of patients in hospital environments through the use of the Internet of Things, but also that the people location within any environment could be tracked using the same concept. However, most of the approaches we have seen concentrate only on identifying the location and current/future health status of patients, neglecting the potential benefits that efficient health resources allocation could bring to the patients (Orimaye et al., 2015; Zamanifar et al., 2017). As we have seen before, one of the major challenges faced in hospital environments refers to the large waiting queues. Moreover, considering that doctors reaction time plays a role on patients recovery, long waiting times may compromise any guarantees about the patient's future health.

Based on this background we introduce ElHealth, a multilevel predictive model for efficient allocation of human resources based on expected flow of patients within hospital environments. In particular, ElHealth adapts the concept of elasticity in cloud computing to the context of human resources, adjusting the hospital's attendance capacity to the expected demand of patients, where professionals are allocated, de-allocated and reallocated according to the expected hospital needs. ElHealth groups information from several sources: patients arrivals and needs (from a real hospital dataset), patients movement (using IoT sensors spread over the hospital environment), and medical staff availability (from a dataset). Using these data, we employ a time-series prediction algorithm (which we discuss in Section 3.3.1) to anticipate the future demand of patients. This information is then useful for applying the concept of elasticity-based allocation of resources. Based on that model, ElHealth computes an efficient allocation of hospital resources (medical staff and equipment), which contributes towards minimizing patients waiting queues.

The next subsections detail our model, bringing the main design decisions (Section 3.1), the proposed architecture (Section 3.2), and the definition of Multi-level Predictive Elasticity of Human Resources at room-level (Section 3.3.1) and hospital-Level (Section 3.3.2).

3.1. Design decisions

Our model is based on the premise that there are sensors scattered around the hospital, which can identify patients who pass through them. Firstly, they must be in all the entrances and exits, so that whenever a patient enters or leaves the hospital it is possible to identify him. To detect the movement and location of patients, we assume the presence of sensors at the doors of all hospital rooms. Each patient must have a *Patient Identification Wristband* linked in the system and must carry it through all time in the hospital's internal environment. The attendant responsible for the reception of patients should be able to perform the linking of a wristband to a given patient as soon as the patient is admitted in the hospital. Thus it is possible to identify when and where a given patient is as soon as he enters at the healthcare environment, along with the time he remains in each room while being attended to. In addition, each healthcare professional must have a tag linked to him in the system and must carry it with him throughout his active period in the hospital. Thus, all available attendants can also be located inside the hospital in the same way as patients.

We use a Real-Time Location System (RTLS) (Boulos and Berry, 2012) with room-level localization accuracy. According to Boulos and Berry (2012), RTLS are systems for identifying and tracking location of assets and/or people in real time or near real time. The choice of an RTLS is based on its ability to allow automatic identification, avoiding the existence of a human error in identification processes. ElHealth should be transparent to patients, in the sense that it does not need to report any conditions related to its movement through the hospital environment, being an activity performed automatically by the system.

ElHealth model adapts the predictive elasticity strategy using upper and lower thresholds for the context of people, based on predicting patient demand, as will be discussed in detail in Section 3.3. Fig. 2 demonstrates the use of thresholds where ElHealth forecasts that the upper threshold will be reached (meaning that human resources should be reallocated to fulfill that needs) and soon after ElHealth forecasts that the lower threshold will be reached (meaning that human resources could be released to other sectors). For this process, ElHealth should be able to alert people to allocate. However, the final decision should always be made by the health professional or hospital manager.

With respect to the data prediction strategy, ElHealth uses a statistical-based approach through an implementation of the ARIMA model (Box and Jenkins, 1970). According to Nisha and Sreekumar (2017), ARIMA model uses historical information to predict future patterns. ARIMA represents a general class of models for forecasting time series (Nisha and Sreekumar, 2017). In addition, we also opted by using ARIMA because it presents competitive results in terms of prediction accuracy when compared to Neural Networks and Random

Table 1
Related work comparison.

Work	Focus	Proposed solution	Support for internet of things	Data prediction model	Human resources deficiency
Orimaye et al. (2015)	Use of sensors to collect data from patients	Enable sensors to perform non-invasively health diagnostics	Uses sensors to identify patients' location	Uses a Support Vector Machine (SVM) to predict future patients' location	Does not address this problem, but help indirectly, proposing automatic diagnosis
Vieira and Hollmén (2016)	Deficiency of resources to perform patients' care	Identify the resources needed to ensure the patient's care flow	Not applicable	Uses Nearest Neighbors and Random Forest to predict future resources usage	Does not propose concrete solutions, only provides data to support decision-making
Ishikawa et al. (2017)	Deficiency of doctors for current patients' care demand in Japan (Hokkaido)	Identify health doctors distribution and sufficiency to propose ways for guarantee care for demand	Not applicable	Uses System Dynamics (SD) and Geographic Information System (GIS) to predict distribution and sufficiency of doctors	Proposes a plan for training physicians that considers geographic requirements
Liu et al. (2017)	Deficiency of doctors for current patients' care demand in global scale	Identify health doctors distribution and sufficiency until 2030 in order to compare with demand projections	Not applicable	Uses an economic model and a Generalized Linear Model to predict distribution and sufficiency of health professionals, and patients' demand	Does not propose concrete solutions, provides data to show the problem escalation, to support solutions proposal by the international community
Zamanifar et al. (2017)	Power usage, overloads and packet losses in data transfer between health sensors	Prepare sensors to collect data from patients, saving energy by keeping off unused sensors	Uses sensors to identify patients' location	Uses a Second-order Hidden Markov model to predict future patients' location	Does not address this problem
Ali et al. (2018)	Need for long-term medical care for patients with chronic diseases	A drugs and food recommendation system based in health status of patients	Uses wearable sensors to collect patients' health status and conditions	Uses a Type-2 Fuzzy logic and ontologies to predict patients' future health status	Does not address this problem, but help indirectly, automatizing care
Graham et al. (2018)	Crowding within emergency departments and the significant negative consequences for patients	Use of data mining using machine learning techniques to predict admissions in a hospital	Not applicable	Uses logistic regression, decision trees and Gradient Boosted Machines to predicts patients' arrival in emergency	Does not propose concrete solutions, provides data to support decision-making of hospital managers

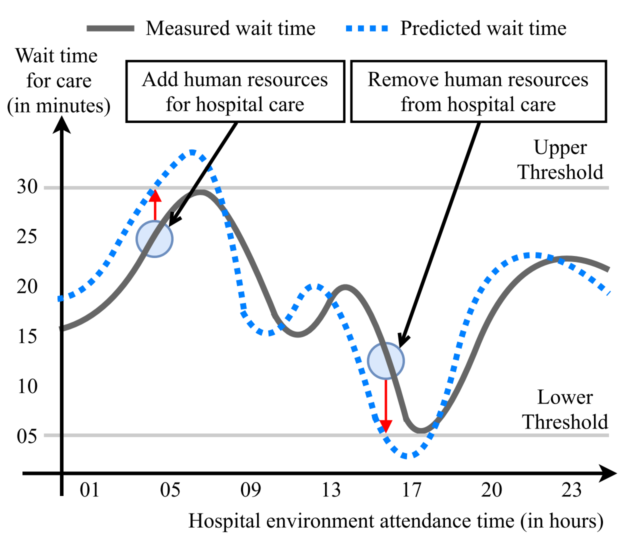

Fig. 2. Predictive elasticity based on number of patients adopted by ElHealth.

Forest engines, but outperforms these two concurrent when analyzing the computational complexity to run the forecasting algorithm (dos Santos et al., 2019). Since the number of patients waiting for care over time can be described as a time series, we chose to use the approach through ARIMA because it is a very flexible mathematical model, with a good predictive performance of time series when compared with other approaches (Nisha and Sreekumar, 2017). ARIMA models are extremely useful in predicting different sectorial series, since they can represent stationary series, and also non-stationary series. We use a non-stationary model based on seasonality in the demand for medical staff, since accidents, epidemics, holidays, and other events, can alter patients' demand for care.

3.2. Architecture

ElHealth architecture model three services: (i) a Web service, responsible for visualization layer, and ElHealth Web Pages interface; (ii) an inference service, responsible for data processing, movement records handling, patients demand prediction, and human resources allocation decisions; and (iii) a database service. These three services are part of our proposed ElHealth Service. Fig. 3 presents the components and the network view in the proposed model.

ElHealth model is subdivided into five modules responsible for information handling from its capture by sensors to the final result displayed in the Web application. Each module has a specific function, having an input information and a specific output result that can be used as input from other modules. Fig. 4 presents the proposed modules, detailing the architecture of the model.

ElHealth_Capture receives and pre-process data captured by sensors scattered around the hospital and sends to *ElHealth_Formatter*, responsible for process data, and identify patients' movement through hospital environments and rooms. After, *ElHealth_Predict* identifies patients movement patterns through the hospital environment. Based on previously generated movement records, the path that patients travel during their movement through the hospital and the time spent in each environment are identified. Thus, this module identifies patterns related to the arrival of patients in these environments, and patterns related to the waiting time for care, using this information to predict future patients arrivals.

