



Antonio Iyda Paganelli

**A novel self-adaptive approach for optimizing
the use of IoT devices in patient monitoring
using EWS**

Tese de Doutorado

Thesis presented to the Programa de Pós-graduação em Informática of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Ciências - Informática.

Advisor: Prof. Markus Endler

Rio de Janeiro
March 2023



Antonio Iyda Paganelli

**A novel self-adaptive approach for optimizing
the use of IoT devices in patient monitoring
using EWS**

Thesis presented to the Programa de Pós-graduação em Informática of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Informática. Approved by the Examination Committee:

Prof. Markus Endler

Advisor

Departamento de Informática – PUC-Rio

Prof. Alberto Barbosa Raposo

PUC-Rio

Prof. Adriano Francisco Branco

PUC-Rio

Prof. Paulo Sérgio Conceição de Alencar

University of Waterloo

Prof. Francisco José da Silva e Silva

Universidade Federal do Maranhão

Prof. Nathália Moraes do Nascimento

University of Waterloo

Rio de Janeiro, March 31st, 2023

All rights reserved.

Antonio lyda Paganelli

Majored in Informatics from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio) and in Physical Education from Estacio de Sá University, where he had the opportunity to participate in the Junior Year Abroad program at the University of Exeter (UK) in Sports and Exercise Sciences. He has a Master's in International Business from the Florida International University (USA) and a Master's in Computer Science from PUC-Rio.

In addition, has more than twenty and five years of experience in the industry, working with NGOs, ISPs, Digital Certification Authorities, Telecom, Cloud Computing, and consulting companies.

Bibliographic data

Paganelli, Antonio lyda

A novel self-adaptive approach for optimizing the use of IoT devices in patient monitoring using EWS / Antonio lyda Paganelli; advisor: Markus Endler. – 2023.

158 f: il. color. ; 30 cm

Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática, 2023.

Inclui bibliografia

1. Informática – Teses. 2. Internet das Coisas. 3. Sistemas de pontuação de alertas antecipados. 4. Monitoramento de pacientes. 5. Eficiência energética. 6. Sistemas embarcados. 7. Redução de dados. 8. Alarmes. I. Endler, Markus. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Aos meus avós e ao meu pai (in memoriam),
mãe, irmãos, filha, enteados e família
pelo apoio e encorajamento.

Acknowledgments

Ao meu orientador Professor Markus Endler pelo estímulo, confiança e parceria para a realização deste trabalho.

Ao CNPq, à PUC-Rio, ao Tecgraf, ao INCT, a FAPESP e à Universidade de Waterloo pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Aos meus colegas do Laboratory of Advanced Collaboration, da PUC-Rio e do Instituto Tecgraf por todo apoio, colaboração, e discussões durante a elaboração desta tese.

Aos meus pais, por todo suporte e carinho de todas as horas. Á minha esposa, Adriana, a minha filha Maria pela paciência e incentivos constantes.

Aos professores que participaram da Comissão examinadora.

Aos professores e funcionários do Departamento de Informática pelos ensinamentos e pela ajuda.

A todos os amigos e familiares que de uma forma ou de outra me estimularam ou me ajudaram.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Abstract

Paganelli, Antonio Iyda; Endler, Markus (Advisor). **A novel self-adaptive approach for optimizing the use of IoT devices in patient monitoring using EWS**. Rio de Janeiro, 2023. 158p. Tese de Doutorado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

The Internet of Things (IoT) proposes to connect the physical world to the Internet, which opens up the possibility of developing various applications, especially in healthcare. These applications require a huge number of sensors to collect information continuously, generating large data flows, often excessive, redundant, or without meaning for the system's operations. This massive generation of sensor data wastes computational resources to acquire, transmit, store, and process information, leading to the loss of efficiency of these systems over time. In addition, IoT devices are designed to be small and portable, powered by batteries, for increased mobility and minimized interference with the monitored environment. However, this design also results in energy consumption restrictions, making battery lifetime a significant challenge that needs to be addressed. Furthermore, these systems often operate in unpredictable environments, which can generate redundant and negligible alarms, rendering them ineffective. However, a self-adaptive system that identifies and predicts imminent risks using early-warning scores (EWS) can cope with these issues. Due to its low processing cost, EWS guidelines can be embedded in wearable and sensor devices, allowing better management of sampling rates, transmissions, alarm production, and energy consumption. Following the aforementioned idea, this thesis presents a solution combining EWS with a self-adaptive algorithm for IoT patient monitoring applications. Thus, promoting a reduction in data acquisition and transmission, decreasing non-actionable alarms, and providing energy savings for these devices. In addition, we designed and developed a hardware prototype capable of embedding our proposal, which attested to its technical feasibility. Moreover, using our wearable prototype, we collected the energy consumption data of hardware components and used them during our simulations with real patient data from public datasets. Our experiments demonstrated great benefits of our approach, reducing by 87% the sampled data, 99% the total payload of the transmitted messages from the monitoring device, 78% of the alarms, and an energy saving of almost 82%. However,

the fidelity of monitoring the clinical status of patients showed a mean total absolute error of 6.8% (+/- 5.5%) but minimized to 3.8% (+/- 2.8%) in a configuration with lower data reduction gains. The total loss of alarm detection depends on the configuration of frequencies and time windows, remaining between 0.5% and 9.5%, with an accuracy of the type of alarm between 89% and 94%. In conclusion, this work presents an approach for more efficient use of computational, communication, and energy resources to implement IoT-based patient monitoring applications.

Keywords

Internet of Things; Early-warning scoring system; Patient monitoring; Energy-efficiency; Embedded systems; Data reduction; Alarms.

Resumo

Paganelli, Antonio Iyda; Endler, Markus. **Uma nova abordagem auto-adaptável para otimizar o uso de dispositivos IoT no monitoramento de pacientes usando o EWS**. Rio de Janeiro, 2023. 158p. Tese de Doutorado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

A Internet das Coisas (IoT) se propõe a interligar o mundo físico e a Internet, o que abre a possibilidade de desenvolvimento de diversas aplicações, principalmente na área da saúde. Essas aplicações requerem um grande número de sensores para coletar informações continuamente, gerando grandes fluxos de dados, muitas vezes excessivos, redundantes ou sem significado para as operações do sistema. Essa geração massiva de dados de sensores desperdiça recursos computacionais para adquirir, transmitir, armazenar e processar informações, levando à perda de eficiência desses sistemas ao longo do tempo. Além disso, os dispositivos IoT são projetados para serem pequenos e portáteis, alimentados por baterias, para maior mobilidade e interferência minimizada no ambiente monitorado. No entanto, esse design também resulta em restrições de consumo de energia, tornando a vida útil da bateria um desafio significativo que precisa ser enfrentado. Além disso, esses sistemas geralmente operam em ambientes imprevisíveis, o que pode gerar alarmes redundantes e insignificantes, tornando-os ineficazes. No entanto, um sistema auto-adaptativo que identifica e prevê riscos iminentes através de um sistema de pontuação de alertas antecipados (EWS) pode lidar com esses problemas. Devido ao seu baixo custo de processamento, a referência EWS pode ser incorporada em dispositivos vestíveis e sensores, permitindo um melhor gerenciamento das taxas de amostragem, transmissões, produção de alarmes e consumo de energia. Seguindo a ideia acima, esta tese apresenta uma solução que combina um sistema EWS com um algoritmo auto-adaptativo em aplicações IoT de monitoramento de pacientes. Desta forma, promovendo uma redução na aquisição e transmissão de dados, diminuindo alarmes não acionáveis e proporcionando economia de energia para esses dispositivos. Além disso, projetamos e desenvolvemos um protótipo de hardware capaz de embarcar nossa proposta, evidenciando a sua viabilidade técnica. Além disso, usando nosso protótipo, coletamos dados reais de consumo de energia dos componentes de hardware que foram usados durante nossas simulações com dados reais de pacientes provenientes de banco

de dados públicos. Nossos experimentos demonstraram grandes benefícios com essa abordagem, reduzindo em 87% os dados amostrados, em 99% a carga total das mensagens transmitidas do dispositivo de monitoramento, 78% dos alarmes e uma economia de energia de quase 82%. No entanto, a fidelidade do monitoramento do estado clínico dos pacientes apresentou um erro absoluto total médio de 6,8% (+/- 5,5%), mas minimizado para 3,8% (+/- 2,8%) em uma configuração com menores ganhos na redução de dados. A perda de detecção total dos alarmes dependendo da configuração de frequências e janelas de tempo analisadas ficou entre 0,5% e 9,5%, com exatidão do tipo de alarme entre 89% e 94%. Concluindo, este trabalho apresenta uma abordagem para o uso mais eficiente de recursos computacionais, de comunicação e de energia para implementar aplicativos de monitoramento de pacientes baseados em IoT.

Palavras-chave

Internet das Coisas; Sistemas de pontuação de alertas antecipados; Monitoramento de pacientes; Eficiência energética; Sistemas embarcados; Redução de dados; Alarmes.

Table of contents

1	Introduction	19
1.1	Motivation	19
1.2	Research Topic Definition	21
1.3	Problem Statement	22
1.4	Goals and Research Questions	23
1.5	Solution Approach Components	24
1.6	Main Contributions	25
1.7	Publications	26
1.8	Thesis Organization	27
2	Background & Related work	28
2.1	Patient Monitoring in Hospitals	28
2.2	Internet of Things and Patient Monitoring Applications	30
2.3	Early-Warning Scores and Vital Signs Monitoring	36
2.4	Self-Adaptive Solutions	41
2.5	Related Work	42
2.6	Summary and Gaps	45
3	Proposed Approach	48
3.1	Combined and individual early-warning scores to promote data and alarm reductions and save energy in IoT-PMAs - RQ1	48
3.2	Reducing sensor data generation and measuring monitoring quality - RQ2	54
3.3	Combined early-warning scores for reducing non-actionable alarms - RQ3	60
3.4	Combined early-warning scores and energy-efficient in IoT devices - RQ4	64
3.5	Embedding the solution in IoT devices - RQ5	66
4	Proposed Approach Application	77
4.1	Testing Environment and Datasets	77
4.2	Experiments Organization	82
4.3	Experiment I - Use of combined early-warning scores for data and alarm reductions	84
4.4	Experiment II - Self-adaptive approach effects on energy efficiency	96
4.5	Experiment III - Effects of proposed self-adaptive algorithm principles on patient monitoring and a comprehensive investigation of energy efficiency	102
4.6	General Discussion	115
5	Final remarks and Future Work	119
5.1	Future Work	121
6	Bibliography	124

A	Research Methodology	135
A.1	Research Topic Definition	135
A.2	Related Work Investigation	135
A.3	Problem Statement Definition	136
A.4	Solution Approach Proposal	136
A.5	Solution Approach Evaluation	136
A.6	Research Final Remarks	136
B	SH-Sens Wearable Kit development	137
B.1	WKit development - Objectives	137
B.2	WKit Design - Version 1	137
B.3	WKit Design - Version 2	138
B.4	WKit development - Discussion	142
C	Web and Mobile Application Prototypes	145
D	Preliminary Experiment	150
D.1	Introduction	150
D.2	Objectives	150
D.3	Methodology	150
D.4	Self-adaptive features	151
D.5	Comparison to other algorithms	154
D.6	Results	155
D.7	Discussion	155
D.8	Conclusion	157

List of figures

Figure 1.1	Overview of the proposed approach for adaptive monitoring. It uses combined and individual early-warning scores in an embedded solution applying adaptive sampling and data aggregation to regulate data rates.	24
Figure 2.1	A typical Wireless Body Area Network architecture. Image source: (FILIPE et al., 2015).	32
Figure 2.2	Energy-efficient strategy taxonomy proposed by (RAULT; BOUABDALLAH; CHALLAL, 2014).	35
Figure 2.3	NEWS-2 table guidelines.	38
Figure 2.4	Combined Scores based on NEWS-2 individual scores.	39
Figure 2.5	Elghers et al. (2014) self-adaptive algorithm.	43
Figure 3.1	Proposal approach overview - recap.	48
Figure 3.2	Dissimilarity index computation example.	52
Figure 3.3	Bézier Quadratic Curve properties. P1 is the patient's risk defined by r_0 , which will adjust the shape of the distribution. Source: (LAIYMANI; MAKHOUL, 2013).	53
Figure 3.4	Proposed self-adaptive algorithm flowchart. Steps in blue, yellow, and red are performed at sampling rates, while in green are executed at the period time ($pLen$).	54
Figure 3.5	Strategies to avoid negligible and redundant alarms. Trigger delays postpone alarms. The redundant lockout window avoids repeating the same alarm.	62
Figure 3.6	Alarms procedures. Three asynchronous processes run concurrently: new scores entrance (e1), alarm delay control in pink and yellow, and redundant window control in blue.	63
Figure 3.7	The conceptual architecture of an IoT-based patient monitoring application as a reference for our embedded solution. Based on (PAGANELLI et al., 2022).	67
Figure 3.8	Virtualized implementation of our patient monitoring solution modules.	69
Figure 3.9	WKit hardware modules in the context of a simplified view of our conceptualized architecture.	70
Figure 3.10	Data model summary. SensorInfo represents Heart Rate (HR), Body Temperature, and Blood-oxygen saturation (SO).	71
Figure 3.11	Images of our SH-Sens WKit prototype and base station. The wearable device is a finger clip.	72
Figure 3.12	Embedded software project organization. The nRF5 SDK and the implemented modules in the Data Acquisition Layer.	74
Figure 4.1	Experiments I and III - Total time per combined score recorded in the original dataset (baseline).	81
Figure 4.2	Experiment I and III - Distribution of combined scores per patients (36) in the original dataset (baseline).	82

Figure 4.3	Experiments I and III - Distribution of alarms per patients (36) in the original dataset (baseline).	82
Figure 4.4	Habib et al. (2016) Modified Local Emergency Algorithm, implemented in our simulator.	85
Figure 4.5	Habib et al. (2016) MLED and the self-adaptive procedure, implemented in our simulator.	86
Figure 4.6	Weighted monitoring error of combined scores by patients in configuration (A). The error was weighted by the distribution of monitored combined scores.	91
Figure 4.7	Weighted monitoring error of combined scores by patients in configuration (B). The error was weighted by the distribution of monitored combined scores.	91
Figure 4.8	Distribution of intervals in the reference system by combined scores.	92
Figure 4.9	Collecting energy measurements using the PPK-II and our hardware prototype.	97
Figure 4.10	Selection of target events of the temperature sensor in the PPK-II graph interface.	99
Figure 4.11	Energy requirements after 24h of monitoring in distinct patient profiles.	100
Figure 4.12	NRF Power Profiler Kit II chart. Acquisition of energy requirement of the self-adaptive procedure in the WKit BLE version.	104
Figure 4.13	Metrics to analyze the application of the four principles' effects on monitored data. Patients are sorted by the mean frequency (yellow) from the highest to the lowest.	105
Figure 4.14	Bézier and Linear distributions, the sum of all applied frequencies, and their differences (%) for all records.	107
Figure 4.15	WKit total energy consumption on all records using the ULPR radio. Comparison between the baseline system and our proposal.	108
Figure 4.16	WKit total energy consumption on all records using the BLE radio. Comparison between the baseline system and our proposal.	108
Figure 4.17	WKit total energy consumption on all records using the ULPR and BLE radios running the self-adaptive algorithm.	109
Figure 4.18	Number of oximeter reads classified by intervals between readouts <5s, >=5s, and reads weighted by energy cost.	109
Figure 4.19	WKit total energy cost for transmitting periodic data and alarms versus processing self-adaptive and alarm assessment processes. BLE version - self-adaptive algorithm.	110
Figure A.1	Research Methodology phases	135
Figure B.1	First studies of SH-Sens Wearable Kit and Base Station with NRF-52832 microcontroller. Aimed at infirmaries and COVID-19 patients.	138
Figure B.2	First studies of board designs. On the left is the WKit. On the right is the base station.	138
Figure B.3	First version of devices.	139
Figure B.4	Sensors and Micro-controllers utilized in our tests for the WKit version 2.	140
Figure B.5	WKit version 2 - circuit schema.	140