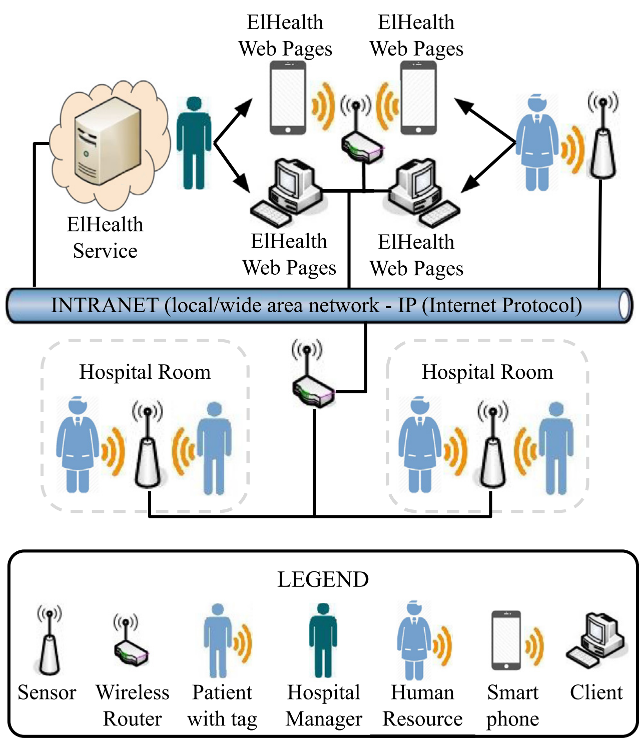

Fig. 3. Components and network view in ElHealth model with (i) ElHealth Web Pages client interface; (ii) ElHealth Service, for information processing and decision making; (iii) a RTLS, for track users' tags; and (iv) Hospital managers, patients, or human resources.

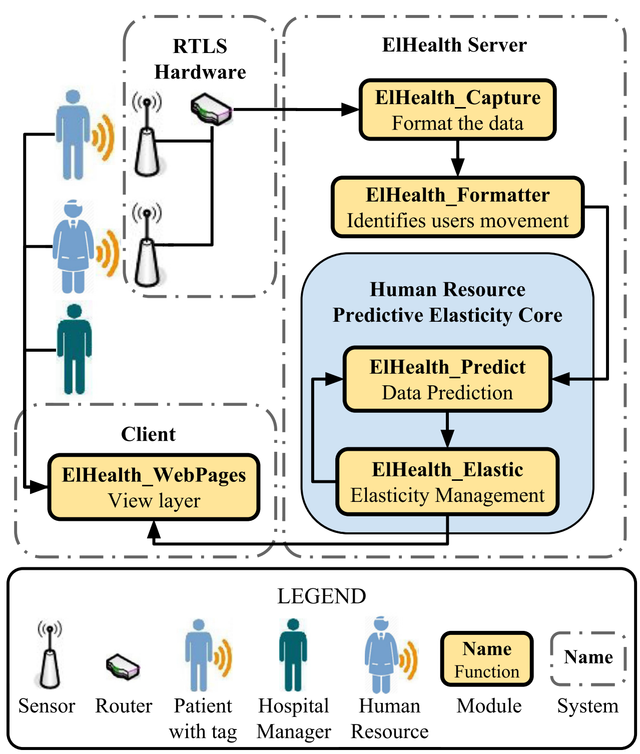

Fig. 4. ElHealth model architecture detail where the information flow starts in ElHealth_Capture module that receives users' movement records from RTLS sensors, and goes through different handlings over proposed modules, until the exhibition of elasticity notifications in ElHealth Web Pages.

ElHealth_Elastic manages system's elasticity. It verifies human resources allocation in each of hospital environments, check the current patients movement and compare with the predictions made by the previous module. Based on this comparison, this module generates an intelligent and automatic allocation of human resources to better meet future patient demand. We want to emphasize that the system generates notifications for human resources to reallocate, but effective reallocation depends on the people accomplishing what was indicated by the application. *ElHealth_Elastic* and *ElHealth_Predict* modules are the most important part and the core of our proposed model, since ElHealth_Elastic requests predictions from the ElHealth_Predict module to take elastic actions, performing resources analysis based on predictions performed by the previous module. In Section 3.3 will be detailed the algorithms and how the elastic management of the human resources in the hospital environment are performed. Finally, *ElHealth Web Pages* displays to human resources the elasticity notifications generated before.

3.3. Multi-level predictive elasticity of human resources

ElHealth introduces the concept of *Multi-level Predictive Elasticity of Human Resources* in healthcare environments, which can be defined as follows.

Definition 1 (*Multi-level Predictive Elasticity of Human Resources*)**.** Multi-level Predictive Elasticity of Human Resources is an extension of the concept of resource elasticity in Cloud Computing (Al-Dhuraibi et al., 2018) to manage the use of human resources at different levels of a healthcare environment, where human resources are allocated and de-allocated according to the expected demand of patients. The Multi-level Predictive Elasticity of Human Resources aims to generate plans of allocation of health professionals in hospital environments based on patients' demand, but always considering the quality of services currently offered by these healthcare environments.

ElHealth employs the term *elasticity* with a slightly different meaning from that used in cloud computing. Here, it refers to the system's ability to allocate/reallocate/de-allocate human resources capable of attending patients in order to adapt to varying patient demand in real time. In particular, in the context of this work, elasticity refers to:

- **Allocation**, which denotes the capacity of the system to request health professionals who are not in hospital to attend the current patients' demand;
- **Reallocation (or migration)**, which denotes the ability of the system to migrate professionals who are allocated to a particular hospital environment to some other environment where more professionals are needed;
- **De-allocation** which denotes the capacity of the system to release human resources no longer needed to attend the current patients' demand.

In order to perform allocation, de-allocation, and reallocation of human resources in an elastic way, ElHealth model makes use of a multilevel approach to predict the future demand of patients and the use of rooms in the hospital. Based on this approach, our model considers predictive elasticity differently at (i) the *room-level*, where our model must identify the future use of a given room, and check if the number of attendants is sufficient to meet patients' demand (as discussed next, in Section 3.3.1), and at (ii) the *hospital-level*, where ElHealth should verify if there are sufficient attendants to meet patients' demand from all rooms in the hospital environment, with attendants moving between rooms (as detailed forward in Section 3.3.2). A diagram of these two levels is presented in Fig. 5.

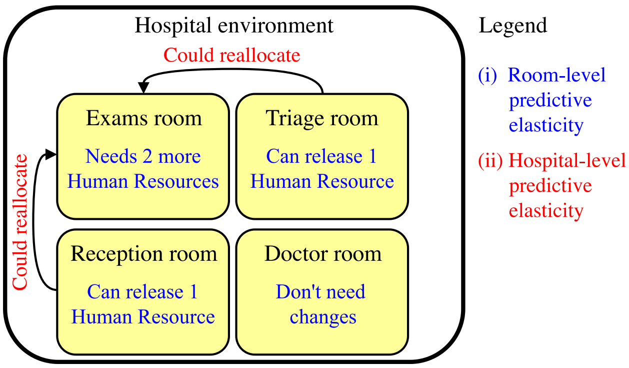

Fig. 5. Multi-level predictive elasticity of human resources example with (i) room-level predictive elasticity, and (ii) hospital-level predictive elasticity.

Table 2
Mathematical notation of ElHealth.

Nomenclature	Description
r	Hospital room
t_n	Specific n time instant
t_i	Initial time instant
t_f	Final time instant
a	Allocated attendants
f_i	Future initial time instant
f_f	Future final time instant
$CV(r,t_i,t_f)$	Care Vector
$size(x)$	Size of a x vector
$CDT(x[i])$	Care Duration Time
$ACT(r,t_i,t_f)$	Average Care Time
$NA(r,t_n)$	Number of Attendants
$ANA(r,t_i,t_f)$	Average Number of Attendants
$NWP(r,t_i)$	Number of Waiting Patients
$NIP(r,t_n)$	Number of Incoming Patients
$ENP(r,t_i,t_f)$	Estimated Number of Patients
$ECT(r,t_i,t_f)$	Estimated Care Time
$HRES(r,t_i,t_f)$	Human Resources Elastic Speedup
$PHRES(r,a,f_i,f_f)$	Proactive Human Resources Elastic Speedup

3.3.1. *Room-level predictive elasticity*

At the room-level, ElHealth needs to predict patients arrival rate at any room based on current and previous arrivals on that room. The prediction is made using the ARIMA model based on the average care time with the current attendants allocation, and the estimated waiting time for the care queue. When ElHealth identifies that the waiting time will become higher or lower than the threshold values set by hospital manager, ElHealth should compute the number of health resources required to meet patients' demand. To this end, ElHealth model introduces the concept of *Proactive Human Resources Elastic Speedup* in smart hospitals, which can be defined as follows.