Figure B.6	WKit version 2 - device design.	141
Figure B.7	WKit and base station version 2. Breadboard tests.	142
Figure B.8	WKit in use with the MIT App Inventor mobile version and the BLE connection.	143
Figure C.1	Infirmary use case - Dashboard of monitored patients	145
Figure C.2	Mobile application logged user screens design	146
Figure C.3	Flutter Mobile application logged user screens	146
Figure C.4	Designed Architecture for supporting the user applications	147
Figure C.5	Mobile application for testing and debugging the BLE connection	148
Figure C.6	Flutter Mobile application - login and sensor configuration screens.	149
Figure D.1	Total time per combined score recorded in the original dataset (baseline).	151
Figure D.2	Elghers et al. (2014) - Local Emergency Detection algorithm.	154
Figure D.3	Harb et al. (2021) proposed algorithm.	154
Figure D.4	Comparison of the number of transitions of increasing combined scores in original patients data (blue) and captured by our proposal (orange) using NEWS-2.	156

List of tables

Table 2.1	Comparative table among previous work.	46
Table 3.1	The four principles utilized for the implementation of our solution proposal based on an early-warning scoring system.	49
Table 3.2	WKit and base station binary application size and memory requirements.	73
Table 4.1	Evaluation metrics.	79
Table 4.2	Simulator parameters.	80
Table 4.3	Parameters of our experiments, setups (A) and (B).	88
Table 4.4	Sum of reads, payload, and number of messages using setups (A) and (B).	89
Table 4.5	Comparative number of alarms to baseline.	89
Table 4.6	Total registered time (sec.) for each score by the baseline and our algorithm, configurations (A) and (B).	90
Table 4.7	Alarms for transition duration $\geq 5s$ (19,433 alarms) by time distance of experimental alarms from reference's alarms.	93
Table 4.8	Parameters of our self-adaptive algorithm.	99
Table 4.9	Total monitoring time (seconds) by combined scores. The highly unstable patient spent most monitoring time in lower combined scores (blue). In contrast, the unstable patient spent the most time at higher scores (red).	100
Table 4.10	Sensors and BLE transmission energy consumption.	100
Table 4.11	Progression of Bézier (Béz) and Linear (Lin) distribution of frequencies (Hz) by the dissimilarity index.	106
Table 4.12	Energy measures from WKit ultra-low-power radio (ULPR) and Bluetooth Low Energy radio (BLE).	107
Table D.1	Proposed self-configuration parameters based on NEWS-2 combined scores.	151
Table D.2	Sum of payloads and the number of messages considering 5 vital signs of 36 patients being monitored across 12 hours in our simulations.	155

List of Abbreviations

AL – Alarms

ALACC – Alarm’s accuracy

ALMDR – Alarm missed detection rate

ANOVA – Analysis of Variance

BF – Behavior function

BTEMP – Body temperature

CC – Clinical condition

CEWS – Combined early-warning score

DG – Data generation

EC – Energy consumption

EWS – Early-warning score

EWSS – Early-warning scoring system

freqMax – Maximum frequency

freqMin – Minimum frequency

HR – Heart rate

ICU – Intensive Care Unit

INCT – Instituto Nacional de Ciência e Tecnologia

indScore – Individual EWS score

IoT – Internet of Things

IoT-PMA – IoT-based patient monitoring application

LAC – Laboratory of Advanced Collaboration

LED – Local Emergency Detection

MINT – Monitoring integrity

MLED – Modified Local Emergency Detection

MQTT – Message Queuing Telemetry Transport

NEWS-2 – National Early-Warning Score 2

pLen – Period length

pCombScore – Period combined score

pScore – Period score

QBC – Quadratic Bézier Curve

QoS – Quality of Service

rCombScore – Readout combined score

RR – Respiratory rate

SBP – Systolic blood pressure

SH-Sens – Smart Health Sensing

SO – Arterial blood saturation

SpO2 – Arterial blood saturation

ULPR – Ultra-Low Power Radio

WBAN – Wireless Body Area Networks

WKit – Wearable Kit of sensors

WSN – Wireless Sensor Networks

*"Education is the most powerful weapon
which we can use in order to prepare our
youth for their role as leaders of tomorrow."*

**Nelson Mandela's speech, Madison Park High School, Boston, 23 June
1990.**

1

Introduction

In this Chapter, we present the motivation for the development of the thesis, followed by the research topic and problem statement definitions, goals, requirements, and research questions. Then, an introduction to the solution approach proposal and the contributions of this work are described. Finally, the organization of the following chapters is presented.

1.1

Motivation

The Internet of Things (IoT) proposes to connect the physical world to the Internet (ATZORI; IERA; MORABITO, 2010). It interconnects devices from industrial equipment, daily utensils, wearables, and environmental sensors. The number of physical objects connected to the Internet is unprecedented, and they are likely to become ubiquitous and pervasive (PERERA et al., 2013).

Information is gathered from sensors to increase efficiency, and monitor, manage, and control systems. Connected cities, homes, offices, vehicles, industrial plants, and people bring a new world of smart services and applications (PERERA et al., 2014). Several IoT solutions are based on wireless sensor networks (WSNs), such as monitoring air quality (GUPTA et al., 2011), water monitoring (SHU et al., 2017), and patient monitoring (AKKAŞ; SOKULLU; ÇETIN, 2020) (KHAN; PATHAN, 2018), to cite some.

However, the deployment of IoT-based applications in actual scenarios faces enormous challenges regarding reliability, availability, mobility, performance, management, scalability, interoperability, and security and privacy (AL-FUQAHA et al., 2015). One notable challenge is the massive data generated from sensors (LEQUEPEYS et al., 2021) (NASHIRUDDIN; RAKHMAWATI, 2022). The increasing number of sensors permanently generating raw data, in many cases redundant and with negligible significance for applications, deteriorates the computational infrastructure. Further, raw data must be analyzed, interpreted, and understood to create value (PERERA et al., 2014), which accrues crescent costs. Likewise, in other domains, a massive amount of data and low bandwidth are issues related to health-distributed systems, such as IoT-based health applications (MOURA et al., 2020).

For example, patient monitoring devices generate a huge amount of raw data. Monitoring the five major vital signs (body temperature, heart rate,

respiratory rate, systolic blood pressure, and oxygen saturation) at 1Hz, as it is typically utilized in hospitals (MOODY; MARK, 1996) (SAEED et al., 2002) (LIU; GÖRGES; JENKINS, 2012), generates 432,000 data points per patient daily. Monitoring a group of 20 patients for one year may produce more than 3 billion data points. Using a remote monitoring solution to track thousands or millions of patients during the COVID-19 pandemic or patients suffering from chronic diseases will reduce or hamper the performance of computer infrastructure as the utilization grows with the transmission of countless data points.

Another specific example of excessive data generation can be noticed in alarm systems of patient monitoring applications, where it is reported that a patient triggers hundreds of alarms per day in general wards. The number of alarms doubles in intensive care units (ICUs) (FERNANDES et al., 2019). In small ICU settings with a few beds, health teams should deal with one alarm every tens of seconds or less. This massive number of alarms causes an extra burden on health professionals, who may start ignoring and reacting automatically to alarms, decreasing service quality and exposing patients to risky situations (NGUYEN et al., 2019). In a scenario with patients being monitored remotely during their daily activities, as proposed by IoT health applications, the number of false alarms is expected to grow.

In addition, the deluge of data comes from battery-operated wireless sensors. Thus, another problem arises with the sensors' energy duration. Monitoring applications can spread sensors in a given area, even underwater (INDU et al., 2014). Moreover, sensors can be mobile and loaded on the human body in wearables (MISHRA; RASOOL, 2019). Especially in patient monitoring systems, wearable sensor devices must be small or tiny, and batteries supply the power (KIMURA; LATIFI, 2005). However, the battery's capacity is proportional to its dimensions (SANISLAV et al., 2021). In such applications, short battery life can hamper and avoid using IoT technology (RAGHUNATHAN; GANERIWAL; SRIVASTAVA, 2006). Thus, a data reduction strategy associated with the best energy use in battery-powered devices is very timely (NASHIRUDDIN; RAKHMAWATI, 2022).

Observing traditional patient monitoring routines in infirmaries, we found the use of early-warning scoring systems (GERRY et al., 2017). Healthcare teams regularly utilize a scoring system that evaluates patients' health conditions based on several health markers, such as the vital signs (PHYSICIANS LONDON, 2017). An individual score is given for each vital sign according to a range of vital sign values. Individual scores can be combined to represent a health condition better. Calculated combined early-warning scores

assist in establishing bed visits frequency by health teams and decision-making of patients' treatment (WONG et al., 2015). Moreover, scores categorize patients in a crescent risk level. Thus, changes in risk levels can generate alerts.

The infirmity procedure of the better the score, the lower the frequency of bed visits can be associated with other simple propositions, such as the more similar the monitored data, the lower the sampling rate. This intelligence can be embedded in smart wearable sensors to assess patients monitored data in real time. So, **we argue that using early-warning scores governed by simple principles can lead to a novel approach and tools to enhance the efficient use of computational, communication, and power resources, promoting data and alarm reductions and energy savings in IoT-based patient monitoring applications.** Therefore, this work focuses on reducing excessive data generation and alarms and promoting energy savings in battery-powered devices in those applications.

To test our argument, we developed software experiments emulating a remote patient monitoring application and a hardware prototype to assess the potential benefits of data, alarm reductions, and energy savings with actual energy measurements. Moreover, once data availability on monitoring is diminished, evaluating the effects of reductions on the operational fidelity of monitored events is also necessary (RAGHUNATHAN et al., 2002). Thus, three quality metrics were also devised: monitored patient health status quality, alarm missed detection rate, and alarm accuracy.

Based on our experiments, our approach was able to reduce samples, transmissions, and alarms drastically, with a small loss in monitoring quality and alarms' accuracy and missed detection rate. Discussions about the reduction mechanisms and energy requirements of battery-powered patient monitoring applications are also addressed. Finally, future work is highlighted, such as combining the proposed approach with other data-reduction and energy-savings techniques.

1.2

Research Topic Definition

My main research interest was investigating IoT-based solutions to maintain and improve people's health. At the beginning of 2020, the world was impacted by the COVID-19 pandemic. Then, our research group was provoked to develop solutions that could help to mitigate the burden faced by health infrastructure.

Therefore, we proposed a comprehensive conceptual architecture of a patient monitoring system for COVID-19 (PAGANELLI et al., 2022). The

architecture considers various use cases and addresses challenges in large-scale IoT solutions, including ensuring interoperability, enhancing reliability, promoting scalability, and safeguarding privacy.

Moreover, the architecture emphasizes functionalities embedded in the monitoring devices, such as supporting flexible and configurable scoring systems. The design of the main functionalities and embedded mechanisms to process the scoring system and self-adaptive schemes are original and were conceived for this thesis. The architecture study (PAGANELLI et al., 2022) revealed some addressed issues (massive data generation and energy constraints in sensor devices) and helped find the solution presented in this thesis.

In parallel, a systematic literature review of real-time data analysis methods used in health monitoring solutions was conducted (PAGANELLI et al., 2022). The review gave a broad and deep understanding of current research, such as the main statistical and machine learning algorithms utilized in real-time for diagnosis, prediction, and anomaly detection in health monitoring applications. It also described the main sensors, properties, and datasets utilized in health monitoring literature. Among the reviewed studies, one recent work by Harb et al. (HARB et al., 2021) applied a scoring system to promote data reductions in a patient monitoring solution in sync with our preliminary solution conceptualization. Although our architecture proposal referred to the same scoring system utilized in infirmaries as (HARB et al., 2021), the approaches were different. Our work utilized combined and individual scores while (HARB et al., 2021) study utilized only individual scores.

Finally, considering the findings from the design and implementation of our conceptual architecture and literature review studies, the general investigation topic was set as IoT-based patient monitoring applications using early-warning scores.

1.3

Problem Statement

Most of the current IoT-based patient monitoring applications can usually generate massive data flows and excessive alarms, which increases the energy consumption in monitoring devices. Then, these sensor devices powered by batteries must implement energy-efficient solutions to extend their running times.

The amount of generated data can compromise the efficiency of the communication and computational infrastructures as these systems grow. Raw sensor data is usually less significant and expressive or even redundant, wasting resources and increasing solution costs. Specifically, alarms can occur at an

excessive frequency and can be redundant, becoming irrelevant. Therefore, the major problems to be addressed in this work are:

- **Problem 1:** Massive production of sensor data in IoT-based patient monitoring applications that track vital signs.
- **Problem 2:** An excessive number of insignificant and redundant alarms.
- **Problem 3:** Low energy autonomy of battery-enabled IoT devices in patient monitoring applications that may be even worsened by Problems 1 and 2.

1.4

Goals and Research Questions

The goal of this thesis is to propose a method based on contextual conditions inferred by early-warning scores to auto-regulate the amount of data generated and promote data reduction. In addition, it also addresses the problems of excessive alarms and energy efficiency in IoT-based patient monitoring applications dependent on wearable sensor devices.

It is beyond the scope of this work to address security, privacy, loss of data in transmissions, and validity of measurements performed by wearable sensors. Hence, in essence, the research questions addressed by this thesis are:

- **RQ1.** How can combined early-warning scores be used to change monitoring frequencies and, as a result, reduce data generation, excessive alarms, and energy consumption?
- **RQ2.** What are the effects of changing the frequencies and reducing sensor data generation on monitoring quality?
- **RQ3.** How can the solution reduce the number of alarms? What are the effects of adaptive frequencies on alarms' accuracy and alarms' missed detection rate?
- **RQ4.** How can the added scoring system embedded into the monitoring device promote energy efficiency?
- **RQ5.** How can the proposed solution be implemented and embedded in an IoT wearable device and integrated into a patient monitoring application?

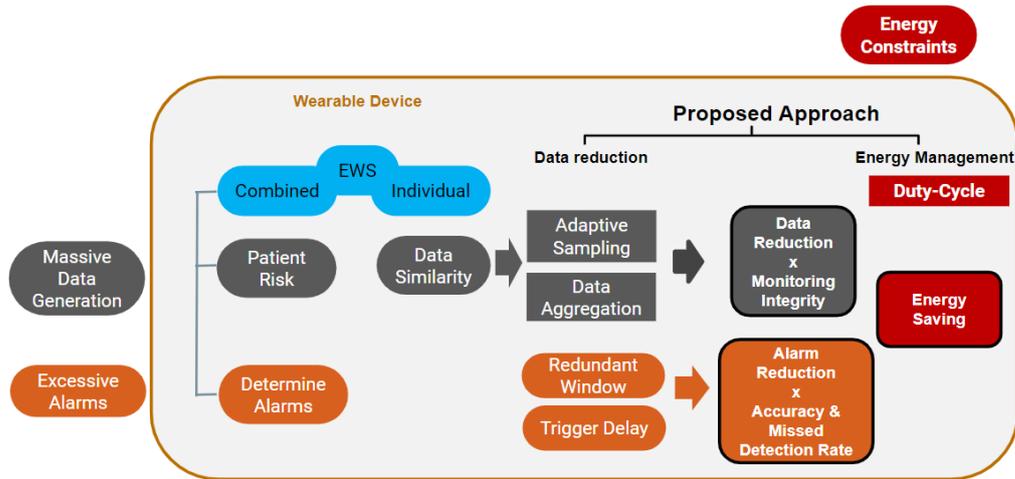


Figure 1.1: Overview of the proposed approach for adaptive monitoring. It uses combined and individual early-warning scores in an embedded solution applying adaptive sampling and data aggregation to regulate data rates.

1.5 Solution Approach Components

Figure 1.1 shows an overview of our proposed approach. The figure depicts the main components of the adaptive software embedded in the IoT sensor device and the addressed issues.

Outside the golden rounded-corner square, the three addressed issues by our thesis. At the right top, the data reduction and energy management approaches focus on adaptive sampling, data aggregation, and using the duty cycle to improve energy efficiency.

Data reduction is implemented using adaptive sampling and data aggregation controlled by early-warning score (EWS) guidelines. These guidelines utilize combined and individual scores. Individual scores categorize vital signs by ranges of values. Additionally, a combined score is a classification of a group of individual scores that can infer a patient's health risk. The latter is used in our proposal to determine frequency ranges to perform the main tasks in the sensor devices, such as data sampling. In contrast, individual scores are utilized to verify data similarity over time. Data similarity is another driver in regulating sampling rates. Increased combined scores determine alarms because it means a deterioration in a patient's overall health condition. Alarms are filtered using two techniques: delays and a redundant alarm lockout window. Then, reduced sampling data and alarms promote energy savings. However, reductions in samplings and alarms are limited by the monitoring fidelity that should be compatible with the patient's health condition and actionable alarms' accuracy, and alarms missed detection rate. We managed to extend the battery life

and the device's autonomy by regulating the tasks' execution frequencies and employing a duty-cycle strategy in the microcontroller and sensor drives.

A detailed description of all these features is provided in Chapter 3

1.6

Main Contributions

The contributions of this thesis are either fully described within this document or else are referenced with a brief description and citations to published papers. While the published papers may cover themes outside the main objectives of our thesis, they have provided significant contributions to my research and the broader field of Health Informatics.

For example, the conceptual architecture of IoT-based patient monitoring applications in the context of COVID-19 and our systematic literature review about real-time data analysis in health monitoring solutions are comprehensive pieces of research that provided valuable insights and defined the scope of problems addressed by the thesis, and which served as a foundation for our proposed solution.

Moreover, most of the experiments described in Chapter 4 were presented at international conferences and specialized workshops and published in their proceedings. The complete list of publications is presented in Section 1.7.

Our main contributions are listed below:

- **C1.** A solution abstraction description based on four principles using combined and individual early warning scores assessed in real-time to govern a self-adaptive algorithm that copes with massive sensor data generation, excessive alarms, and energy efficiency. **Chapter 3.**
- **C2.** The design and implementation of a hardware prototype with our proposal embedded that provides actual energy measurements in distinct use case scenarios and provides pieces of evidence of the feasibility of our solution. **Appendix B.**
- **C3.** A description of several mechanisms to reduce the excessive number of alarms produced by patient monitoring systems. **Section 3.3**
- **C4.** The definition of three error metrics to assess the potential loss in monitoring and alarm fidelities when reducing data samplings in the monitoring of vital signs that compose an early-warning system. **Sections 3.2 and 3.3**
- **C5.** A rich discussion about the application of our proposal in IoT-based patient monitoring applications based on a set of experiments

demonstrating its potential to promote data and alarm reductions and energy savings. **Chapter 4.**

- **C6.** A description of a conceptual architecture of an end-to-end IoT-based patient monitoring solution, which could be integrated with our proposed approach. **Section 3.5 and (PAGANELLI et al., 2022).**
- **C7.** A systematic literature review of real-time data analysis performed in health monitoring systems. **(PAGANELLI et al., 2022).**

1.7

Publications

1. Paganelli A. I., Branco A., Endler M., Velmovitsky P. E., Miranda P., Morita P. P., Alencar P., Cowan, D. (2021). **IoT-Based COVID-19 Health Monitoring System: Context, Early Warning and Self-Adaptation.** *In 2021 IEEE International Conference on Big Data (Big Data)*. Orlando, FL, USA, 2021, pp. 5975-5977, doi: 10.1109/Big-Data52589.2021.9671361.
2. Paganelli A. I., Velmovitsky P. E., Miranda P., Branco A., Alencar P., Cowan D., Endler M. & Morita P. P. (2022). **A conceptual IoT-based early-warning architecture for remote monitoring of COVID-19 patients in wards and at home.** *Internet of Things*, 18, 100399. <https://doi.org/10.1016/j.iot.2021.100399>
3. Paganelli A. I., Mondéjar A. G., da Silva A. C., Silva-Calpa G., Teixeira M. F., Carvalho F., Raposo A., Endler M. (2022) **Real-time data analysis in health monitoring systems: a comprehensive systematic literature review.** *Journal of Biomedical Informatics*, 104009 vol(127), March, 2022. doi: 10.1016/j.jbi.2022.104009.
4. Paganelli A. I., Branco A., Endler M., Velmovitsky P. E., Morita P. P., Alencar P., Cowan, D. (2022). **A novel self-adaptive method for improving patient monitoring with composite early-warning scores.** *In 2022 IEEE International Conference on Big Data (Big Data)*. Osaka, Japan, 2022, pp. 201-208, doi: 10.1109/Big-Data55660.2022.10021046.
5. Paganelli A. I., Sarmiento, A., Branco A., Endler M., Nascimento N., Alencar P., Cowan, D. (2022). **Assessing energy consumption in data acquisition from smart wearable sensors in IoT-Based health applications** *In 2022 IEEE International Conference on Big*

Data (Big Data) - BIGEACPS'22 Workshop. Osaka, Japan, 2022, pp. 2882-2885, doi: 10.1109/BigData55660.2022.10020572.

1.8

Thesis Organization

This Chapter introduces the motivation, problem statement, research questions, and the proposed solution of this research. Chapter 2 reviews underlying technologies, concepts, and related work. Chapter 3 describes the proposed solutions to answer our research questions one by one.

Moreover, Chapter 4 applied our proposed solution and addressed problems 1 and 2 in experiments I (Section 4.3) and Problem 3 in experiments II (Section 4.4) and III (Section 4.5). Furthermore, experiment III also shows the effects of the applied principles that rule our self-adaptive algorithm on the monitoring data. Yet, Appendix B enriched our experiments with an implementation of our proposal in an IoT hardware platform that not only attests to the proposal's feasibility but also provides actual energy measures to run the simulations with real patient monitoring data from the public datasets. A general discussion on the application of our proposal closes Chapter 4.

Finally, Chapter 5 summarizes the main findings and limitations and points out future work.

2

Background & Related work

This Chapter summarizes the technologies and concepts related to our research on IoT-based patient monitoring applications using early-warning scores. It also reviews related work and summarizes the gaps in the literature that will be explored by our proposal.

2.1

Patient Monitoring in Hospitals

Patients being monitored in hospitals are usually bedridden and carry sensors connected to multiparameter monitors via wires. The multiparameter monitors are connected to the electrical power outlet and take readings of physiological markers every second or less, showing the acquired values read on a screen. Embedded in the systems of these multiparameter monitors is the ability to issue audible and visual alarms based on threshold values defined for each health marker. Alarms are manually and individually configured for each patient. In addition, multiparameter monitors can be networked and can send monitored data to central panels, such as a workstation with video monitors in the on-call staff room (GARDNER; SHABOT, 2001).

Although the multiparameter monitors can be interconnected with the monitoring centers, the monitoring of the clinical condition is carried out through visits to the patient's bed by the nursing team. (PRUTSACHAIN-IMIT; JANTHONG; TUANROPI, 2020). Moreover, general wards normally are not instrumentalized, and monitoring depends on the nursing team shifts, usually once or twice a day (LEENEN et al., 2020). Health data are recorded in forms that facilitate the calculation of individual and combined scores following a scoring system. The result of the scores and the assessment of the health team can modify the routine of visits to the patient's bed or even indicate the need for an immediate intervention by a specialized team (WONG et al., 2015).

These monitoring devices are proprietary and do not allow customization to include new features, such as automatic calculation by a scoring system. Then, the scoring system annotations are made manually and depend on the correct data entry to record and process this information (KYRIACOS; JELSMA; JORDAN, 2011).

In sophisticated systems, the information collected from the monitoring can feed databases and be integrated with other data from the patient's medical

records, such as personal data, health history, tests performed, clinical notes, medication prescriptions, and other information relevant to the treatment of individuals (GARDNER; SHABOT, 2001).

The cost for the operation and maintenance of these monitoring systems is very high, as it depends on expensive equipment and generates a huge amount of data that must be manually or automatically transmitted, processed, and stored in the hospital infrastructure. In addition, multiparameter monitors are intended for cases where the patients remain bedridden most of the time, as wires connect the sensors to the individuals. In this way, the patient is only monitored when she/he is inside the hospital environment, occupying a bed, drastically increasing the treatment cost due to the entire infrastructure involved. Therefore, traditional hospital systems for monitoring physiological data are intended only for the most severe cases, usually in ICU beds (LEENEN et al., 2020).

In summary, the benefits of hospital monitoring systems are:

1. The possibility of monitoring physiological parameters in a controlled environment increases the accuracy of the data collected, as the sensor connections are wired to the monitor equipment.
2. The patient is standstill most of the time, allowing stability in the fixation of the sensors.
3. The power supply of multiparameter monitors is permanent as they are connected directly to the electrical grid.
4. Possibility of configuring visual and audible alarms based on threshold values per sensor.
5. Visualization of monitored data in real-time.
6. Data transmission to monitoring centers and storage in databases.

The disadvantages are:

1. High cost of equipment and the related infrastructure.
2. Poor mobility of monitored patients.
3. Proprietary systems with low customization.
4. Devices do not support score systems automatically.
5. Generally, they are exclusive to the most severe cases of patients admitted to ICUs.

2.2

Internet of Things and Patient Monitoring Applications

Several technological advances supported the IoT revolution, such as the miniaturization of electronic components, wireless communication technologies and protocols, the development of low-cost sensors and actuators, and the progress of flexible components and circuits (KHAN; PATHAN, 2018) (ELANGO; MUNIANDI, 2020).

Following this trend, wearable devices with physiological sensors, also known as biosensors, and radio communication, aim to obtain and analyze human physiological and pathological information online and in real-time (LU et al., 2020), providing ways to monitor health parameters remotely in a pervasive and non-invasive way (ELANGO; MUNIANDI, 2020). Then, remote healthcare systems have become associated with IoT technologies.

In addition to the development of underlying technologies, population aging imposes more pressure on the already overloaded health system infrastructure. This fact leverages the search for alternative out-of-hospital solutions, such as those provided by IoT-based health applications (PHILIP et al., 2021).

Moreover, increasing evidence from studies on the remote monitoring of patients has been produced in recent years, as can be seen in the number of reviews and surveys on IoT-based health monitoring systems released recently. Mainly, they present an extensive overview of the solutions (DIAS; CUNHA, 2018) (QI et al., 2017) (THILAKARATHNE; KAGITA; GADEKALLU, 2020), or focus on target groups or specific monitoring aspects such as applied data mining techniques (BANAEI; AHMED; LOUTFI, 2013), fog computing for healthcare (MOURA et al., 2020) (SILVA; JUNIOR, 2018) (MUTLAG et al., 2019), Internet edge computing (HARTMANN; HASHMI; IMRAN, 2019), in epidemics (MOHAMMADZADEH et al., 2020), non-communicable diseases (KRISTOFFERSSON; LINDÉN, 2020), wearable technologies and products for baby monitoring (HASAN; NEGULESCU, 2020), wearable devices in health care (LU et al., 2020), general technical aspects (G.S; P.B, 2020), clinical evidence in hospitalized adults (LEENEN et al., 2020), and healthcare services (USAK et al., 2019). We also performed a systematic literature review focused on real-time data analysis methods in IoT-based health monitoring systems (PAGANELLI et al., 2022)

IoT health monitoring systems can have many purposes, such as fitness, wellness, specific activities, and medical (DIAS; CUNHA, 2018), and they are linked to hospitals, clinics, and smart homes (MOURA et al., 2020). They are also very convenient for people suffering from long-term diseases, such as diabetes and hypertension, to cite a few. These patients need constant

monitoring by a doctor to adjust their treatments. Then, IoT technologies are appropriate for remote patient monitoring (GÓMEZ; OVIEDO; ZHUMA, 2016).

IoT-based patient monitoring applications (IoT-PMA) are presumed to minimize healthcare costs, increase patient care, and enhance patient service delivery (THILAKARATHNE; KAGITA; GADEKALLU, 2020). Moreover, recording vital-sign data allows extracting detailed reports of patients' health status (Zainol et al., 2019). The frequent data acquisition may detect changes in symptoms during treatment and evaluate its efficacy, contributing to the individualization of the treatment plan (LU et al., 2020). Also, real-time monitoring of patients has the potential to detect risk conditions early (Nubenthan; Ravichelvan, 2017) and trigger alarms (HARTMANN; HASHMI; IMRAN, 2019). Thus, these applications aim to act as a preventive tool detecting or predicting health degradation and symptoms changes (KRISTOFFERSON; LINDÉN, 2020), being a strong ally of the modern medicine 4P model: preventive, predictive, personalized, and participatory (LU et al., 2020).

The advantages of IoT-based patient monitoring applications can be summarized:

1. Low cost of its components.
2. Mobility enhanced by wireless connections and batteries.
3. Potential for application in several use cases outside clinical settings.
4. Use of existing communication infrastructure (smartphones and Internet gateways).

The disadvantages are:

1. Low autonomy of battery-powered sensors.
2. Production of an excessive amount of data.
3. Cost for the operation and maintenance of these systems, either by changing or recharging batteries or by the cost of transmission, processing, and storage of data.
4. Quality of captured data (low-cost sensors, uncontrolled environment).

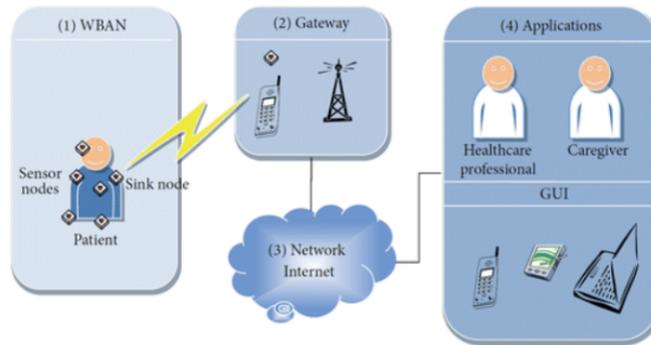


Figure 2.1: A typical Wireless Body Area Network architecture. Image source: (FILIPPE et al., 2015).

2.2.1

IoT-based Patient Monitoring Applications Architecture

IoT-PMAs combine wearable sensors, information, and communication technologies, massive data generation, big data algorithm applications, machine learning, and artificial intelligence (RGHIOUI et al., 2020). Most IoT-PMAs have a multi-layer architecture composed of sensor, fog, and cloud layers (BAIG; GHOLAMHOSSEINI, 2013) (SILVA; JUNIOR, 2018) (MOURA et al., 2020). However, a fourth application layer is also usually seen in the literature (SILVA; JUNIOR, 2018) (QI et al., 2017).

The sensor layer comprises the sensors on the patient's body and a device that has a radio to transmit the acquired data to an internet gateway, which may also be a portable device such as a smartphone with 3G, 4G, and WiFi access. Each sensor node is composed of a sensing unit that connects the component to the physical world, a microprocessor or microcontroller, a short-range radio, and a power supply unit (RAGHUNATHAN et al., 2002).

Sometimes, the sensors attached to the body are interconnected and form a local network with a portable central node that controls the sensing network and the transmissions to the Internet. The interconnection of these sensors is known as a Body Area Network. If each sensing unit also has a microcontroller, the network is named a Body Sensor Network (DIAS; CUNHA, 2018). These networks are also named Wireless Body Area Networks (WBANs) (FILIPPE et al., 2015) (KHAN; PATHAN, 2018). Figure 2.1 depicts a typical WBAN architecture.