Definition 2 (*Proactive Human Resources Elastic Speedup*)**.** Proactive Human Resources Elastic Speedup is an extension of the Speedup concept of parallel computing (Amdahl, 1967) to identify the gain of medical care time with the dynamic parallel use of human resources for care in a hospital environment. The Proactive Elastic Speedup uses a predictive approach to determine the future demand of patients and dynamically define the adequate number of attendants, identifying the gain of future medical care time in a hospital environment.

ElHealth proposes some mathematical formalism to estimate the Proactive Human Resources Elastic Speedup, which will be described in the sequence. Table 2 presents a summary of such mathematical notation.

Let $CV(r,t_i,t_f)$ denote the care vector of room r for the time interval between t_i and t_f. The size of any such vector is defined by $size(x)$. Using these two functions, the average care time in the hospital's room r between t_i and t_f times can be formulated as in Eq. (1), where $CDT(x[i])$ refers to a care duration time $x[i]$ that has already occurred in that room and $x[] = CV(r,t_i,t_f)$ is a care vector that occurred in that room.

$$ACT(r,t_i,t_f) = \frac{1}{size(x)} \sum_{i=0}^{size(x)-1} CDT(x[i]) \tag{1}$$

Eq. (1) results in a numerical value of time. An example would be any room r, between 1 and 5 times, where the result could be defined as: $ACT(r,1,5) = 15$ min. Using this equation, it is possible to estimate the average time of a care in a certain hospital room.

Due to the elasticity of human resources, at different time instants, there is a different number of attendants allocated to care in each of the hospital rooms. The average number of attendants in the hospital's room r between times t_i and t_f is defined as in Eq. (2), where $NA(r,t_n)$ refers to the number of attendants allocated to care in the room r at the instant of time n.

$$ANA(r,t_i,t_f) = \frac{1}{t_f - t_i} \sum_{t_n=t_i}^{t_f-1} NA(r,t_n) \tag{2}$$

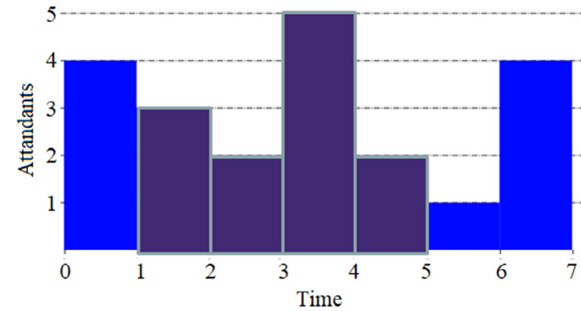

Fig. 6. Calculating ANA equation in a hospital room between times 1 and 5, where the number of attendants allocated are varying, reaching an average number of 3 attendants allocated between these times.

Fig. 6 presents an example of how ANA is computed. In the example, a given room r has a varying number of attendants between times 0 and 7 (e.g., three attendants at the time 1, two attendants at time 2, and so on). The average number of attendants can then be computed using Eq. (2) as follows: $ANA(r,1,5) = \frac{1}{5-1}\sum_{t_n=1}^{5-1} NA(r,t_n) = \frac{NA(r,1)+\cdots+NA(r,4)}{4} = \frac{3+2+5+2}{4} = 3$ attendants.

The same idea of the previous function is useful for patients' reality because in different moments of time there are different amounts of patients awaiting care in each of the hospital rooms. Thus, the estimated number of patients waiting for care in the hospital's room r between t_i and t_f times is defined by Eq. (3), where $NWP(r,t_i)$ refers to the number of waiting patients for care in a room r at t_i time instant, and $NIP(r,t_n)$ refers to the number of incoming patients in a room r at t_n time instant.

$$ENP(r,t_i,t_f) = NWP(r,t_i) + \sum_{t_n=t_i+1}^{t_f-1} NIP(r,t_n) \tag{3}$$

Using the equations previously proposed, our model calculates the estimated care time of all patients waiting, and estimates the time that a new incoming patient needs to wait to be attended. The $ECT(r,t_i,t_f)$ is defined by Eq. (4), where $ACT(r,t_i,t_f)$ refers to the average care time for room r between t_i and t_f times, and $ENP(r,t_i,t_f)$ refers to the estimated number of patients who are waiting for care in a room r between t_i and t_f instants.

$$ECT(r,t_i,t_f) = ACT(r,t_i,t_f) \cdot ENP(r,t_i,t_f) \tag{4}$$

An example would be the room r, between two times $t_i = 0$ and $t_f = 40$ that would result in an average number of 4 patients and an average

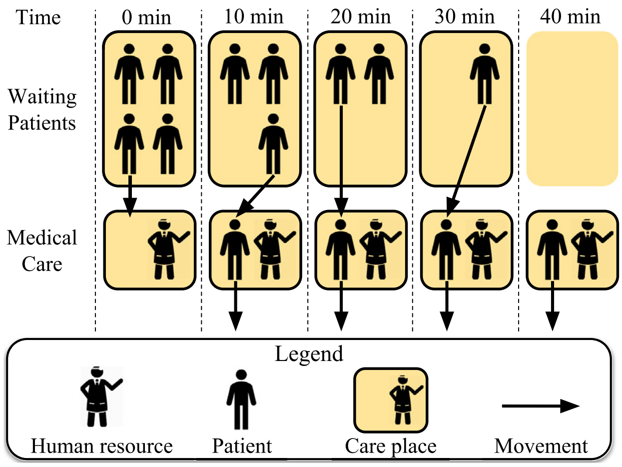

Fig. 7. Calculating *ECT* in a hospital room with 4 patients waiting, and average care time of 10 min, where an Estimated Care Time of 40 min.

care time of 10 min as shown in Fig. 7. In this hypothetical situation, at 0 min instant the first patient was called to the care. In 10 min instant, the first patient ends their care and goes away, so the second patient is designated to care, and so on, until instant 40 min, when the last patient is released. Thereby, all patients are attended within 40 min. Applying Eq. (4), we obtain $ECT(r,t_i,t_f) = ACT(r,t_i,t_f) \cdot ENP(r,t_i,t_f) = 10 \cdot 4 = 40$ min.

Knowing $ECT(r,t_i,t_f)$, we can analyze the average time for care of all patients waiting in the room r between t_i and t_f times. However, this value refers to a hospital room with a single attendant allocated for care, but in most cases will be more than one health professional working in that room, making it necessary to identify the average time with different numbers of attendants. In this context, ElHealth model uses a parallel allocation of human resources, such as the parallel allocation of virtual machines used in elastic systems (Al-Dhuraibi et al., 2018) or the use of parallel processors in high-performance computing (Rosa Righi et al., 2016). Thus, based on the Reactive Elastic Speedup proposed by Rosa Righi et al. (2016), ElHealth introduces Eq. (5) for *Human Resources Elastic Speedup*.

$$HRES(r,t_i,t_f) = \frac{ECT(r,t_i,t_f)}{ANA(r,t_i,t_f)} \tag{5}$$

Consider again the previous example (Fig. 7), with room r between two times t_i and t_f with an average number of 4 patients, an average care time of 10 min and with 2 health professionals allocated, as shown in Fig. 8. In this hypothetical situation, at 0 min time instant, there were 4 patients waiting and none in attendance by doctors, so the first two patients were called to care. In 10 min instant, the first two patients are released, and the last two patients are designated to care. Thus, at 20 min instant, the last two patients are released. Thereby, all patients are attended in only 20 min. Using Eq. (5), we obtain: $HRES(r,t_i,t_f) = \frac{ECT(r,t_i,t_f)}{ANA(r,t_i,t_f)} = \frac{ACT(r,t_i,t_f) \cdot ENP(r,t_i,t_f)}{ANA(r,t_i,t_f)} = \frac{10 \cdot 4}{2} = 20$ min.

$HRES(r,t_i,t_f)$ returns the estimated care time of a room r between the t_i and t_f times, considering a parallel allocation of attendants in that period of time, through the use of $ANA(r,t_i,t_f)$ function. Thus, with the increase in the average number of attendants allocated, the estimated care time decreases, inversely proportional.