The sensor layer is one hop distant from the Internet edge. The Internet edge is the network infrastructure that provides connectivity to the internet space, composed of gateways, edge routers, switches, and firewalls (ESCAMILLA-AMBROSIO et al., 2018). Around the Internet's edge is where fog computing occurs (MOURA et al., 2020). Fog computing can be viewed

as an Internet access point fitted with enhanced networking and computing capabilities (MOURA et al., 2020), a layer between the sensing devices and the cloud (SILVA; JUNIOR, 2018) (YOUSEFPOUR et al., 2019). The cloud layer is a pool of configurable and virtualized software and hardware resources located in data centers serving the Internet (ESCAMILLA-AMBROSIO et al., 2018). Finally, the end-user applications can be executed at the sensor, fog, and cloud layers (SILVA; JUNIOR, 2018). A detailed description of edge, fog, and cloud definitions can be found in Escamilla-Ambrosio et al. (ESCAMILLA-AMBROSIO et al., 2018), and Yousefpour et al. (YOUSEFPOUR et al., 2019).

The scale of computational resources is associated with the cloud, fog, and sensor layers. On the cloud side, there is the presence of high availability of computing resources and high power consumption. At the same time, the fog layer provides the intermediate availability of computing resources and lower-power computing. The sensor devices extend fog computing, storage, and networking capabilities, though with limited availability of computing resources and very low power consumption. The computation performed by the sensor devices is known as mist computing or things computing (YOUSEFPOUR et al., 2019), and it has the objective of lowering application latency, providing local responses, increasing network throughput (i.e., using data compression) and enhancing the independence of a solution (BARIK et al., 2018).

In many IoT-PMAs, only data acquisition and transmission services are performed on wearable devices. However, although wearable-based sensor devices are characterized by having low energy, low bandwidth, low processing power, and constrained hardware nodes (MOURA et al., 2020), following Banaee et al. (BANAE; AHMED; LOUTFI, 2013), any wearable sensor solution for health monitoring must implement a simple rule-based method to recognize patterns, anomalies, and specific events based on predefined rules and conditions. Nevertheless, several studies implemented methods for recognizing abnormal health conditions based on sophisticated machine learning algorithms, such as ensemble classifiers (ANI et al., 2017), neural networks (FOUAD et al., 2020), and other techniques (BANAE; AHMED; LOUTFI, 2013). In IoT-PMAs, these algorithms are usually executed remotely to take advantage of powerful computers at the fog (MOURA et al., 2020) or cloud layers (HARTMANN; HASHMI; IMRAN, 2019) (BAIG; GHOLAMHOSSEINI, 2013) (Darcini S.; Isravel; Silas, 2020). However, the detection of risk and emergencies is based on the prior identification of abnormal conditions and patterns of sensed data, demonstrating a physiological deterioration in a patient's health (VISHNUPRASAD et al., 2014). Recent studies (EDDAHCHOURI et

al., 2022) show that continuous wireless monitoring of vital signs is associated with reducing unplanned ICU admissions and rapid response team calls in hospitalized patients.

Ideally, identifying such risk situations should be performed closer to where the signals were acquired to reduce latency and increase reliability. In IoT-PMAs, wearable devices with microcontrollers and processing capabilities can capture vital sign data. A low-cost solution that permanently monitors and automatically calculates scores opens new possibilities for care and treatment inside and outside hospitals. However, continuous remote monitoring of patients can generate a massive amount of data (MICHARD; SESSLER, 2018). Regardless patient's clinical status, data can be significantly redundant, as patients may be left without substantial changes for long periods.

2.2.2

Power Management and Data Reduction in IoT-PMAs

Monitoring physiological parameters constantly requires a robust energy supply unit (YU; LI; ZHAO, 2021). However, wearable sensors impose limitations on energy consumption in IoT-PMAs (ZHOU et al., 2014) (LOUKATOS et al., 2021). The battery capacity is proportional to its weight and size (SANISLAV et al., 2021), and sensors are expected to be small (KIMURA; LATIFI, 2005) with limited battery capacities.

Moreover, there is a fundamental trade-off between prolonged monitoring with a low accuracy versus a higher one during a shorter period. Note that in the latter case, the total error is potentially unknown because it may have a period where there is no monitoring. For example, it can be impossible to recharge or replace a battery, or while the battery is recharging or being replaced, there is no monitoring.

Many strategies were devised to enhance power management in wireless sensor networks that can be applied to WBANs and IoT-PMAs. These strategies received different classifications, such as hardware and software design, network protocols, and middleware services. In addition, energy-saving schemes can also be viewed as radio optimization, battery repletion, duty cycles (sleep/wake-up) schemes, energy-efficient routing, and data reduction (SHU et al., 2017).

A comprehensive taxonomy of energy-saving strategies in wireless sensor networks (WSNs) is presented by Rault et al. (RAULT; BOUABDALLAH; CHALLAL, 2014) as shown in Figure 2.2. The authors also review the physical and network layers approaches regarding energy efficiency in WSNs according to application needs.

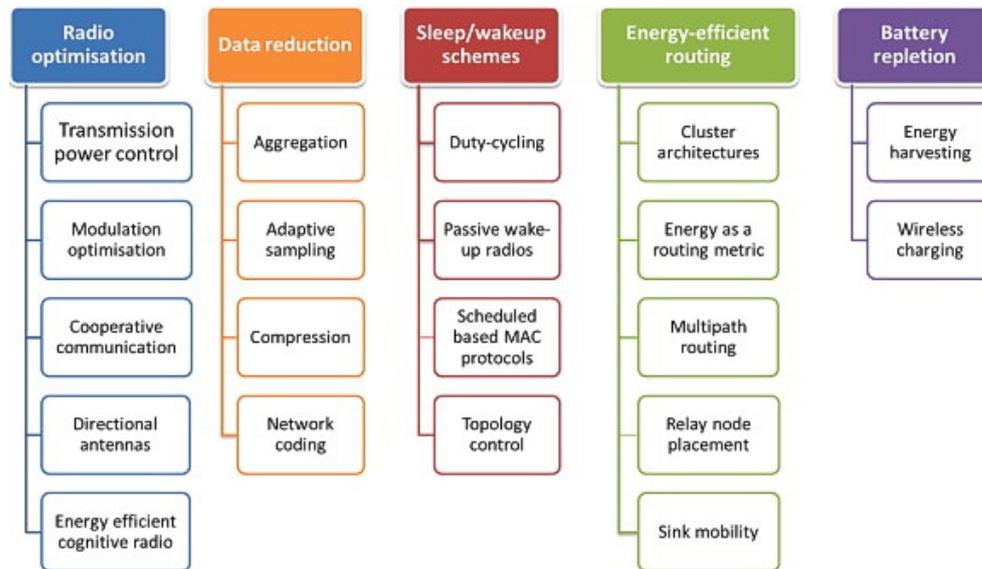


Figure 2.2: Energy-efficient strategy taxonomy proposed by (RAULT; BOUABDALLAH; CHALLAL, 2014).

Although multiple strategies can be applied to the same application, each represents a wide field of study exploring trade-offs between energy consumption, system performance, and operational fidelity (RAGHUNATHAN et al., 2002). Some of them, such as energy-efficient routing, potentially would not bring significant gains in WBANs because these networks have one or a maximum of two hops (ZHOU et al., 2014). This and other techniques emphasize maximizing the lifetime of an entire network composed of thousands of nodes. However, since our research is focused on the data reduction approach and IoT-PMAs supported by WBANs with a few nodes, we will focus our review on studies that utilized related techniques.

Considering that the data analysis objective is to transform raw data acquired by sensors to a high-level knowledge representation (BANAEI; AHMED; LOUFI, 2013), moving part of the analysis task close to where the data acquisition is performed can reduce the network traffic (HARTMANN; HASHMI; IMRAN, 2019). Thus, data reduction in transmissions will save energy on power-constrained devices (HARTMANN; HASHMI; IMRAN, 2019) (MUTLAG et al., 2019). Data reduction strategies try to reduce unneeded samples and sensing tasks. Energy requirements of sampling and transmissions are relatively high in IoT-PMAs (RAULT; BOUABDALLAH; CHALLAL, 2014). Moreover, in health applications, some sensors, such as an accelerometer, may have higher energy requirements than radio transmissions because data acquisition can take much longer than performing radio operations (HABIB et al., 2016).

In data aggregation, data fusion is utilized for reducing payload in

transmissions, getting the average, minimum, maximum, or another statistical measure instead of sending all the measures, for example. This technique may reduce the accuracy of measurements. Most data aggregation studies in WSNs also focus on large networks with different topologies and several hops (RAJAGOPALAN; VARSHNEY, 2006). However, some data fusion principles can be applied to IoT-PMAs, for example, consolidating data points in a high-level representation such as a category.

Another approach, named adaptive sampling, regulates sampling rates according to the coverage of information and precision. Then, potentially reducing data generation when not required by the application. In addition, using the network coding method, a linear combination of several packets in broadcast scenarios avoids sending many copies of each packet. Finally, data compression reduces the number of bits necessary to represent some information. Then, fewer data will be transmitted (RAULT; BOUABDALLAH; CHALLAL, 2014). Because processing energy cost, in general, is lower than transmitting, compression is an energy efficiency strategy. However, in most cases, it does not apply to sensor devices because compression algorithms' computational costs exceed resources available (KIMURA; LATIFI, 2005).

Furthermore, a practical approach to developing an energy-efficient device for any WSN involves designing hardware and software components. An overview of such factors can be found in (RAGHUNATHAN et al., 2002). Although we considered some of those factors in our proposal, such as the duty cycle of all components and tests with different radio protocols, our primary focus is on adaptive sampling and data aggregation for reducing operations in any sensor node platform aimed at IoT-PMAs.

IoT-based studies that utilized data reduction techniques such as adaptive sampling will be present in Sections 2.4 and 2.5.

2.3

Early-Warning Scores and Vital Signs Monitoring

Monitoring vital signs in patients is essential to assess and identify a worsening health condition (KYRIACOS; JELSMA; JORDAN, 2011) (FU et al., 2020). Subtle changes in vital signs occur 8 to 24 hours before a life-threatening event, such as ICU admission and cardiac arrest (WEENK et al., 2017). Vital signs fluctuation may form specific patterns that require increased attention and notifications to health teams (VARSHNEY, 2008).

In the late 90s, early-warning score (EWS) references based on vital signs were created to reduce predictable deaths in infirmaries (MORGAN et al., 1997). They are intended to streamline responses, guide interventions, and

monitor the effectiveness of interventions (KYRIACOS; JELSMA; JORDAN, 2011) (SUBBE, 2001). They became an essential tool for assessing patients' conditions, adequate bed visits, and predicting outcomes (WEENK et al., 2017) such as ICU admissions (GOLDHILL et al., 2005), cardiac arrest (QUARTERMAN et al., 2005), and mortality (SMITH et al., 2014). The National Institute for Health and Clinical Excellence (NICE) recommended the utilization of EWS in the UK in 2007 (FANG; LIM; BALAKRISHNAN, 2020). Then, EWS based on vital signs monitoring became a routine in hospitals as part of a rapid response system that triggers alarms and pre-planned responses by healthcare professionals.

Since the introduction of EWS in hospitals, several EWS guidelines have been proposed, such as the Acute Physiology and Chronic Health Evaluation II (APACHE II), Rapid Emergency Medicine Score (REMS), Sequential Organ Failure Assessment (SOFA), Multiple Organ Dysfunction Score (MODS), Pre-disposition, Infection, Response, and Organ Dysfunction (PIRO), Mortality in Emergency Department Sepsis (MEDS), Sepsis in Obstetrics Score, Search Out Severity (SOS), Leed's Early-Warning Scores (LEWS), Patient-at-Risk score (PARS), Triage Early-Warning Scores (TEWS), Modified Search Out Severity (M-SOS), Modified Early Warning Score (MEWS), and the National Early-Warning Score II system (NEWS-2) (BADRINATH et al., 2018) (CHER-ANAKHORN et al., 2016) (ALBRIGHT et al., 2014) (KYRIACOS; JELSMA; JORDAN, 2011) (SUBBE, 2001) (PHYSICIANS LONDON, 2017) (P et al., 2022).

Each early-warning system utilizes a set of health parameters to stratify patients. Figure 2.3 shows the NEWS-2 created by the Royal College of Physicians in England (PHYSICIANS LONDON, 2017). The first version of NEWS was launched in 2012 and reviewed based on feedback and application in practice in December 2017. It became the unique EWS utilized in acute hospitals and ambulances in the UK (WILLIAMS, 2019). Then, NEWS-2 is widely used in hospitals worldwide for detecting patients at risk of worsening (GERRY et al., 2017). It classifies each vital sign in ranges of values that vary from normal to abnormal, receiving a score. Score zero represents the normal status, while three is a highly abnormal condition.

Further, NEWS-2 can be an effective tool to avoid harm and costs, avoiding death and ICU admissions (YE et al., 2019). These NEWS-2 scores are typically calculated three times a day by nursing teams, and at least once every 12h in hospitals in England (NICE, 2007 July), which may not be sufficient to detect early clinical worsening (SAAB et al., 2021). A delay in detecting worsening is known to be associated with increased mortality (WEENK et al.,

Physiological parameter	Score						
	3	2	1	0	1	2	3
Respiration rate (per minute)	≤8		9–11	12–20		21–24	≥25
SpO ₂ Scale 1 (%)	≤91	92–93	94–95	≥96			
SpO ₂ Scale 2 (%)	≤83	84–85	86–87	88–92 ≥93 on air	93–94 on oxygen	95–96 on oxygen	≥97 on oxygen
Air or oxygen?		Oxygen		Air			
Systolic blood pressure (mmHg)	≤90	91–100	101–110	111–219			≥220
Pulse (per minute)	≤40		41–50	51–90	91–110	111–130	≥131
Consciousness				Alert			CVPU
Temperature (°C)	≤35.0		35.1–36.0	36.1–38.0	38.1–39.0	≥39.1	

Figure 2.3: NEWS-2 table guidelines.

2017).

Individual scores are summed, and the result is combined in groups to form a combined score. Combined scores are aggregated weighted scoring systems normally used in tracking and trigger procedures (GAO et al., 2007) and vastly utilized in studies (FU et al., 2020). Then, the combined score defines the periodicity of the successive measurements (WONG et al., 2015) as shown in Figure 2.4. Additionally, early-warning scoring systems help to reduce the excessive number of false alarms generated by threshold-based alarms (SUBBE; DULLER; BELLOMO, 2017). EWS, in general, and NEWS-2, in particular, are simple to implement and present low computational complexity. Moreover, when a scoring system such as the NEWS-2 is used embedded in sensing devices for constant monitoring, previous studies pointed to potential savings in energy and data transmission (ELGHERS; MAKHOUL; LAIYMANI, 2014) (HABIB et al., 2016) (HARB et al., 2021). These studies will be described in Section 2.5. Finally, comprehensive reviews of using NEWS-2 in clinical sets can be found in (MANN et al., 2021) (GERRY et al., 2020) (FANG; LIM; BALAKRISHNAN, 2020).