A problem of reactive elasticity is that the elasticity actions are taken after the upper threshold are reached, causing a state of overload in hospital throughout the professionals' movement period. Thus, an alternative to this problem is the use of proactive elasticity (Righi et al., 2019). Thus, anticipating the moment when the upper threshold will be reached, people's movement can occur in advance, minimizing or

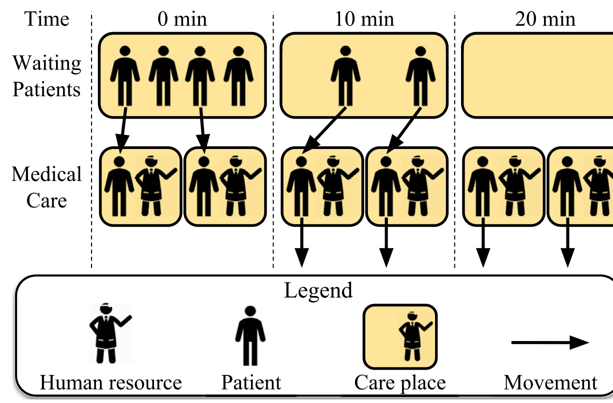

Fig. 8. Calculating the *ECT* in a hospital room using parallel allocation of attendants, with 4 patients waiting, average care time of 10 min, and 2 attendants, with an estimated care time, through *Human Resources Elastic Speedup*, of 20 min.

avoiding patients' overloads in hospital. In this context, we propose Eq. (6) for *Proactive Human Resources Elastic Speedup* as follows:

$$PHRES(r,a,f_i,f_f) = \frac{ECT(r,f_i,f_f)'}{a}, \tag{6}$$

where a is the number of attendants allocated between the future times f_i and f_f, and $ECT(r,f_i,f_f)'$ is a prediction of the future care time for this room using ARIMA. We can compute ECT as:

$$ECT(r,f_i,f_f)' = ACT(r,f_i,f_f)' \cdot ENP(r,f_i,f_f)', \tag{7}$$

where $ACT(r,f_i,f_f)'$ and $ENP(r,f_i,f_f)'$ are predictions of the average care time and future patients at room r, respectively. Thus, for each room r being calculated, we generate a time series of $ACT(r,t_i,t_f)$ that occurred in the past, and we use it to predict $ACT(r,f_i,f_f)'$. In addition, for each room we also generate a time series for $NIP(r,t_i,t_f)$, and can predict future patient input and find $ENP(r,f_i,f_f)'$.

Using the aforementioned equations, ElHealth are able to predict the waiting time of any hospital room. Varying a attribute in $PHRES$ equation, with the increase and decrease of the number of health professionals in attendance, ElHealth can identify how many attendants would be needed to adjust the waiting time of any room to the proposed thresholds, as defined by the hospital manager. Algorithm 1 presents our method to verify the need to allocate or de-allocate human resources in any room r in a smart hospital.

Algorithm 1: Room-Level Predictive Elasticity

Data: Room r, a attendants, future initial time f_i, future final time f_f
Result: Quantity of attendants to allocate or de-allocate

```
begin
    upper ← Upper Threshold of waiting time in r;
    lower ← Lower Threshold of waiting time in r;
    n ← 0;
    a' ← a;
    if PHRES(r, a, f_i, f_f) > upper then
        while a' < limit(r) e PHRES(r, a', f_i, f_f) > upper do
            n ← n + 1;
            a' ← a + n;
        end
    else if PHRES(r, a, f_i, f_f) < lower then
        while a' > 0 e PHRES(r, a', f_i, f_f) < lower do
            n ← n − 1;
            a' ← a + n;
        end
    end
    return n;
end
```

3.3.2. Hospital-level predictive elasticity

At the hospital-level, ElHealth needs to test different allocations for the attendants so as to ensure that all rooms identified in the previous step (Section 3.3.1) have enough attendants, and to minimize overcrowding. Our algorithm considers the possibility of moving health professionals between different hospital environments in order to optimize the medical care time. To this end, the available options refer to: allocating new attendants, reallocating health professionals between different sectors, or de-allocating human resources that are no longer necessary. ElHealth's first option should always be the possibility of reallocating human resources already allocated to hospital care. The reallocation is prioritized because is the option that brings fewer costs to the hospital since it performs adjustment of medical care without additional attendants. To redistribute such health attendants between different hospital rooms, our model uses some strategies known from other contexts of scientific computing, and adapts them to the predictive elasticity of human resources needs. Algorithm 2 presents the pseudo-code for hospital-level predictive elasticity.

Algorithm 2: Hospital-Level Predictive Elasticity

Data: Hospital room list h, vector v with all attendants of hospital, future initial time f_i, future final time f_f
Result: Updated hospital room list h

1 **begin**
2 $l \leftarrow$ a new vector of rooms and quantity of attendants to allocate or de-allocate;
3 **forall** *Room r on hospital room list h* **do**
4 $a \leftarrow$ number of attendants allocated in r;
5 $q \leftarrow$ run Algorithm 1 for *Room-level Predictive Elasticity* using r, a, f_i and f_f as Data;
6 $l.add(r, q)$;
7 **end**
8 sort l, quantity of available attendants;
9 $l \leftarrow$ run Algorithm 5 for *Human Resources Deallocation* using l and allocated attendants of v as Data;
10 sort l, quantity of available attendants;
11 **forall** *Room r on list l* **do**
12 $l_r \leftarrow$ sort l, quantity of available attendants with room r specialty;
13 $available_r \leftarrow$ list of all human resources available for allocation with room r specialty;
14 run Algorithm 4 of *Human Resources Allocation* using r, l_r and $available_r$ as Data;
15 **end**
16 $h \leftarrow$ rooms of l vector;
17 **return** h;
18 **end**

In what follows, we firstly discuss the reallocation concept, followed by the allocation procedures, rules and algorithms. Lastly, we present the de-allocation process. We note that, although de-allocation appears first in the algorithm (line 10), it actually builds upon the human resources allocated during the preceding iteration of the algorithm. In ElHealth model, each room has a required specialty to the human resources that are allocated in it. In parallel, each health professional has a list with all its specialties. The process of reallocating or allocating human resources is only performed between professionals who have the required destination room specialty. This is necessary because in a laboratory exams room is required a nursing professional accustomed to blood tests for example, and even if we have x-ray technicians available for reallocation, they are not able to improve the attendance in the aforementioned room.

In order to achieve a balanced reallocation of human resources, we developed a variation of the dynamic List Scheduling algorithm (Wang and Sinnen, 2018), which was originally used for process scheduling. Here, all hospital rooms are in a list ordered by the number of attendants available for reallocation. In that way, whenever a room r needs more attendants, the elasticity manager checks for available attendants, with room r specialty, in the first room of the list . If attendants are available, then they are reallocated to the room lacking them, and the list is sorted again. If more attendants are needed, the algorithm checks the first room in the list again, and so forth, until the room obtains all the required attendants. This whole process is presented in Algorithm 3.

Fig. 9 illustrates the reallocation process, where room 1 needs five more attendants and rooms 2 and 4 have some free attendants. Following the logic of adapted List Scheduling algorithm, in the first round, room 2 is the first in the list, with five available attendants, and gives an attendant for room 1. In the second round, room 2 remains the first on the list, now with five free health professionals. Among these, gives another attendant to room 1. Soon after, in the third round, room 2 remains at the top of the list with three free human resources. Now, another attendant is reallocated from room 2. Then, in the fourth round, even though all rooms in the list have the same number of free attendants, room 2 remains at the top of the list, so another attendant is reallocated. Finally, in the fifth round, room 4 becomes the first on the

Algorithm 3: Human Resources Reallocation through the adapted List Scheduling algorithm

Data: Room r that requires attendants, and the sorted list l_r with all rooms
Result: Final situation of room r

begin
$next \leftarrow$ first room in list l_r;
while *r needs attendants and there are attendants available in l_r with room r specialty* **do**
r receives an attendant from $next$;
$l_r \leftarrow$ list l_r sorted again;
if *r still needs attendants* **then**
$next \leftarrow$ first room in list l;
else
return *r is with adequate allocation*;
end
end
return *r still needs attendants*;
end

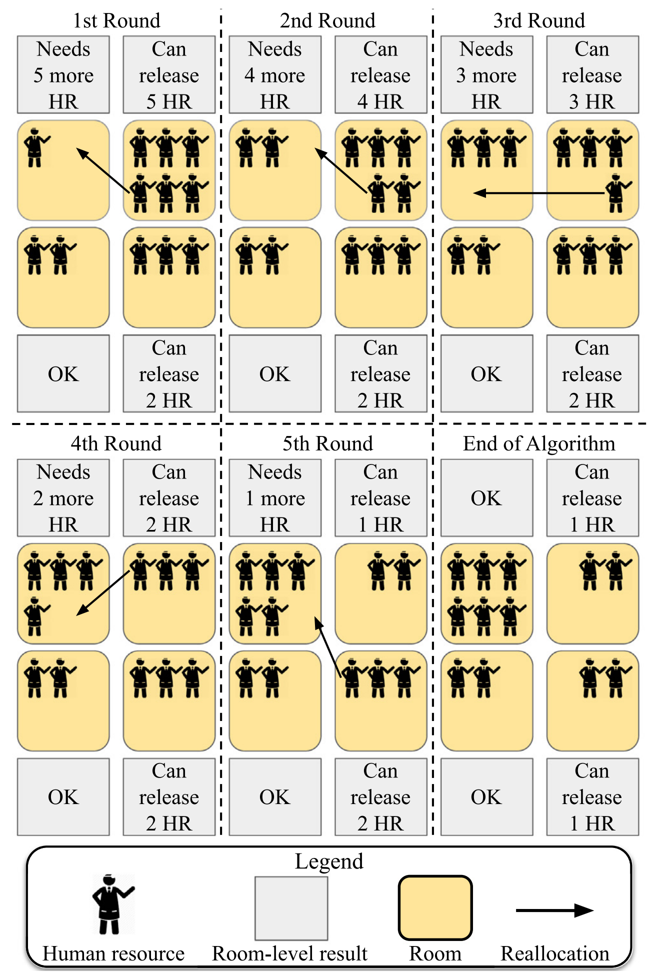

Fig. 9. Reallocation through the adapted List Scheduling algorithm, with a sorted list of 4 rooms, and 12 attendants, where Room 1 needs to allocate more 5 attendants.

list, since it has two free human resources (as opposed to room 2, which has only one), and an attendant of room 4 is reallocated to room 1.

A potential problem that arises in the context of elasticity is the so-called *hysteresis*, which refers to the tendency of the system to return to the previous state in the absence of the impulse that caused the change. In the context of human resources elasticity, hysteresis occurs if a resource reallocated from a given room A to another room B and, in the subsequent time-step, room B needs that resource back. This kind of situation happens when the stimulus that led to the reallocation ceases to exist. However, when the resource is returned to the original room, the stimulus will emerge once again, leading the resource to be constantly reallocated between the two rooms.

In order to prevent hysteresis of human resources, we employ a cooldown-based strategy (Kejariwal, 2013). In particular, whenever a resource is reallocated from a given room A to another room B, should room B need that resource back in the subsequent time-step, its need will only be met if another room has free resources. In other words, the resource cannot be immediately returned, which avoids the hysteresis effect.

In some situations, the reallocation process performed by Algorithm 3 may not be enough to improve the attendance level of the hospital. In such situations, allocation of new resources may be necessary. Algorithm 4 presents procedure for human resource allocation. In the algorithm, line 2 attempts to reallocate human resources according to the priority proposed in this model. We emphasize that, in order to minimize operational costs, allocation is only performed if reallocation is not able to meet the patients demand. We highlight that the hospital must have a strategy to define human resources available for external allocation. Since different countries have different labor laws, the rules that can make available for allocation the hospital staff on rest time can vary.

Algorithm 4: Human Resources Allocation

```
Data: Room r that requires attendants, the sorted list l_r with all rooms, and a vector
      v with all externally available attendants for allocation with room r specialty
Result: Final situation of room r
1  begin
2      Execute Algorithm 3 of Human Resources List Scheduling using r and l_r as Data;
3      if r do not need more attendants then
4          return r is with adequate allocation;
5      else
6          while r needs attendants and there are attendants available in v do
7              r receives an attendant from v;
8              if r do not need more attendants then
9                  return r is with adequate allocation;
10         end
11         return r still needs attendants;
12     end
13 end
```

Finally, if the algorithm identifies that the demand for care of all hospital rooms is very low and that the de-allocation of attendants of some room does not harm the whole, ElHealth must identify which attendants were allocated outside of their normal working hours and de-allocate them to lower the hospital's financial costs. Algorithm 5 presents the steps towards human resource de-allocation.

4. Evaluation methodology

We assess the performance of ElHealth through simulations in a virtual hospital environment. Considering the unavailability of data, the hospital environment was defined based on synthetic workloads. These data and its parameters are detailed in Section 4.2. According to Islam et al. (2012), synthetic workloads can be considered a representative form to evaluate elasticity in computational clouds.

ElHealth was implemented mainly in Java, except for the ARIMA method, which was implemented in Python. For hospital queues simulation, we used a clock with discrete increments of ten seconds. At each

Algorithm 5: Human Resources De-allocation

```
Data: Sorted list l with all rooms, and a vector a with all allocated attendants
Result: Room list l updated
begin
    forall Human Resource hr on list a do
        if tr ≤ a maximum time limit for allocated human resource then
            release hr;
    end
    qd ← number of attendants available in rooms in l;
    qf ← number of attendants missing in the l ;
    while qd < qf and size(a) > 0 do
        sort a, by time of care in descending;
        hr ← a.get(0);
        release hr;
    end
    return l;
end
```

advance in the simulation clock, our simulator verifies the patients who are in care and those who should leave the care. At each monitoring cycle, the arrival of patients should be checked. The data probability distributions were generated using triangular distributions (more details in Section 4.2), as implemented by *StdRandom* (Sedgewick and Wayne, 2017).

4.1. Considered scenarios

Given the hospital simulation procedure, we consider three different scenarios:

- **S1**: Hospital without ElHealth
- **S2**: Smart hospital with ElHealth for only human resources reallocation
- **S3**: Smart hospital with ElHealth for human resources allocation, reallocation and deallocation

4.2. Performance evaluation parameters

To perform the simulation of the hospital environment, we use the data collected in the study of Capocci et al. (2017) performed in a hospital environment located in Guarulhos City, in the state of São Paulo in Brazil. According to Capocci et al. (2017), all patients upon entering the unit first go through reception, where a Personal Health Record (PHR) (Roehrs et al., 2017) is prepared. After this preparation, patients are referred to waiting for triage. In the triage procedure, the patients are examined by the nursing team and classified into priorities according to the urgency of the health problem and are referred to waiting for medical attention. In polyclinic analyzed by Capocci et al. (2017), after first medical attention, 24% of patients are referred for x-ray exam, 37% for laboratory examinations (blood test, for example), 8% for electrocardiograms (ECG) exam, and 31% do not need more than physician examination. Also after doctor treatment room, only 1% of patients do not take medication and are released with only one prescription, but 50% of patients require intravenous medication, 30% intramuscular injection and 19% inhalation medication. After the exams, 60% of patients need to return to doctor, and 40% are released. After a return care, 78% of patients are released, 2% need new exams, and 20% require new medication.

Also according to Capocci et al. (2017), the care time in each room of the hospital environment follows a triangular distribution, with minimum and maximum times and a more frequent average time. Table 3 shows the distributions for all possible care in this hospital unit, as identified by Capocci et al. (2017) in their study. All other parameters used in our simulation can be found in Capocci et al. (2017).

As our case study is based on Brazilian hospital data, we have set thresholds appropriate to our reality. So, based in Brazilian Law Project of June 14, 2018 (Fabio, 2018) that proposes a maximum waiting time

Table 3
Triangular distributions of probability for care times.

Attendance	Attendance time		
	Lower	Mode	Upper
Reception room			
PHR preparation	2 min	3 min	5 min
Triage room			
Triage process	5 min	8 min	10 min
Doctor treatment room			
First care with doctor	5 min	11 min	16 min
Return care with doctor	4 min	7 min	10 min
Collection exams room			
Laboratory exams	6 min	8 min	13 min
X-ray exams room			
X-ray exam	10 min	15 min	23 min
Electrocardiogram exams room			
ECG exam	30 min	45 min	60 min
Medication room			
Intramuscular injection	3 min	3.5 min	5 min
Intravenous and inhalation preparation	0.5 min	1.5 min	2.5 min
Intravenous medication	40 min	70 min	120 min
Inhalation medication	8 min	10 min	13 min

for care in hospitals, clinics, and laboratories of 30 min on normal days (from Monday to Sunday), we set ElHealth's upper threshold in 30 min. Since works such Rostirolla et al. (2018), Righi et al. (2016), Rosa Righi et al. (2016) and Al-Haidari et al. (2013) uses lower threshold for elasticity of 30% of the maximum system load, we set ElHealth's lower threshold in 9 min (30% of maximum waiting time).

In Brazil, the working model adopted for hospital environments is the so-called 12x36 h. According to Brazilian Law No. 13,467 (Brazilian government, 2017), under this work regime, an employee can work for twelve consecutive hours (with a one-hour pause for lunch) and must rest for thirty-six hours before a new work shift of twelve hours starts. Under this regime, four health professions alternating shifts is enough to ensure a single position for twenty-four hours, seven days a week. Also according to the understanding of the law, if for any reason an employee needs to work within their rest period, it should be treated as overtime, unless the hours are compensated at another time.

Thus, while a human resource of the hospital is in working time, there are three other employees who perform the same function in their paid-rest period. According to Brazilian Decree-Law No. 5,452 (Brazilian government, 1943), the minimum rest period between two working days must be eleven consecutive hours. In that way, even if there are overtime hours, an employee must rest eleven hours to return to the next work shift. Thus, these three resting employees shall not be arbitrarily available to a new allocation. In particular, any resting employee is only available under the following rules:

Rule 1: The minimum rest period for a human resource to be available for allocation is eleven hours;
Rule 2: An allocated employee cannot works outside of the normal work shift for a long time period. The largest possible work period allowed in Brazilian legislation is twelve hours. Thus, an allocated employee cannot work more than twelve hours;
Rule 3: Allocated employees must be de-allocated no later than 11 h before they next normal work shift; and
Rule 4: Each employee must meet one of the 36 h rest periods within the same week in order to comply with a law determination that requires all workers to have a 24 h paid-rest period per week.