2.3.1

Excessive Alarms and Alarm Fatigue

Alarms should capture the attention of healthcare professionals assuring awareness, assessment, and supportive intervention. Thus, alarm systems are designed to cause cognitive distress. However, alarms can be actionable and

NEWS score	Frequency of monitoring	Clinical response
0	Minimum 12 hourly	<ul style="list-style-type: none"> Continue routine NEWS monitoring
Total 1-4	Minimum 4-6 hourly	<ul style="list-style-type: none"> Inform registered nurse, who must assess the patient Registered nurse decides whether increased frequency of monitoring and/or escalation of care is required
3 in single parameter	Minimum 1 hourly	<ul style="list-style-type: none"> Registered nurse to inform medical team caring for the patient, who will review and decide whether escalation of care is necessary
Total 5 or more Urgent response threshold	Minimum 1 hourly	<ul style="list-style-type: none"> Registered nurse to immediately inform the medical team caring for the patient Registered nurse to request urgent assessment by a clinician or team with core competencies in the care of acutely ill patients Provide clinical care in an environment with monitoring facilities
Total 7 or more Emergency response threshold	Continuous monitoring of vital signs	<ul style="list-style-type: none"> Registered nurse to immediately inform the medical team caring for the patient – this should be at least at specialist registrar level Emergency assessment by a team with critical care competencies, including practitioner(s) with advanced airway management skills Consider transfer of care to a level 2 or 3 clinical care facility, ie higher-dependency unit or ICU Clinical care in an environment with monitoring facilities

Figure 2.4: Combined Scores based on NEWS-2 individual scores.

non-actionable. Actionable alarms are triggered by valid abnormal physiological states requiring attention and interaction. Non-actionable alarms do not require immediate attention or any attention at all (HRAVNAK et al., 2018). It is estimated that only 5-13% of alarms in patient monitoring systems are actionable (RUPPEL; FUNK; WHITTEMORE, 2018). The excessive number of non-actionable alarms causes what is known as alarm fatigue.

Alarm fatigue has been defined as a decline in a health team's response to alarms due to excessive alarms, sensory overload, desensitization, and apathy, among other occupational variables (NGUYEN et al., 2019). It was considered the number one problem with device technology, which poses an enormous risk to patient safety (CVACH, 2012) because valid alarms that need intervention are missed (RUPPEL; FUNK; WHITTEMORE, 2018).

Studies reported that hospital monitoring systems generate, on average, 350 alarms per bed per day. In Intensive Care Units (ICUs), the average climbs to 771 per bed per day (JONES, 2014). Some studies reported that 80% to 99% of alarms are false or clinically insignificant (FERNANDES et al., 2019). In a pediatric ICU, 87% of alarms were non-actionable, reaching 99% in pediatric wards (BONAFIDE et al., 2015). This fact may encourage healthcare teams

to turn off alarm apparatus (CVACH, 2012).

To address the problem, a multidisciplinary approach is necessary involving clinicians, computer scientists, industry, and regulatory agencies (HRAVNAK et al., 2018). Monitors are designed with too-tight thresholds for high sensitivity and do not miss a true abnormal event. However, this configuration also generates several clinically insignificant alarms. While sensitivity is high, reaching 97%, specificity is low at 58% (CVACH, 2012). The alarms may be true, but they are viewed as a "nuisance" by health teams, jeopardizing quick responses when needed. As the number of non-actionable alarms increases, the delays also increase (HRAVNAK et al., 2018).

Therefore, solutions for mitigating the problem of excessive redundant and insignificant alarms are necessary to be implemented in patient monitoring devices. Using multiple parameters, rate of changes and better signal quality in conjunction with a smart algorithm and/or artificial intelligence systems may mitigate the problem (CVACH, 2012). Prioritization of alerts and the development of sophisticated alerts, and including end-user opinion in alert selection were also suggested as means to reduce the number of non-actionable alarms (KANE-GILL et al., 2017).

Other strategies focus on notifying only known stage changes, changing the baseline frequently, loosening alarm threshold parameters, and delaying announcements, which would avoid short-lived transitions. Efficient alarm management also should consider professional training and standardization to configure alarms better and bring confidence to caregivers (HRAVNAK et al., 2018).

Some machine learning techniques were applied (FERNANDES et al., 2019) (CHEN et al., 2016) to address this problem. For example, (CHEN et al., 2016) utilized random forest models. The authors considered an alarm consecutive when two or more vital sign alerts exceeded thresholds in intervals shorter than 40 seconds between them and a vital sign alert event when it lasted at least 3 minutes, with 2/3 of the consecutive observed values crossing the threshold. In addition, reviewers utilized a scoring system to label alarms. In some cases, four reviewers assessed alarm events to determine if the alarm was real or an artifact. Generalization of the generated model may be limited because it was not tested using different monitoring platforms. Further, for building the models, the method is highly dependent on labeled data and powerful computation, which is unavailable in sensor devices.

Fernandes et al. (FERNANDES et al., 2019) proposed an automatic reasoning algorithm to aggregate or not alarms, provide a false alarm probability metric, and select the best targets of caregivers to receive alarms. Using a

message-queue producer and consumer mechanism, the algorithm basically distributes alarms to caregivers using aggregation and avoids sending redundant alarms in a time window. The solution also is dependent on cloud computing to process the algorithm mechanism.

In another study (SCHMID et al., 2015), using a rule-based algorithm, two anesthesiologists, supported by video recordings, re-evaluated a database of intra-operative elective cardiac surgery monitoring data to find relevant alarms and rate the severity of these events. The alarms were classified as true/false, clinically relevant/irrelevant, and their importance as a warning, serious, or life-threatening. Beyond that, the authors defined a mild violation when monitored values exceeded less than 4% beyond the configured thresholds. An adaptive time delay algorithm was devised based on this analysis. The method reduces in 73.51% the rate of false positive alarms. The adaptive algorithm removed alarms triggered by negligible events characterized by short-duration mild violations. Then, the higher the violation, the shorter the period, and vice-versa. Although using specialists to label data, this solution could be implemented on sensor devices because it requires low computational resources.

A broad review of interventions to reduce alarm frequency in clinical settings can be found in (PAINÉ et al., 2015). In addition, an overview of medical device technologies regarding alarm management can be found in (IMHOFF et al., 2009). A recent investigation of alarm fatigue solutions and future directions is described in (HRAVNAK et al., 2018).

The machine learning studies are very specific to the context where they were used and are difficult to generalize. Furthermore, although there is no consensus and direct metrics to measure alarm fatigue, it is assumed that reducing the number of alarms and/or improving the specificity of alarms will mitigate alarm fatigue (WINTERS et al., 2018). Simple, generic, direct methods that go beyond setting thresholds and be easily interpreted and configured by health teams for managing the excessive number of alarms in IoT-PMAs are still rare in the literature.

2.4

Self-Adaptive Solutions

Self-adaptation refers to a system's ability to modify its runtime behavior to achieve its objectives. A self-adaptive system continuously monitors itself and, based on the data it gathers and the analysis of this data, decides whether adaptation is required. These systems need to make changes at runtime and fulfill the system requirements satisfactorily.

Moreover, self-adaptive software has the capacity to adapt its behavior according to newly realized situations from the surrounding environment autonomously. An open adaptive system can introduce new execution plans during runtime, while a closed adaptive system is self-contained and does not support addition behaviors. Conditional expressions are a form of self-adaptation. The expression is evaluated and alters its behavior depending on expression outcomes (OREIZY et al., 1999).

Oreizy et al. (OREIZY et al., 1999) formulated a set of questions that must be answered when developing a self-adaptive application, as follows:

- Under what conditions does the system undergo adaptation?
- Should the system be open-adaptive or closed-adaptive?
- What type of autonomy should be supported?
- Is the adaptation cost-effective? In which conditions?
- How often is adaptation considered?
- What kind of information must be collected to make adaptation decisions?
- How accurate and current must the information be?

We intend to answer these questions in Chapter 3, where we introduce our proposed solution.

Self-adaptive behavior can be applied to adaptive sampling, a strategy utilized in many studies in WSNs (NASHIRUDDIN; RAKHMAWATI, 2022). However, there is a trade-off between the sampling cost and monitoring error rate (RAHIMI; SAFABAKHSH, 2010). Applying the adaptive sampling to the time dimension by modifying the sampling rate according to the signal history. However, when the signal changes quickly, reductions in fidelity are acceptable (RAGHUNATHAN; GANERIWAL; SRIVASTAVA, 2006).

2.5

Related Work

In the IoT-PMAs domain, some studies (ELGHERS; MAKHOUL; LAIYMANI, 2014) (HABIB et al., 2016) (IDA et al., 2020) (HARB et al., 2021) explored the use of scoring systems to support local data analysis and promote self-adaptation to reduce transmissions using the data reduction approach.

Elghers et al. (ELGHERS; MAKHOUL; LAIYMANI, 2014) propose reducing data transmission and detecting emergencies on the nodes. The authors claimed to be one of the first studies to promote sensor raw data transmission reduction and detection of emergencies locally on the node in

Algorithm 2 Emergency Detection and Adpating Algo *LED**

Require: m (1 round = m periods), R_{max} (maximum sampling rate).

Ensure: R_t (instantaneous sampling speed).

$R_t \leftarrow R_{max}$

2: **while** $Energy > 0$ **do**

for each round **do**

4: Run *LED* (Emergency Detection) for each period m
 compute SR , SF and F .

6: find F_t

8: **if** $F < F_t$ **then**

$R_t \leftarrow BV(F, F_t, r^0, S_{max})$ (BV behavior function).

else

10: $S_t \leftarrow R_{max}$

end if

12: **end for**

end while

endalgorithmic

Figure 2.5: Elghers et al. (2014) self-adaptive algorithm.

patient monitoring applications. Each sensor processes the NEWS-2 locally based on monitored data. The sensor sends data at fixed periods. The authors proposed the Local Emergency Detection algorithm (LED), which sends the first collected measurement in each period, followed by abnormal measures between periods. Normality is checked using the NEWS-2, and any value classified with a score greater than zero was considered abnormal.

Figure 2.5 extracted from (ELGHERS; MAKHOUL; LAIYMANI, 2014) presents the proposed algorithm. In order to optimize the algorithm further, the work utilized a self-adaptive algorithm to adjust sensor sampling frequencies based on the variance of acquired values. It uses one-way ANOVA and a Fisher distribution $F(J-1, N-1)$, where J is the number of periods and N is the number of data points. To calculate F , the within-period variation (SR) and the between-period variation (SF) are found. Then, F will be the result of SF divided by $J - 1$ over SR divided by $N - J$. Given a false-rejection probability α for the Fisher distribution, $F_t = F_{1-\alpha}(J-1, N-J)$. So, F can be tested. The level of criticality is represented by a behavior function (BV). The BV is a function of the Quadratic Bézier curve (QBC). If the mean and variance are different among periods $F > F_t$, the method configures the sensor to its maximum sampling rate. However, if the periods have the same values, the frequency is reduced to the minimum rate or some frequency in between according to the F value and the QBC.

Then, the sampling rate can increase if the mean and variance are not equal but still below the given threshold. In addition, the speed of adjusting the

frequencies should follow the classification of patients' health determined by a physician manually. The adaptive algorithm is compared to a baseline (sending all data and running without the adaptive algorithm). Results demonstrated a variable benefit, which can reduce 87% of data transmissions from the baseline.

Furthermore, in a hypothetical energy model, the authors assume that reading and sending each value would consume 0.1 and 0.7 units, respectively. Then, they calculated the energy requirements and demonstrated that the proposed algorithm would extend battery life.

Habib et al. (HABIB et al., 2016) utilized a similar model as (ELGHERS; MAKHOUL; LAIYMANI, 2014) to self-adapt frequencies and NEWS-2 scores to select data to be sent, developing the Modified LED (MLED) algorithm. However, to avoid redundancies, the algorithm only transmits values that have distinct scores from previous ones consecutively. For example, a sequence of SO values (98, 97, 97, 98, 95, 95, 93, 95, 96, 98) will receive scores (0, 0, 0, 0, 1, 1, 2, 1, 0, 0) following NEWS-2, and MLED will send (98, 95, 93, 95). The algorithm was able, in some instances, to reduce data transmissions by 50% more than the LED algorithm using MLED and adaptive sampling. Additionally, following a hypothetical energy model, the authors assumed that each data acquisition would cost 0.3 energy units while transmission cost 1 unit. Then, following the experiments, the proposed algorithm would consume 3 to 4 times less energy than the algorithm proposed by (ELGHERS; MAKHOUL; LAIYMANI, 2014).

Ida et al. (IDA et al., 2020) explored the use of early-warning scores in hospitals to personalize monitoring of patient's vital signs, and to enhance time-response in case of problems, proposing a self-adaptive approach. The proposal utilizes a Pub-Sub communication model and gateways to process early-warning score information locally. The gateway utilizes a local broker and a configuration service where medical staff can personalize data processes according to the patient's situation. In addition, the gateway runs a virtual sensor agent responsible for calculating the scores, sending notifications in case of problems, and reconfiguring the system. The virtual sensor subscribes to its configuration topic of the configuration service and runs accordingly.

Moreover, a score manager agent receives the scores from all virtual sensors to calculate the patient's risk. It sends notifications if needed based on calculated scores. Finally, a configuration manager can increase saving data frequencies if scores exceed predefined values. It also allows the administrator to reduce frequencies when scores are within the predefined thresholds and stable. Experts define the frequencies. The gateway was implemented using a Raspberry Pi 3 loaded with a Mosquitto MQTT Broker and InfluxDB. The authors simulated hypothetical scenarios with and without the self-adaptive

model monitoring the temperature parameter. They analyzed the memory and CPU usage in both scenarios. The authors concluded that the non-adaptive model could result in the unnecessary use of gateway processors and memory, affecting the reaction time, especially in emergencies.

Harb et al. (HARB et al., 2021) also utilized the LED algorithm but proposed to perform a linear regression sending the two coefficients along with critical values within a period and reconstructing the values on the receiver side. The authors highlighted that the (HABIB et al., 2016) solution only sends critical data to the coordinator node. Then, monitoring a healthy patient will not produce enough data for archiving historical data, and ANOVA and Fisher tests are not sensitive for patients with medium criticality. It is assumed that patients at low and high risk yield more redundancy, while in medium risk, the variation is more prominent. For each round composed of two or more periods, where each period corresponds to a set of reading values, the degree of stability among the set of NEWS-2 scores within those periods is calculated. The level of stability varies from 0 to 100, where 0 represents full stability while 100 is the opposite. Then, the new frequency is calculated as a percent of the highest frequency down to a minimum according to the stability level found in the previous round. The algorithm was compared to the solution proposed by (HABIB et al., 2016) using different round sizes, period sizes, and minimum frequencies. The results demonstrated that the new experimental algorithm could reduce transmissions from 39% to 94% depending on the analyzed vital sign.

All these studies mainly utilized individual scores to promote sensor data transmission reductions and energy savings. Moreover, Elghers et al., Habib et al., and Harb et al. used actual data from patients extracted from the MIMIC database (MOODY; MARK, 1996) (GOLDBERGER et al., 2000). Table 2.1 summarizes the main characteristics of the described studies, except by Ida et al. (IDA et al., 2020) because this study utilized only hypothetical data scenarios and parameters based on health staff decisions, not being clear about threshold values, notifications, and monitoring frequencies. Moreover, that study aimed at running on gateways not embedded in wearable devices.

2.6

Summary and Gaps

This Chapter reviewed essential concepts of patient monitoring applications in hospitals and outside clinical settings supported by IoT devices. Additionally, it reviewed IoT-PMAs underlying technologies, proposed architectures, data reduction, and power management methods. Moreover, early-

Table 2.1: Comparative table among previous work.