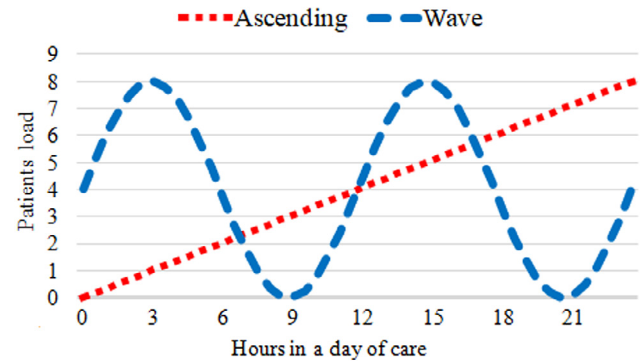

Fig. 10. Graphical representation of workloads used in ElHealth tests, where x axis expresses time available in one day of care, while y axis represents the arrival of patients at each time instant.

4.3. Workload

We use the human resources allocation found in Capocci et al. (2017) research, where 11 health professionals were allocated, 24 h a day, seven days a week, through more than one work shift. To be specific, health professionals were allocated as follows: 2 attendants in a reception; 1 nurse working in patient triage; 2 doctors acting in doctors treatment rooms; 2 nurses working with collection exams; 2 nurses working throughout the medication area; 1 nurse acting on the electrocardiogram; and 1 radiology technician acting with the x-ray exams.

Regarding patients load, we modeled two workloads: ascending and wave. The idea of using different load behaviors for the same application is used to observe how the input load can impact saturation points, bottlenecks, and the addition or removal of resources (Righi et al., 2016). These two behaviors of workload are based on those proposed by Righi et al. (2016). Thus, wave workload are the most closely to the hospital reality, and the ascending workload was chosen to identify the behavior of the model in a situation of increased patient load, which could be caused, for example, by a viral outbreak. Fig. 10 presents a representation of each workload of the model. The x axis expresses the time available in one day of care in the hospital unit, while the y axis represents the arrival of patients at each instant of time.

Since that the workloads generate decimal numbers, we established a strategy to generate integers for the arrival of the patients in the hospital environment. This occurs because in a real environment, is not possible the arrival of 0.2 patients or 1.7 patients, for example. Thus, we adopted a load accumulation strategy, where if at any given moment there is something between 0.1 and 0.9 patient, this value is accumulated with next instant load. An example would be any instant with a load of 0.6 patient. Since there would not be an integer charge, a patient would not be introduced into the system and the charge would accumulate for the next instant of time. At the next moment, with a new load of 0.6 patient, the accumulated load would be 1.2 patient, resulting in the entry of 1 patient in the hospital. Thus, there would be still 0.2 patient, which would be accumulated for the next instant and so on.

4.4. Performance evaluation metrics

In order to evaluate the proposed model, the following metrics are considered:

- Maximum waiting time for care;
- Elastic number of human resources used.

Table 4
Evaluation metrics and expected results in each scenario.

Scenario	Maximum waiting time	Elastic number of human resources used
S1	Current	11 by work shift
S2 (Expected)	Less than S1	11 by work shift
S3 (Expected)	Less than S2	11 or more by work shift

To evaluate the waiting time, we used as parameter the variation of the maximum waiting time between the scenarios and the adequacy of the maximum waiting time to the established limits. With regard human resources number, we expect that our model uses the existing health professionals in the hospital in an optimized way. Thus, static allocation of S1, with eleven employees working, can be compared to ElHealth elastic allocation, with the number of human resources varying throughout the day. Table 4 presents all the evaluation metrics described above, relating the results expected for the second and third scenario with the use of ElHealth, when compared to the current hospital environment, without the ElHealth model.

5. Performance evaluation and results analysis

Based on the evaluation methodology proposed for the ElHealth model, we performed six simulations of the proposed hospital environment in order to collect results for analysis. For each proposed scenario, between S1, S2 and S3, a simulation was performed for each of the workloads, Ascending and Wave.

For the maximum waiting time metric, we expected a decrease in patient waiting for care. Fig. 11 shows the maximum waiting time identified for each workload in the proposed scenarios over the simulated one-week period. We perceive a significant reduction in the maximum waiting time between S1 and S2, and a second diminution when comparing S2 and S3, regardless of the workload used. After a thorough analysis, we can identify that in S3 for reception, triage, doctor treatment, and collection exams rooms, at no time was measured waiting time longer than 30 min, regardless of the workload used. As for medication, x-ray, and electrocardiogram rooms, there were a few moments when this limit was exceeded. Through the collected data, we identify a significant reduction in waiting time with the use of the Multi-level Predictive Elasticity of Human Resources when compared to the hospital without the use of the elasticity.

For elastic number of human resources used metric, we expected an increase in the number of professionals in the hospital, as well as an variation of this number over the hospital care period. Fig. 12 presents the elastic number of human resources used for hospital care in S3, the only scenario where the number of employees can variate. We can observe that the elastic number of human resources ranged from 11 to 14 per hour. Although there are moments with the allocation of up to 14 health professionals in care, the average per hour of care professionals turns out to be slightly lower depending on the time it takes for an employee to be allocated or reallocated in the hospital.

Furthermore, as exposed in aforementioned Fig. 12, whenever ElHealth reallocates or allocates people for care, the patients' waiting time decreases. Thanks to the reallocation and allocation procedures, ElHealth has shown to decrease the waiting time by 95.5% and 96.71% for wave and ascending workloads, respectively, as compared to scenario where no reallocation is performed.

5.1. Discussion

Based on established metrics, we can note that the ElHealth model was able to improve the performance of the simulated hospital environment in all workloads used. Table 5 presents all the results found in each of proposed evaluation metrics. As proposed in our evaluation methodology, we expected that the maximum waiting time presented a gradual decrease between scenarios S1, S2, and S3, and this in fact occurred, fulfilling the objective of this metric. For the elastic

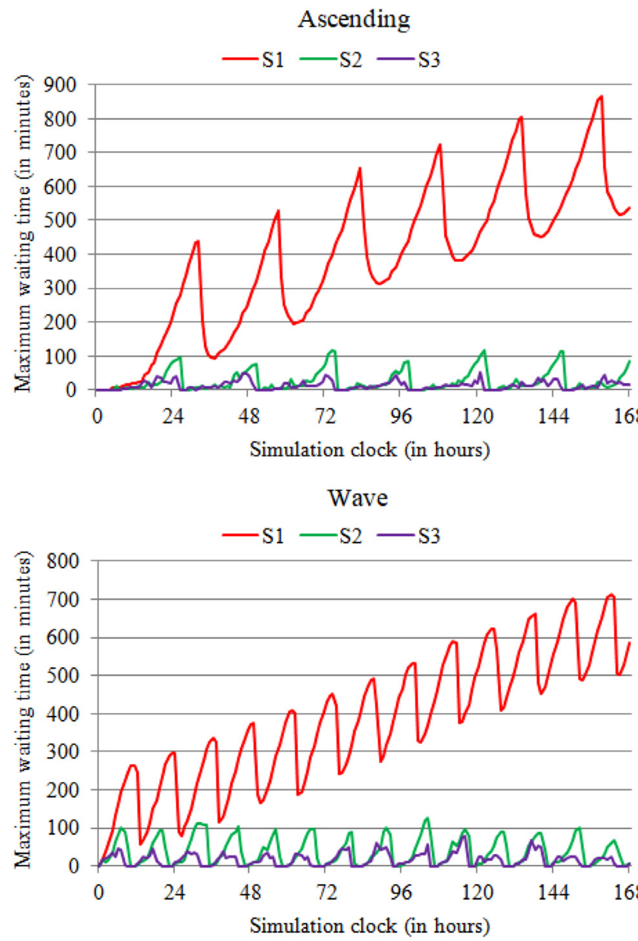

Fig. 11. Maximum waiting time at the hospital for each of the proposed scenarios, S1, S2, and S3, using ascending and wave workloads.

Table 5
Evaluation metrics and results found in each of the proposed scenarios, using ascending and wave workloads.

Workload	Scenario	Maximum waiting time (in minutes)		Elastic number of human resources
		Average	Upper	
Ascending	S1	388.81 (±215.8)	868	11
	S2	25.32 (±30.4)	117	11
	S3	12.88 (±12.4)	50	11.36
Wave	S1	384.18 (±171.7)	711	11
	S2	39.15 (±35.9)	126	11
	S3	17.21 (±17.2)	77	11.68

number of human resources used, an increase in the result was expected between scenarios S2 and S3, and our model once again was able to meet the proposed goal. Thus, the expected results in the evaluation methodology were achieved through the use of the ElHealth model in the proposed hospital environment.

For maximum waiting time metric, there were two objectives: time reduction and the framing of the time within the established upper limit of 30 min. As already shown, the ElHealth model was able to significantly reduce waiting time for the proposed hospital environment. However, although the average maximum waiting times for the S3 scenario were within the established limit (12.88 min with Ascending workload and 17.21 with Wave workload), when we analyzed the longer waiting time identified in all the simulation period, the upper limit was exceeded (50 min with Ascending workload and 77 min with Wave workload). We believe that this occurred due to the limitations

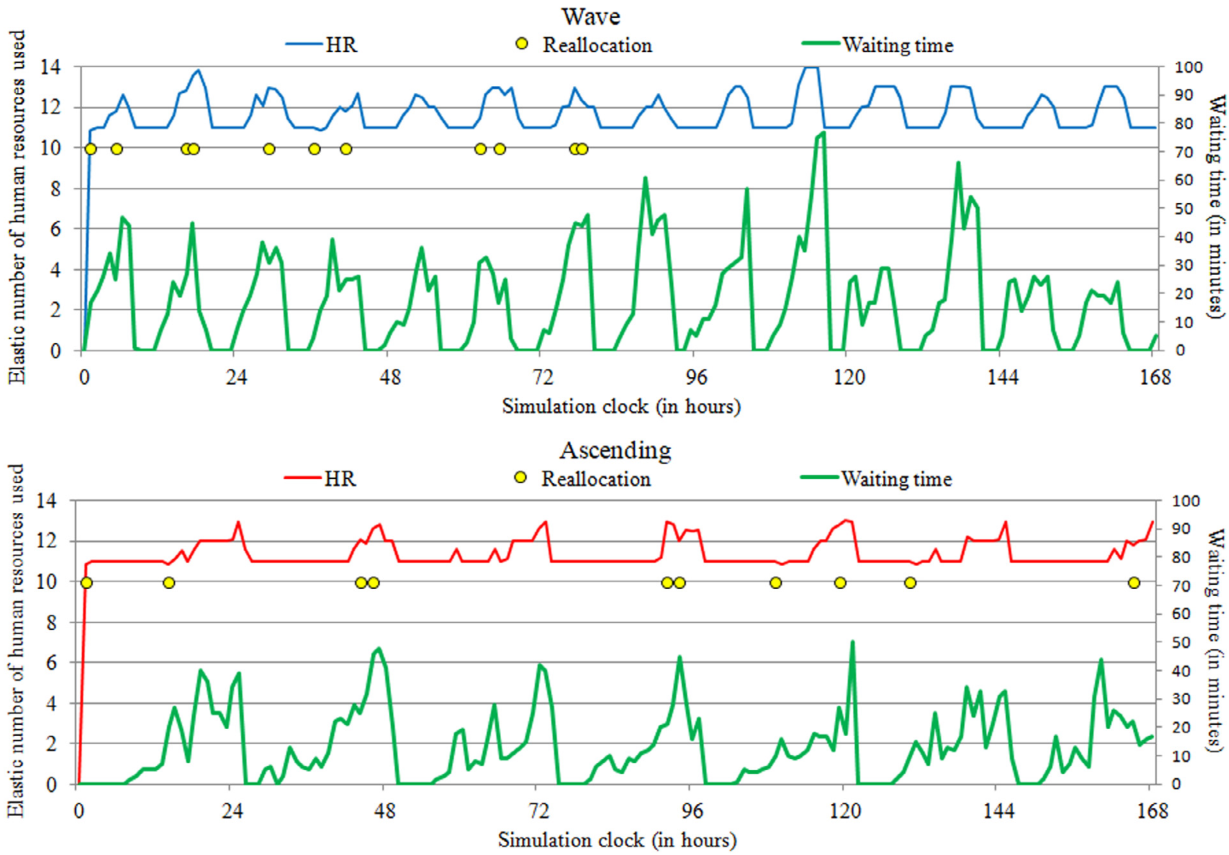

Fig. 12. Elastic number of human resources used compared with maximum waiting time at the hospital using ascending and wave workloads in S3, where Multi-level Predictive Elasticity of Human Resources are completely used.

of the hospital environment used as the basis for this simulation. As there were not many care stations available to be allocated new human resources, our model was not able to reach the goal in this hospital environment. For elastic number of human resources used metric, we expected an increase in the average number of human resources between scenarios C2 and C3, and this also actually occurred. Based on the data collected through the simulations and tests, we can note that there is evidence of the ElHealth model's functioning and its ability to adjust the workforce of a hospital environment to the patients' demand, reducing the waiting time for care and the average number of patients waiting in the queues. However, taking into account that the tests were based on a simulated environment, is important to point out that the results obtained cannot be generalized. Thus, we can say that for an absolute validation of the system would be necessary its effective implantation in a hospital, in order to capture real data of the demand of patients to be analyzed. So, from the technical point of view, there is evidence of the prototype developed to be functional, but we cannot only prove with simulation its effectiveness in the real world.

5.2. Deployment considerations

In order to deploy ElHealth, we could use radio frequency identification (RFID) sensors. Nowadays, RFID is one of the most widely used IoT technologies to detect the proximity of a person to a particular object. We could deploy our approach with a fixed RFID reader at each entrance of a room in the hospital. Each such reader could then be connected to two or three antennas to capture movement around that entrance, so minimizing the possibility of false-negative occurrences. In particular, antennas could be attached on each side (left and right) and on the top of each door. Besides the RFID reader and antennas, each patient should have an RFID-enabled wristband. Building upon such configuration, we have studied the deployment costs of our approach based on several manufacturers, so resulting in the following average costs per unit.

- Fixed reader: US$1338.00
- Antenna: US$103.00
- RFID-enabled wristband: US$1.00

The above values should then be multiplied by the number of rooms to be monitored, following the ElHealth algorithms. In addition to the aforementioned costs, we must also consider the IT software and hardware installation. Nevertheless, RFID is a mature technology and well-established systems to manage RFID data are nowadays available at reasonable costs.

6. Conclusion and future works

This article presented the ElHealth model. Unlike related work, ElHealth not only proposes a use of data prediction to anticipate eventual problems in the future, but also presents a model to allocate, migrate and deallocate people in hospitals in such a way to provide benefits at patients viewpoint. Using sensors and a ARIMA-based prediction engine, we can instrument a smart hospital to collect data in time-series, so better arranging professionals and either preventing or mitigating patient treatment problems, which sometimes are related to live or death issues. In this way, we extended the concept of elasticity from cloud computing to the context of human resources management, while proposing new mathematical formalisms, algorithms and definitions to provide a dynamic and elastic allocation of professionals in hospital environments.

We expect that the model proposed in this work can help to decrease the waiting time of patients for healthcare. The idea is to provide such facility in transparent way for the patients, *i.e.* , they do not need to follow additional procedures in the hospital, but only wear a wristband which serves as identification. We also hope to, with the use of ElHealth, we can identify bottlenecks in the patients care flow and help optimize processes in healthcare environments. Moreover, the provided data can also be used for decision making in terms of changes in the hospital capacity and infrastructure. In ElHealth's case study, the waiting time is decreased by 95.5% and 96.71% for wave and ascending workloads, respectively.

Although presenting encouraging results, we envisage some limitations that must be addressed on implementing ElHealth model in a real hospital environment: (i) employees and patients must carry their identification tags throughout their time in the smart hospital; (ii) ElHealth only generates notifications for human resource, but the effective movement of staff in hospital environments depends on their individual decision to follow the recommended guidance; (iii) previous installation of RTLS sensors in corridors and doors of the hospital.

As future work, we envisage the development of a prototype that implements all the modules and algorithms proposed by ElHealth, so enabling the deploying in a real hospital environment. Another possibility concerns the adaptation of the model to use other prediction algorithms including Artificial Neural Networks and Random Forest approaches.

Acknowledgments

We thank the anonymous reviewers for their valuable suggestions. This work was partially supported by the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES, finance code 001); by the Foundation for Research of the State of Rio Grande do Sul (FAPERGS); by Fundação para a Ciência e a Tecnologia (FCT, through the UID/EEA/50008/2019 Project); and by the Brazilian National Council for Research and Development (CNPq, grant no. 309335/2017-5).

References

Akeju, O., Butakov, S., Aghili, S., 2018. Main factors and good practices for managing BYOD and IoT risks in a K-12 environment. Int. J. Internet Things Cyber-Assur. 1 (1), 22–39.

Al-Dhuraibi, Y., Paraiso, F., Djarallah, N., Merle, P., 2018. Elasticity in cloud computing: State of the art and research challenges. IEEE Trans. Serv. Comput. 11 (2), 430–447.

Al-Haidari, F., Sqalli, M., Salah, K., 2013. Impact of cpu utilization thresholds and scaling size on autoscaling cloud resources. In: 2013 IEEE 5th International Conference on Cloud Computing Technology and Science, Vol. 2. IEEE, Bristol, UK, pp. 256–261.