	Elghers et al. (2014)	Habib et al. (2016)	Harb et al. (2021)
EWS	Individual	Individual	Individual
Data reduction Strategy	Send first read in fixed intervals. Between these intervals send only abnormal values (non-zero scores)	Same as Elghers et al. but between intervals send non-repeated consecutive abnormal values based on EWS scores.	Same as Elghers et al., and also send coefficients of linear regression
Self-adaptive Strategy	ANOVA, Fisher Test based on read values.	ANOVA, Fisher Test based on read values.	Degree of stability index based on individual scores
Adaptive Distribution	Quadratic Bézier Curves with static distribution	Quadratic Bézier Curves with static distribution	Fixed, proportional to the stability index.
Patient's condition	Manually assigned	Manually assigned	N/A
Utilized Sensors	HR	RR and BTemp	HR, RR, SBP, SO
Main parameters	2 patients, 1 hour, freq. max=6, freq. min=1, risk = [0.4 0.9]	2 patients, 70 periods (2h), freq. max=50, freq. min=10, risk = [0.4 0.9]	72 patients, 100,000 reads, period size=3600s, round size=2. 3 patients for energy tests. low, medium, high criticality.
Data reduction results	Measured by rounds, from 0% to 87%.	BTemp: 75.8% when risk= 0.9, and 85.7% when risk=0.4, RR: 63.5% and 71.8%, respectively	HR=94%, SysBP=52%, RR=39%, SpO2=40% compared to Elghers et al.
Error metrics	Visual, emergency detection	Distribution of scores	Visual, data regeneration
Emergency Detection	Individual scores > 0	Individual scores > 0	Individual scores > 0
Alarm Procedures	No	No	No
Energy model	Hypothetical	Hypothetical	Hypothetical
Energy gains	From 0 to 7 times compared to baseline	Up to 7x RR from the baseline.	HR=6x, SBP=6x, RR=5x, SpO2=4x from baseline

warning scores and vital signs monitoring were explained, and the issues related to excessive alarms during patient monitoring, such as alarm fatigue. In addition, basic concepts of self-adaptive solutions and related work using early-warning scores on IoT-PMAs were reviewed.

Finally, some gaps in the literature were identified, which this thesis will explore. The list below presents those topics.

1. Direct use of combined scores in adaptive sampling and data aggregation;
2. Combining self-adaptive methods with alarm reduction strategies;
3. Risk inference through combined scores performed in real-time during monitoring;
4. The use of objective error metrics to assess self-adaptive effects on monitoring integrity.
5. Energy models based on actual measurements from hardware prototypes applied to a self-adaptive scoring-based solution.

3 Proposed Approach

This Chapter introduces our method to cope with excessive sensor data generation (DG), alarms (AL), and energy consumption (EC) of wearable sensor devices in IoT-PMA using an early-warning scoring system. The solution’s objective is to **minimize(DG, AL, EC)**. However, the potential gains of the solution are limited by operational fidelity parameters such as monitoring information integrity (MINT), alarms accuracy (ALACC), and alarms missed detection rate (ALMDR), which will be explained in this Chapter. To recap from Chapter 1, an overview of our proposal is shown in Figure 3.1.

In our proposed solution, we must highlight that we assume that patients will carry the IoT wearable device permanently, sensors provide valid measurements, data is protected against violations, and communication links are reliable. The following sections explain our proposed approach according to our Research Questions.

3.1 Combined and individual early-warning scores to promote data and alarm reductions and save energy in IoT-PMAs - RQ1

Our proposal takes advantage of early-warning scoring systems (EWSS) widely utilized in infirmaries to promote data aggregation and adaptive sampling by configuring, in real time, parameters that regulate changes in frequencies of wearable sensor operations, such as data sampling, processing, and transmissions.

Our solution assumes that the lower the combined early-warning score (CEWS), the better the patient’s health condition and vice-versa as it is utilized in infirmaries (SUBBE, 2001) (PHYSICIANS LONDON, 2017). Remem-

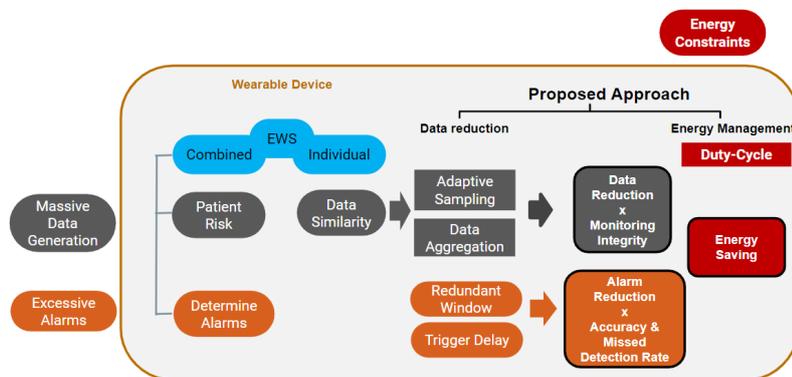


Figure 3.1: Proposal approach overview - recap.

Table 3.1: The four principles utilized for the implementation of our solution proposal based on an early-warning scoring system.

P1	The higher the combined early-warning score, the higher the monitoring frequency, and vice-versa.
P2	The more similar the monitored data along the time, the lower the monitoring frequency, and vice-versa.
P3	In higher combined early-warning scores, frequencies will increase faster, and vice-versa.
P4	The higher the combined early-warning score, the shorter the interval to re-assess the patient's condition, and vice-versa.

bering that a CEWS is a classification of ranges of the result of summed individual early-warning scores, which can capture different health markers to infer patients' health condition better. Thus, CEWS is also known in the literature as aggregated weighted scores (FU et al., 2020). Using the CEWS, the proposed self-adaptive approach follows four principles as described in the next paragraphs and summarized in Table 3.1.

Principle (P1) defines that the more severe the patient's health condition, represented by a higher CEWS, the higher the monitoring frequency and vice-versa; and (P2) the higher the dissimilarity of monitored data along the time, the higher the monitoring frequency and vice-versa. Beyond that, the monitored data can be very similar and redundant, and the patient is either in good health or not. Thus, principle P1 precedes P2 and limits the P2 frequency variation. Principle P2 will take effect within ranges of frequencies pre-determined by principle P1.

Moreover, the severity of the patient's health condition will also affect the distribution of frequencies within the configured ranges. In more severe conditions, frequencies will increase faster with the slight dissimilarity of monitored data and vice-versa (P3).

For the implementation of P1, the patient's health condition is evaluated regularly but in variable intervals defined as period lengths (pLen). pLen also varies according to the patient's health condition. Then, principle (P4) is defined as the more severe the patient's health condition, the shorter pLen and vice-versa. Shorter re-evaluation intervals allow the system realizes changes more quickly. So, there is a higher probability of reducing frequencies whenever possible and capturing any alterations in health conditions. Therefore principle P1, in conjunction with P3 and P4, aim at turning the monitoring more sensitive when a patient's health worsens, while P2 increases the sensibility when data is less redundant.

Following these principles, our solution proposal runs as follows. During

pLen, several samples are acquired by each sensor at a current sampling rate. At each reading, the corresponding individual score (indScore) of the sample is found, and its occurrence is registered. At the end of pLen, a data aggregation process is performed to infer the individual score of the period for each sensor as follows: if any of the scores occurred in at least 50% of the readings, the period aggregated score is set to that score, from the highest to the lowest score. Otherwise, the period aggregated score receives the score with the highest occurrence. Then, having the period aggregated score (pScore) of each sensor, the period combined score (pCombScore) can be found to infer a patient's health condition based on the utilized classification of the EWSS.

$$summedScore := \sum^{sensor} pScore_{sensor} \quad (3-1)$$

$$sensor = \{HR, SpO2, RR, SBP, BTemp\}$$

$$pCombScore := ScoringSystem(SummedScore) \quad (3-2)$$

Having the pCombScore, the device behavior in the next period can be defined. If pCombScore differs from the previous one, the following parameters are adjusted: pLen, minimum frequency (*freqMin*), maximum frequency (*freqMax*), and risk factor (*r0*). Parameters *freqMin* and *freqMax* define the minimum and maximum frequencies of sampling for a specific patient's health condition determined by the CEWS. At the same time, the *r0* is used to define the frequency distribution between *freqMin* and *freqMax* calculated using a Quadratic Bézier Curve (QBC). However, if pCombScore does not differ from the previous one, the above parameters do not change.

Then, after adjusting those parameters if needed, a dissimilarity index is calculated anyways. The described sequence is shown in the pseudo-code below.

```

1 if current pCombScore != previous pCombScore {
2   Update pLen, freqMin, freqMax, r0 according to current
   pCombScore
3 }
4 dIdx = CalculateDissimilarityIndex()

```

Monitoring data dissimilarity is observed using the set of pScores for each vital sign (HR, SO, RR, SBP, BTemp) calculated in the current and previous periods at the end of pLen. The dissimilarity index is computed using

the absolute difference of pScore over n periods. For each period, there are s pScores, where s is the number of monitored health parameters. Therefore, there will be n periods, and for each period, s scores. The set of pScores of each period is transformed to a string with s positions. So, pScore of health parameter one in string position one, pScore of health parameter two in string position two, and so on. Then, the Hamming distance across the n periods is utilized to calculate the dissimilarity index. Nonetheless, instead of considering each distinct character as a distance of one, it is weighted with the absolute difference between the score integer values.

The weighted distance is accumulated, period(t) is compared to the period($t-1$), period($t-1$) to period($t-2$), and so on up to period($t - n - 1$). Thus, the maximum weighted distance will be:

$$\text{MaximumWeightedDistance} := \text{maxScoreValue} * s * (n - 1) \quad (3-3)$$

maxScoreValue is the highest score in the scoring system, s is the number of sensors, and n is the number of periods.

Most EWSS utilize scores from zero to three. Considering that substantial variations in scores are rare, even for very unstable patients for the proposed frequency ranges, the maximum weighted distance is limited to $(2 * s)$. Remembering that this limit can be configured to attend different contexts. Yet, increasing this limit, frequencies will vary slower, and it would be necessary to achieve higher dissimilarity to reach the maximum frequency of the configured range. Finally, the weighted distances are normalized between 0 and 1, dividing the sum of the weighted distance by $(2 * s)$. If the result exceeds one, it will be set to one. Otherwise, it is a fraction between 0 and 1.

Figure 3.2 presents an example of how the dissimilarity index is calculated using the NEWS-2 reference. In this example, four periods ($n=4$) and five health parameters ($s=5$) exist. Each health parameter has one pScore for each period. In period(T), all pScore equal zero, while in period($T-1$), RR has a pScore of 1 and SO of 2. The differences of each pScore, one by one, for each health parameter are computed, resulting in the sum of differences. For example, for the health parameter SO, the difference between T and $T-1$ is 2, between $T-1$ and $T-2$, is 0, and finally, between $T-2$ and $T-3$ is 1. Then, the total difference is $2 + 0 + 1 = 3$.

All health parameters' sum of differences is also summed, resulting in the sum of the sum of the differences ($ssDiff$). In our example, it equals 7. To obtain the dissimilarity index, if the $ssDiff$ is less than 10 ($2 * s$, $s=5$, the number of health parameters), it is divided by 10. Otherwise, the dissimilarity

Periods	HR	RR	SO	TE	BP	Combined. Score
T-3	0	0	1	2	0	1
T-2	0	0	2	1	0	1
T-1	0	1	2	0	0	1
T	0	0	0	0	0	0
Σ Diffs	0	2	3	2	0	$\Sigma \Sigma$ diffs = 7
Dissimilarity index						if $\Sigma \Sigma$ diffs < 10, \Rightarrow $\Sigma \Sigma$ diffs / 10, Otherwise, 1.0

Figure 3.2: Dissimilarity index computation example.

index will be one.

Finally, the sampling frequency for the next period can be defined using freqMin , freqMax , r_0 , and the dissimilarity index ($dIdx$) applied to a QBC. As shown in Figure 3.3, to define the QBC, it is necessary three points: P_0 is the origin, which will correspond to the minimum frequency (freqMin) and minimum dissimilarity ($dIdx=0$); P_2 is given by the maximum frequency (freqMax) and maximum dissimilarity ($dIdx=1$), while P_1 is the risk factor (r_0), a point in the second diagonal of the imaginary parallelogram. Note that freqMin , freqMax , and r_0 are parameters configured according to the period CEWS ($pCombScore$), and $dIdx$ is calculated as explained above. A Behavior Function was implemented based on the algorithm explained in (LAIYMANI; MAKHOUL, 2013) to return a frequency by $dIdx$ in the Bézier curve.

The Bézier distribution is convenient (LAIYMANI; MAKHOUL, 2013) because if the score is high, r_0 will be high, and a small variation in $dIdx$ will quickly increase the sampling frequency. At the same time, if the risk is low, it is necessary to have a larger variation in $dIdx$ to increase the sampling frequency from the minimum frequency. Note that if r_0 equals 0.5, the curve will represent a linear distribution. An r_0 value above 0.5 will accelerate frequencies with small dissimilarities, while an r_0 below 0.5 will decelerate frequency increases.

Now, it is described how the CEWS copes with alarms. Alarms are controlled by the CEWSs calculated at a sampling rate. Differently from the period combined scores ($pCombScore$) calculated at the end of each period in $pLEN$ intervals. Every time a cycle of sensor reads occurs, the readout combined score ($rCombScore$) is calculated. If this $rCombScore$ is greater than the previous one, an alarm will eventually be triggered.

Nonetheless, two strategies were developed to avoid insignificant and

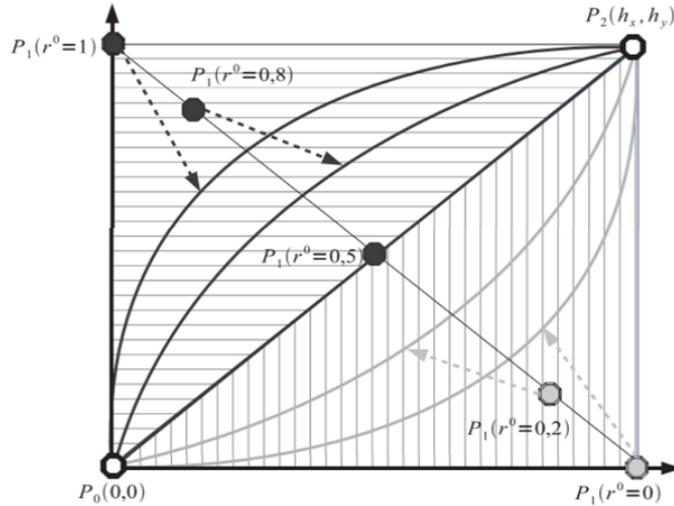


Figure 3.3: Bézier Quadratic Curve properties. P_1 is the patient's risk defined by r_0 , which will adjust the shape of the distribution. Source: (LAIYMANI; MAKHOUL, 2013).

redundant alarms. These strategies are explained in Section 3.3. The flowchart 3.4 depicts the self-adaptive algorithm described above integrated with the alarm procedure call.

Finally, to reduce our solution's energy requirements, a duty cycle approach was implemented in our prototype in addition to the strategies described above. A controlling loop process wakes up periodically and checks the queue of tasks to be performed. Then, it performs the queued tasks, if any. Otherwise, it configures the microcontroller to sleep mode. Additionally, sensor drivers were rewritten to enter sleep mode when not reading data. All these features are explained in detail in Section 3.4.

In summary, the described approach will reduce data sampling because the highest possible frequency will be used when the health parameters characterize the highest CEWS. However, conditions vary over time. If the health condition improves, data sampling will be reduced. Moreover, if the patient is constantly in the same health condition, data represented by $indScores$ will be very similar, and the frequencies will also be reduced. Finally, the number of alarms tends to be reduced because they are triggered based on CEWS ($rCombScore$), not in individual variations. So potential individual alarms are aggregated by the combined score. Furthermore, two strategies to avoid negligible and redundant alarms were implemented.

In the next sections, every feature described above will be detailed and analyzed from the perspective of solving the selected problems.