Ali, F., Islam, S.R., Kwak, D., Khan, P., Ullah, N., jo Yoo, S., Kwak, K., 2018. Type-2 fuzzy ontology-aided recommendation systems for iot-based healthcare. Comput. Commun. 119, 138–155.

Amdahl, G.M., 1967. Validity of the single processor approach to achieving large scale computing capabilities. In: Proceedings of the April 18-20, 1967, Spring Joint Computer Conference. In: AFIPS '67 (Spring), ACM, New York, NY, USA, pp. 483–485.

Boulos, M.N.K., Berry, G., 2012. Real-time locating systems (RTLS) in healthcare: a condensed primer. Int. J. Health Geogr. 11 (25), 1–8.

Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis: Forecasting and Control. In: Holden-Day series in time series analysis, Holden-Day, San Fransisco, CA.

Brazilian government, 1943. Decreto-Lei No 5.452, de 1 de maio de 1943.

Brazilian government, 2017. Lei No 13.467, de 13 de julho de 2017.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. In: The Wadsworth statistics/probability series, Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA.

Butean, A., David, A., Buduleci, C., Daian, A., 2015. Auxilum medicine: A cloud based platform for real-time monitoring medical devices. In: 2015 20th International Conference on Control Systems and Computer Science. IEEE, Bucharest, Romania, pp. 874–879.

Capocci, N.R., Nascimento, B.S., Lopes, F.B., Rodrigues, E.F., Maiellaro, J.R., 2017. Simulation as a hospital management support tool. Indep. J. Manag. Prod. 8 (5), 798–811.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297.

da Costa, C.A., Pasluosta, C.F., Eskofier, B., da Silva, D.B., da Rosa Righi, R., 2018. Internet of health things: Toward intelligent vital signs monitoring in hospital wards. Artif. Intell. Med. 89, 61–69, URL http://www.sciencedirect.com/science/article/pii/S0933365717301367.

Cover, T., Hart, P., 1967. Nearest neighbor pattern classification. IEEE Trans. Inform. Theory 13 (1), 21–27.

Cox, D.R., 1958. The regression analysis of binary sequences. J. R. Stat. Soc. Ser. B Stat. Methodol. 20 (2), 215–242.

Cross, V.V., 2014. Fuzzy ontologies: The state of the art. In: 2014 IEEE Conference on Norbert Wiener in the 21st Century (21CW). IEEE, Boston, MA, USA, pp. 1–8.

Darshan, K.R., Anandakumar, K.R., 2015. A comprehensive review on usage of internet of things (iot) in healthcare system. In: 2015 International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT). IEEE, Mandya, India, pp. 132–136.

Fabio, Z., 2018. Projeto de Lei, de 14 de junho de 2018.

Fix, E., Hodges, Jr., J.L., 1951. Discriminatory analysis-nonparametric discrimination: consistency properties, Tech. rep., California Univ Berkeley.

Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Statist. 29 (5), 1189–1232.

Graham, B., Bond, R., Quinn, M., Mulvenna, M., 2018. Using data mining to predict hospital admissions from the emergency department. IEEE Access 6, 10458–10469.

Ho, T.K., 1995. Random decision forests. In: Proceedings of 3rd International Conference on Document Analysis and Recognition, Vol. 1. IEEE, Montreal, Canada, pp. 278–282.

Ishikawa, T., Fujiwara, K., Ohba, H., Suzuki, T., Ogasawara, K., 2017. Forecasting the regional distribution and sufficiency of physicians in japan with a coupled system dynamics—geographic information system model. Hum. Resour. Health 15 (1), 64.

Islam, S., Lee, K., Fekete, A., Liu, A., 2012. How a consumer can measure elasticity for cloud platforms. In: Proceedings of the 3rd ACM/SPEC International Conference on Performance Engineering. In: ICPE '12, ACM, New York, NY, USA, pp. 85–96.

Kejariwal, A., 2013. Techniques for optimizing cloud footprint. In: 2013 IEEE International Conference on Cloud Engineering (IC2E). IEEE, Redwood City, CA, USA, pp. 258–268.

Kundu, A., He, Y., Bahl, P., 1989. Recognition of handwritten word: First and second order hidden markov model based approach. Pattern Recognit. 22 (3), 283–297.

Liu, J.X., Goryakin, Y., Maeda, A., Bruckner, T., Scheffler, R., 2017. Global health workforce labor market projections for 2030. Hum. Resour. Health 15 (1), 11.

Moreira, M.W.L., Rodrigues, J.J.P.C., Korotaev, V., Al-Muhtadi, J., Kumar, N., 2019. A comprehensive review on smart decision support systems for health care. IEEE Syst. J. 1–10.

Nierop-van Baalen, C., Grypdonck, M., van Hecke, A., Verhaeghe, S., 2019. Health professionals' dealing with hope in palliative patients with cancer, an explorative qualitative research. Eur. J. Cancer Care 28 (1), e12889, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ecc.12889.

Nisha, K.G., Sreekumar, K., 2017. A review and analysis of machine learning and statistical approaches for prediction. In: 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT). IEEE, Thondamuthur coimbatore, India, pp. 135–139.

Orimaye, S.O., Leong, F.C., Lee, C.H., Ng, E.C.H., 2015. Predicting proximity with ambient mobile sensors for non-invasive health diagnostics. In: 2015 IEEE 12th Malaysia International Conference on Communications (MICC). IEEE, pp. 6–11.

Pinto, S., Cabral, J., Gomes, T., 2017. We-care: An iot-based health care system for elderly people. In: 2017 IEEE International Conference on Industrial Technology (ICIT). IEEE, Toronto, ON, Canada, pp. 1378–1383.

Righi, R.R., Rodrigues, V.F., da Costa, C.A., Galante, G., de Bona, L.C.E., Ferreto, T., 2016. Autoelastic: Automatic resource elasticity for high performance applications in the cloud. IEEE Trans. Cloud Comput. 4 (1), 6–19.

Righi, R.D.R., Rodrigues, V.F., Nardin, I.F.D., Costa, C.A.D., Alves, M.A.Z., Pillon, M.A., 2019. Towards providing middleware-level proactive resource reorganisation for elastic HPC applications in the cloud. 10 (1), 76–92.

Roehrs, A., da Costa, C.A., da Rosa Righi, R., de Oliveira, K.S.F., 2017. Personal health records: A systematic literature review. J. Med. Internet Res. 19 (1).

Rosa Righi, R., Costa, C.A., Rodrigues, V.F., Rostirolla, G., 2016. Joint-analysis of performance and energy consumption when enabling cloud elasticity for synchronous hpc applications. Concurr. Comput.: Pract. Exp. 28 (5), 1548–1571.

Rostirolla, G., d. R. Righi, R., Barbosa, J.L.V., da Costa, C.A., 2018. Elcity: An elastic multilevel energy saving model for smart cities. IEEE Trans. Sustain. Comput. 3 (1), 30–43.

dos Santos, U.J.L., Pessin, G., da Costa, C.A., da Rosa Righi, R., 2019. Agriprediction: A proactive internet of things model to anticipate problems and improve production in agricultural crops. Comput. Electron. Agric. 161, 202–213, http://www.sciencedirect.com/science/article/pii/S0168169917313856, BigData and DSS in Agriculture.

Sarhan, Q.I., 2018. Internet of things: a survey of challenges and issues. Int. J. Internet Things Cyber-Assur. 1 (1), 40–75.

Sedgewick, R., Wayne, K., 2017. Class StdRandom. Princeton University. https://introcs.cs.princeton.edu/java/stdlib/javadoc/StdRandom.html (accessed 12.01.19).

Singh, P., 2018. Internet of things based health monitoring system: opportunities and challenges. Int. J. Adv. Res. Comput. Sci. 9 (1), 224–228.

Singh, K.J., Kapoor, D.S., 2017. Create your own internet of things: A survey of iot platforms. IEEE Consum. Electron. Mag. 6 (2), 57–68.

Srinivas, J., Das, A.K., Kumar, N., Rodrigues, J., 2018. Cloud centric authentication for wearable healthcare monitoring system. IEEE Trans. Dependable Secure Comput. 1.

Vieira, D., Hollmén, J., 2016. Resource frequency prediction in healthcare: Machine learning approach. In: 2016 IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS). IEEE, Dublin, Ireland, pp. 88–93.
Wang, H., Sinnen, O., 2018. List-scheduling versus cluster-scheduling. IEEE Trans. Parallel Distrib. Syst. 29 (8), 1736–1749.
Zadeh, L.A., 1975. The concept of a linguistic variable and its application to approximate reasoning-I. Inform. Sci. 8 (3), 199–249.
Zamanifar, A., Nazemi, E., Vahidi-Asl, M., 2017. Dmp-iot: A distributed movement prediction scheme for IOT health-care applications. Comput. Electr. Eng. 58 (Supplement C), 310–326.