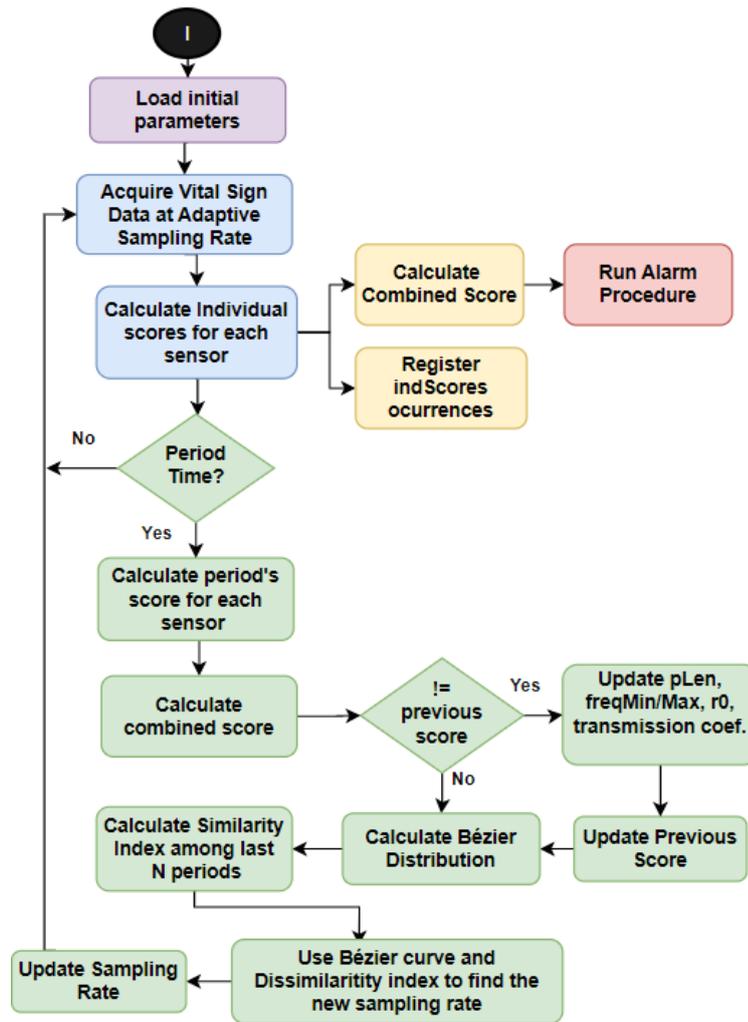


Figure 3.4: Proposed self-adaptive algorithm flowchart. Steps in blue, yellow, and red are performed at sampling rates, while in green are executed at the period time (pLen).

3.2

Reducing sensor data generation and measuring monitoring quality - RQ2

This section details and formalizes how the adaptive frequency using combined and individual scores of an EWSS will tackle the problem of excessive sensor data generation (DG). In addition, a quality metric is defined to assess the effects of reducing frequencies and sensor data generation on monitoring patients' health conditions according to the early-warning scoring system.

3.2.1

Sensor data generation - problem definition

An IoT-PMA can be defined as a system composed of individuals (INDs) being monitored by a device (DEV) loaded with sensors that collect s health parameters (HP) from IND at a rate (rx).

For simplification, let's assume that each *IND* will carry only one *DEV*, all *DEV*s collect the same number of *HP*s, and all *HP*s are acquired at the same *rx* for the same *IND*. Then, every *IND* has a *DEV* collecting *s* *HP*s at *rx* rate as shown in expression (3-4).

$$\forall IND, \exists DEV : DEV \text{ collects } s \text{ HPs at an } rx \text{ rate.} \quad (3-4)$$

In our problem specification, the number of *IND* is sufficiently large and tends to increase, *s* is fixed among different *IND*. So, sensor data generation (DG) can be defined as:

$$DG := x * s * rx \quad (3-5)$$

, where *x* is the number of individuals using the IoT-PMA, *s* is the number of *HP*s monitored by a *DEV*, and *rx* is the sampling rate. The sampling rate is defined as the number of samplings *n* over time. Then, $rx = n / t$.

$$DG := x * s * n/t \quad (3-6)$$

The variables *x* and *s* are assumed to be given and tend to grow as more diseases can be monitored, and more sensors are added to IoT-PMAs. Nonetheless, the variable *rx* depends on the events to be observed by the monitoring application.

The adaptive sampling will be applied by a self-adaptive procedure based on the EWSS to reduce *n* with minimal loss in the monitoring fidelity. Following the principles P1-P4 (Table 3.1), our proposal expects to reduce *DG*.

Observing the principles, *n* will be reduced compared to a baseline system running at a maximum sampling rate whenever the combined score is lower than the highest score. In addition, the calculation of combined scores occurs in shorter intervals when the combined score is higher. Then, the solution will capture quicker changes in the patient's health. Moreover, even if the combined score does not change or improve, principle P2 assesses the similarity of monitored data. If the similarity is reasonable, reflecting some redundancy in monitored data, *n* will also be reduced. These mechanisms lead to *DG* reduction.

In the next subsections, it is defined the utilized scoring system based on NEWS-2 guidelines (PHYSICIANS LONDON, 2017) and the self-adaptive solution design.

3.2.2

Individual and Combined Scores based on NEWS-2

Early-warning scoring systems utilized in infirmaries are based on physiological parameters such as heart rate (HR), respiratory rate (RR), body temperature (BTEMP), systolic blood pressure (SBP), and arterial oxygen saturation (SO), defined as health parameters (HP). The NEWS-2 defines two other parameters, one used to adequate the range of nominal values of SO for chronic hypercapnic respiratory failure (HCRF) patients, and the other one to observe the consciousness level (CL). HCRF patients have a second SO scale exclusively for these cases.

Each HP receives a nominal value (v_{HP}) at a given monitored time (t). The v_{HP} can also be called sampling or reading value, and it can be represented by $v_{HP}(t)$. For example, the HR of 80 bpm at time $t=1$ can be represented by $v_{HR}(1) := 80$ bpm. Likewise, $v_{SO}(1) := 98$ represents SO at time $t=1$ at 98%.

Clinical conditions CL and $HCRF$ are specific parameters verified and assigned by health professionals. The latter is a condition that does not vary over time. The parameter CL can be categorized as true (normal) or false (altered), $v_{CL}(t) := \text{true} \mid \text{false}$. Additionally, the $HCRF$ can also be represented as true (present) or false (absent), $v_{HCRF} := \text{true} \mid \text{false}$.

An EWSS determines a score (s) for unique ranges of nominal values for each HP. Each range has a maximum v_{HPmax} and a minimum v_{HPmin} . Different ranges for each HP do not overlap. In EWSS, each HP has r ranges of values. For example, using the arterial blood saturation (SO) health parameter, r equals 4:

```

If  $v_{HCRF} == \text{false}$ 
  If  $v_{SO}(t) > 0.96$ , then  $s_{SO}(t) := 0$ 
  If  $v_{SO}(t) \leq 0.96$  and  $v_{SO}(t) > 0.93$ , then  $s_{SO}(t) := 1$ .
  If  $v_{SO}(t) \leq 0.92$  and  $v_{SO}(t) > 0.88$ , then  $s_{SO}(t) := 2$ .
  If  $v_{SO}(t) \leq 0.88$ , then  $s_{SO}(t) := 3$ .

```

(*) HCRF = chronic hypercapnic respiratory failure. If **this** condition is present, another scale of values should be used.

In our experiments, we assumed that v_{HCRF} equals false and $v_{CL}(t)$ equals true. Then, it was considered the scale I for SO ("SpO2 scale 1 (%)") as shown in Figure 2.3).

Each $v_{HP}(t)$ corresponds to an individual score within the EWSS, represented by $s_{HP}(t)$. The scores and ranges of values for each HP are determined by specialists and clinical research, establishing the relationship between a de-

gree of health risk and the score assigned. It is not the scope of this work to question, investigate or improve these relationships. However, the proposed solution can serve as a tool for investigations in this direction. Below follows the summary of utilized health parameters in our proposal based on the NEWS-2.

$$v_{HP}(t) \in \mathbb{R}^+, \quad (3-7)$$

$$HP = \{HR, SO, RR, BTEMP, SBP\}$$

$$t \in \mathbb{N}^+$$

$$\forall v_{HP}(t), \exists! s_{HP}(t) \in \mathbb{N}_{[0,3]} \quad (3-8)$$

$$v_{HP2} = \{true|false\} \quad (3-9)$$

$$HP2 = \{CL, HCRF\}$$

$$\forall t v_{CL}(t) == true \implies s_{CL}(t) == 0 \quad (3-10)$$

$$v_{HCRF} == false \quad (3-11)$$

The lower the early-warning score, the lower the risk of health worsening, and the closer the v_{HP} is to the values found in healthy people. The higher the early-warning score, the greater the probability of a severe health event occurring in the monitored patient.

In ICUs, several physiological parameters are monitored. The combination of values found at a given moment can better evidence the risk associated with a patient than individual early-warning scores. Only one s_{HP} may not be accurate enough to indicate the clinical picture of a monitored patient. Thus, the individual early-warning scores s_{HP} are added, and the result is also classified by ranges of values. This ranked classification is named the combined early-warning score (**CEWS**). Then, the CEWS is used to infer the patient's clinical condition (CC) at a given moment.

$SUMs(t)$ represents the sum of individual early-warning scores for each $v_{HP}(t)$. For example:

$SUMs(t) < 1,$	$CEWS(t) := 0, CC == \text{"normal"};$
$SUMs(t) \geq 1 \text{ and } SUMs(t) < 5,$	$CEWS(t) := 1, CC == \text{"worrisome"};$
$SUMs(t) \geq 5 \text{ and } SUMs(t) < 7,$	$CEWS(t) := 2, CC == \text{"severe"};$
$SUMs(t) \geq 7,$	$CEWS(t) := 3, CC == \text{"critical"};$

The $CEWS(t)$ not only estimates the patient's risk of worsening and even dying but also guides health teams on the level of care that the patient

should receive. The higher the CEWS, the more attention should be given to the patient, and more immediately or frequently should be visited.

3.2.3

Self-Adaptive solution design

As explained in subsection 2.4 the following questions should be answered when designing a self-adaptive solution.

A Under what conditions does the system undergo adaptation?

When the patient's health condition changes. The CEWS represents the health condition. Changes in CEWS are assessed in variable intervals depending on the last calculated CEWS. Additionally, when data is very similar, the similarity is assessed using inferred individual scores in the same variable intervals as well.

B Should the system be open-adaptive or closed-adaptive?

The system is closed-adaptive. It does not allow new execution plans during run-time. Our execution plan is explained in Section 3.1, and the implementation in subsection 3.5.4.

C What type of autonomy should be supported?

Once configured, the solution is fully automatic and auto-contained, running autonomously.

D Is the adaptation cost-effective? In which conditions?

Yes, in any condition that allows reducing sampling rates characterized by lower CEWS and containing highly similar and redundant data.

E How often is adaptation considered?

The adaptation takes place in variable intervals named periods. Period lengths are defined in seconds. The higher the CEWS, the shorter the pLen, and vice-versa.

F What kind of information must be collected to make adaptation decisions?

Adaptation is based on EWSS that utilizes health parameters such as vital signs. Then, the solution must acquire or sense vital signs regularly and calculate or infer the early-warning scores.

G How accurate and current must the information be?

It depends on the clinical case being monitored. Some diseases and profiles may require low latency and data to be very accurate. However, clinical cases are out of the scope of our solution. It is allowed the configuration of parameters such as minimum and maximum frequencies and re-assessment time for running the self-adaptive procedure for each combined score. Configuration of parameters gives flexibility to adequate the solution to different cases.

3.2.4

Quality metric for early-warning score monitoring

Reducing the monitoring frequency can skip important events only captured using higher frequencies. To measure the potential loss of the adaptive frequency approach used in our proposed solution, the absolute difference in recorded time of each CEWS was observed between a system running at the maximum allowed frequency and using our adaptive approach. The lower the absolute difference, the better the quality.

The EWSS defines a set of CEWS $0,1,..,n$ computed based on individual scores. There is one individual score for each health parameter. The health parameters are read in intervals. The data rate is fixed in a system running without our adaptive sampling proposal. Then, for each reading, the individual and CEWS are calculated. The elapsed time between readings is accumulated and grouped by each CEWS according to the current calculated score.

On the other hand, in our proposal running the adaptive sampling algorithm, the data rate is variable, but the method for recording the time is the same. It accumulates the time difference between readouts according to the last calculated CEWS.

Then, the total time recorded in each CEWS is compared between the baseline and our solution, and the absolute difference is divided by the corresponding time in the baseline. Absolute differences are calculated for all evaluated INDs, and the error rate is found by dividing the absolute differences by the total recorded time in each CEWS. However, when analyzing the error rate of monitoring quality in individuals, the rates are weighted by the distribution of recorded time in each CEWS. Then, error rates are proportional to the total recorded time in each CEWS, and the total error is the sum of the weighted error rates.

- **Monitoring Quality (MQ):** The total monitored time in each CEWS using our solution should be as close as possible to the registered in the baseline system running without our solution for the same *IND*. The

absolute and percentage differences in monitored time in each CEWS are quality parameters assessed in our proposal.

3.3

Combined early-warning scores for reducing non-actionable alarms - RQ3

Patient monitoring applications generally send alarms when monitored values are outside their normal thresholds. Every health parameter (HP) sample is represented by a numerical value v read in time t . So, each sample can be represented as $v_{HP}(t)$, as presented in expression 3-7. Time t is updated regularly at fixed or variable intervals. Each HP is configured with a superior (sup_{HP}) and an inferior threshold (inf_{HP}) to distinguish normal and abnormal values. When $v_{HP}(t)$ is below or above these thresholds, an alarm is triggered at time t ($Alarm(t)$). Then, the threshold-based alarm conditions can be expressed as:

$$(v_{HP}(t) < inf_{HP}) \vee (v_{HP}(t) > sup_{HP}) \Rightarrow Alarm(t) \quad (3-12)$$

However, using an EWSS, it is possible to represent a scale of data abnormality and handle alarms accordingly. Instead of sending alarms when v_{HP} is abnormal, alarms can be sent only when the current early-warning score is greater than the previous score. This procedure will reduce the number of redundant alarms.

$$(s_{HP}(t) > s_{HP}(t-1)) \Rightarrow Alarm(t) \quad (3-13)$$

For example, using the threshold's formula (3-12), with sup_{HR} equals to 99bpm, a sequence of HR reads: 90bpm, 110bpm, 111bpm, 110bpm, 111bpm, will generate four alarms. However, using the formula for scores (3-13), the correspondent scores are 0, 1, 1, 1, 1, following the NEWS-2 reference (Figure 2.3), only one alarm Will be generated, in the transition from zero to one.

Furthermore, some variations can also be filtered using combined scores instead of individual scores. For example, in a health condition where $s_{HR}(t1)=1$, and $s_{SBP}(t1)=0$, the combined score will be $CEWS(t1)=1$. In the next reading cycle, $s_{HR}(t2)=0$ and $s_{SBP}(t2)=1$, $CEWS(t2)=1$, and no alarm will be triggered because $CEWS(t1) == CEWS(t2)$. Differently, using the individual scores, one alarm will be triggered because $s_{SBP}(t2) > s_{SBP}(t1)$.

Moreover, using adaptive sampling, some noise that generates alarms can be avoided. For example, let's suppose that data is read every second in the baseline system. Our adaptive sampling algorithm reads data every two seconds according to the patient's health condition. For simplification, it will be observed only the HR parameter with the following sequence:

Time	Baseline		Adaptive Sampling	
	HR Value	Score	HR Value	Score
1s	95bpm	0		
2s	95bpm	0	95 bpm	0
3s	95bpm	0		
4s	105bpm	1 (AL)	105 bpm	1 (AL)
5s	105bpm	1 (AL)		
6s	105bpm	1 (AL)	105 bpm	1
....				

Two alarms will be triggered at seconds 4 and 6 in the baseline system. However, in the adaptive sampling version, only one alarm would be triggered at 4s.

3.3.1 Strategies to avoid negligible and redundant alarms

In addition to using adaptive sampling, combined scores, and triggering alarms only when the combined score increases, two strategies were applied to reduce the number of alarms. One aimed at avoiding negligible alarms using delays. Another one is to avoid redundant alarms using redundant lockout time windows.

When an alarm event occurs, the algorithm delays a few seconds (dt) to trigger the alarm and only sends it if the alarm condition remains after dt . For example, Figure 3.5 depicts a hypothetical monitoring context where in period A, the combined score is 0, changing to 1 at the beginning of period B. The alarm is not triggered at that moment. The system waits for dt elapsed, and the combined score remains 1. So, the alarm is triggered. After that, the combined score increases to 2, but this time, after dt , the combined score returns to 1. Then, the alarm is not triggered. Using delays, very short variations in scores are avoided, not triggering negligible alarms.

Furthermore, a timer is started when the alarm is triggered in period B of our example. Eventually, the patient's health condition returns to the combined score of zero at the end of period C. It increases to 1 again at the beginning of period D. The alarm is not triggered because the redundant lockout window is still enabled, avoiding sending alarms of the same type within that time window. If the combined score increases to two instead of one and stays at two after dt , alarm type 2 will be sent, and a parallel new redundant lockout window of alarm type 2 will start. Redundant lockout window time is a parameter to be configured in our algorithm. When this time

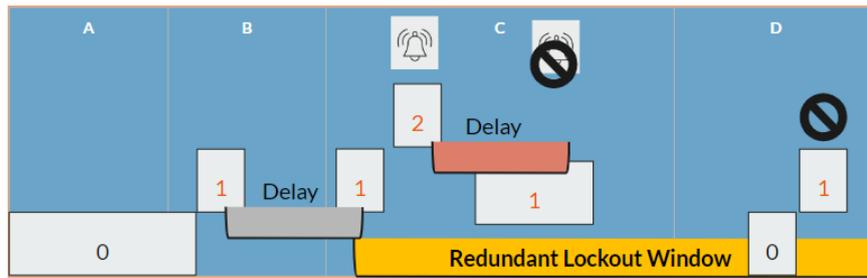


Figure 3.5: Strategies to avoid negligible and redundant alarms. Trigger delays postpone alarms. The redundant lockout window avoids repeating the same alarm.

expires, alarms will be sent normally, and a new timer will be configured.

Figure 3.6 shows the alarm procedure. After calculating the combined score and checking that it has increased, it is verified if the new potential alarm condition is within the redundant lockout window time. If not, it follows to check if there is already a previous alarm of the same type being delayed. If not, the delay timer is started according to the new score. When this alarm delay timer expires, it is checked if the current combined score is lower than the original alarm type. If not, the alarm is triggered. Otherwise, it is checked if the current score is greater than the combined score when the alarm was created. If it is, the redundant window for the current score is also checked for this new alarm type, and the alarm is triggered. A correspondent redundant lockout window is enabled whenever an alarm is triggered and a timer is started. When the timer for the lockout window expires, the correspondent redundant lockout window is disabled.

3.3.2

Quality metrics for alarm reductions

The previous subsections presented several characteristics and strategies to reduce the number of excessive alarms in IoT-PMAs. However, it is necessary to check the effect of such strategies regarding the accuracy and missed detection rate of alarms in relation to a benchmark. In our case, the original patient monitoring system. This system is also named the baseline or reference system, considering it will provide the "ground truth" to assess our proposed approach.

Alarms' accuracy is related to alarm types. Our proposed solution should trigger alarms reflecting the original condition which provokes the alarm. For example, the combined score is 0 and changes to 1. Then, the alarm is of type 1. If the condition changes to 2, the alarm is of type 2, and so on. Therefore, the new score greater than the previous reflects the alarm type. Our proposed

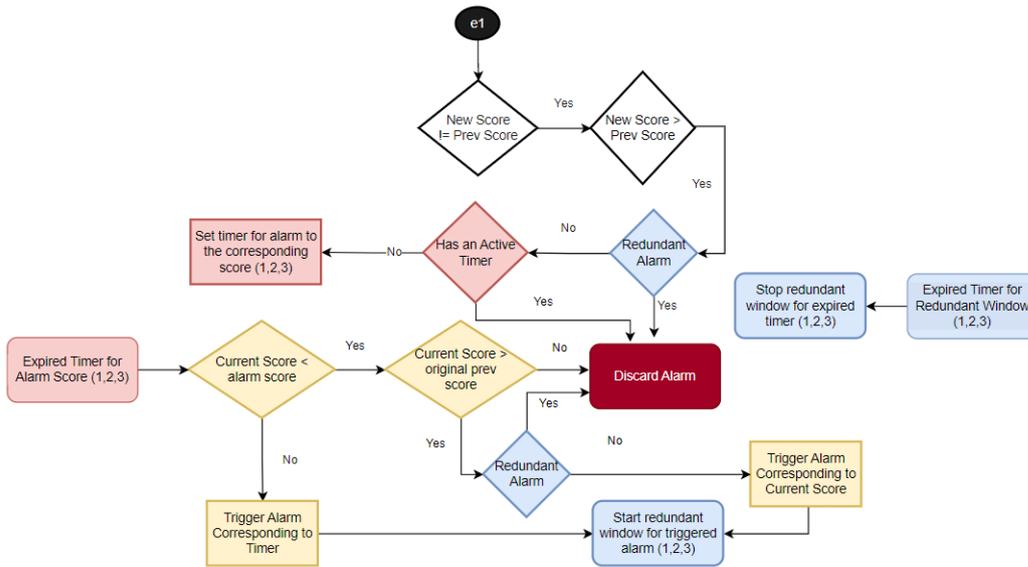


Figure 3.6: Alarms procedures. Three asynchronous processes run concurrently: new scores entrance (e1), alarm delay control in pink and yellow, and redundant window control in blue.

solution should trigger the same alarm types as the baseline patient monitoring system.

Moreover, the missed alarm detection rate regards the rate of missed alarms by our solution in comparison to the baseline system considering different detection time frames. The complement of this rate represents all non-negligible alarms triggered in the original system that was also triggered using our proposed solution within a time frame. Alarms are considered negligible when they come from very short transitions in combined scores. This short transition time can be configured in our solution. In Chapter 4, it will be described the utilized references to set up this parameter.

Firstly, it is defined what alarms are considered actionable in the baseline system. For example, transitions that last at least a time t . Secondly, a time window (tw) to identify the alarms in our solution is also defined. Then, when an actionable alarm happens in the baseline, it is checked within the time window tw after the corresponding original triggered time in our system if there is an alarm in our system. We include alarms avoided by the lockout window strategy because we wish to check the capacity of our system to detect that alarm situation in the baseline system. Even if the alarm was considered redundant, meaning that we voluntarily avoided it. So, if there is an alarm triggered by our solution, its type is checked to register the accuracy metric. Otherwise, if there is no alarm within the time window in our system, the missed alarm detection counter is increased.

To summarize the adopted alarm quality metrics:

- **Missed alarm detection rate (ALMDR)**: Using adaptive sampling, our solution may delay triggering or even skip alarms that are not realized at a lower sampling rate. The missed alarm detection rate measures how many alarms are missed by our solution within a given time frame. Then, the missed detection rate is given by different time window sizes. Therefore, the missed detection rate is composed of two metrics, a time window and the percent of missed alarms within this time window.
- **Alarm Accuracy (ALACC)**: The alarm triggered by our solution should match the type of alarm triggered by the baseline monitoring system. The accuracy is the percentage of alarms in our solution derived from the same transitions of combined scores in the baseline system within the time windows defined by the ALMDR metric.

3.4

Combined early-warning scores and energy-efficient in IoT devices - RQ4

Embedding our proposed approach on the wearable device has the intuition of bringing more intelligence to where the data are acquired. However, it may require more processing and memory utilization, leading to more energy requirements in the constrained device. On the other hand, a smart algorithm may also manage resources more efficiently, reducing the number of other operations, such as data acquisition and transmission. This section will explain how embedding an EWSS can be energy-efficient in our wearable sensor device.

Each wearable device (DEV) is powered by a battery with a capacity \mathbf{C} . Capacity is described as an amount of energy \mathbf{E} for a time \mathbf{T} . Then, for the same \mathbf{C} , higher \mathbf{E} , lower \mathbf{T} , and vice-versa.

$$C = E * T \quad (3-14)$$

DEV executes several functions: reading (r), processing (p), and transmitting (t) that require energy. So, executing each function in the DEV consumes an amount of energy q , and \mathbf{qr} , \mathbf{qp} , \mathbf{qt} represent the energy consumption to execute once each function, respectively. However, there are two types of transmissions: regular v_{HP} and alarms. Because alarms will not occur at the same rate as the regular transmissions, it is represented the energy required for transmitting alarms as \mathbf{qal} and for regular transmissions \mathbf{qt} .

Additionally, even without performing any functions, the DEV has a basic energy consumption \mathbf{qb} when the microprocessor and drivers are awakened. The DEV operating modes are known as sleep mode and awake mode, with $\mathbf{qbSleep}$ representing consumption during sleep mode and $\mathbf{qbAwake}$ during

the awake mode. The energy requirement during awake mode is much higher than during sleep mode.

However, it is proposed to utilize the sleep mode whenever the DEV is not performing any operation, as explained in Subsection 3.5.4. As a scheduler will only enter AWAKE mode when one of the reading, processing, or transmission functions is executed, it can be disregarded in our model the $qbAwake$ term since the energy consumption will be covered by qr , qp , qt and qal . Then, the energy requirement formula should be:

$$Q = (qbSleep * timeSleep) + (Xqr * timeRead) + (Yqp * timeProc) + (Zqt * timeTrans) + (Wqal * timeTransAl) \quad (3-15)$$

X = number of reads performed

Y = amount of processing performed

Z = number of transmissions performed

W = number of alarm transmissions performed.

$timeSleep$ = amount of time DEV spent in sleep mode

$timeRead$ = time to perform the readout operation

$timeProc$ = time to perform the processing

$timeTrans$ = time to transmit the periodic data

$timeTransAl$ = time to transmit an alarm

It is necessary to include the time to perform each operation in our equation to calculate the time in sleep mode ($timeSleep$) as the total running time - ($X * timeRead + Y * timeProc + Z * timeTrans + W * timeTransAl$). Additionally, it is necessary to adjust the amount of energy qr , qp , qt according to the utilized time unit scale.

Moreover, in sleep mode, as said, the energy requirement is much lower than performing any operation. Then, updating the battery capacity equation 3-14, it will be noticed that by reducing Q , battery capacity will last longer.

$$C = (E + Q) * T \quad (3-16)$$

Our algorithm should decrease the time to perform tasks or reduce the number of operations to increase the operating time given a fixed capacity. Our proposal focuses on reducing the number of operations because it will benefit any chosen sensor technology. In Section 3.2, it is explained how the adaptive sampling procedure reduces the sampling rates. Then, reducing the number of readouts, the X coefficient in formula 3-15 will be lowered. Processing is also reduced compared to fixed processing rates because it occurs in pLen

intervals, which are variable according to the combined scores. Whenever the combined score is not maximum, $pLen$ is longer, so processing will take place less frequently, lowering \mathbf{Y} coefficient.

In the same way, periodic transmissions occur in a time proportional to $pLen$. Variations in patients' health conditions that represent lower combined scores will also reduce the number of periodic transmissions, represented by the \mathbf{Z} coefficient. Finally, in Section 3.3, several strategies were presented to reduce the number of alarms, lowering \mathbf{W} coefficient. Therefore, by reducing the number of operations, Q will be reduced, and T must be higher to maintain equality.

Because the EWSS and proposed algorithms have small complexity with fixed operations independent of the data size, very little time is added to **timeProc**. The reductions in coefficients \mathbf{X} , \mathbf{Y} , \mathbf{Z} , and \mathbf{W} will overcome the additional processing load. Therefore, our approach will extend battery life in wearable sensor devices thanks to the early-warning scores, the self-adaptive and adaptive sampling procedures, and the strategies to reduce negligible and redundant alarms. In addition, the time saved by reducing the number of operations will increase **timeSleep**. However, the energy requirement during sleep mode is very low, which will benefit even more by extending the device's battery life.

3.5 Embedding the solution in IoT devices - RQ5

This section presents how our proposed solution can be embedded in an IoT device and how the proposal can be integrated into the IoT-based patient monitoring application (IoT-PMA). Firstly, we have conceptualized an end-to-end IoT-PMA and the interfaces with our device. Secondly, in the context of this general architecture, we have developed a proposal for the embedded system. In parallel, we devised and built a hardware prototype to monitor patients remotely. Finally, the solution was implemented and embedded in our devised prototype.

Figure 3.7 shows the conceptual architecture of an IoT-PMA composed of three layers. This conceptualized application was named Smart Health Sense (SH-Sens). The SH-Sens was partially implemented in a virtualized environment to check the proposed concepts and their feasibility before designing a hardware prototype and embedded software integrated into that application. The external API interfaces of the data distribution layer were not implemented, nor were the functionalities of the data processing and external application modules, because they will not interfere directly with our proposal.

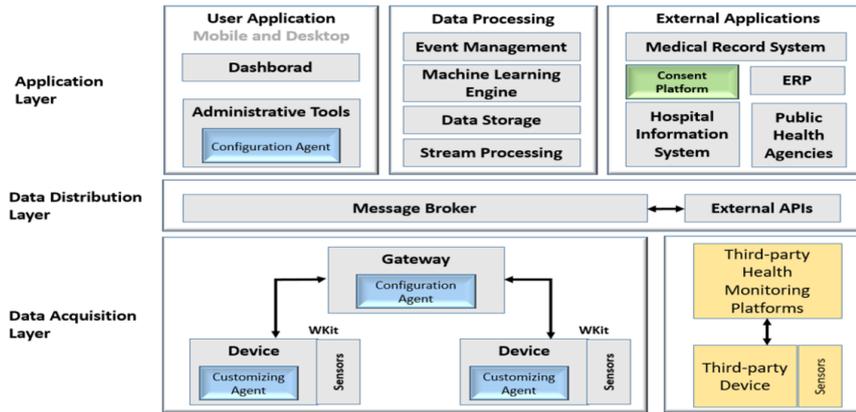


Figure 3.7: The conceptual architecture of an IoT-based patient monitoring application as a reference for our embedded solution. Based on (PAGANELLI et al., 2022).

A complete description of this architecture can be found in (PAGANELLI et al., 2022).

3.5.1

SH-Sens modules: solution virtualization

The virtualized environment was essential to verify the integration of our embedded proposal within a scalable IoT-PMA, compare our proposal to related work, and design our hardware solution. However, it is not the main focus of this thesis.

Figure 3.8 shows the implemented modules of the conceptual architecture.

The modules were deployed using Docker containers¹ grouped by functionalities. Each module has its specific container: SH-Sens Wkit (Wearable Kit of Sensors), SH-Sens Base Station, SH-Sens controller, ContextNet, Kafka/Zookeeper, MQTT Proxy, MQTT Broker, Postgresql, and Django (HTTP Server). The SH-Sens containers (WKit, Base Station, Controller) are based on a common image with the SH-Sens API. The docker-compose tool² was used to orchestrate the container's deployment.

Each WKit is a thread running inside the SH-Sens WKit container that reads data from a patient's record of public datasets such as MIMIC (MOODY; MARK, 1996), MIMIC II (SAEED et al., 2002), and Queensland (LIU; GÖRGES; JENKINS, 2012). These databases provide actual vital sign time series captured from inpatients in hospitals. Data is processed within each thread, emulating our proposed algorithms. We also implemented the

¹<https://www.docker.com/>

²<https://docs.docker.com/compose/>

basic functionalities of base stations in other threads running within the SH-Sens Base Station container. Those functionalities included receiving data from several WKit and redirecting them to the MQTT Proxy server. The transmissions of periodic data and alarms between the WKits and the base station are emulated using broadcast UDP in a virtualized network. The communication protocol between the WKit and base stations is described in (PAGANELLI et al., 2022). However, message formats are described in subsection 3.5.3.

The base station receives the messages and translates them to MQTT (Message Queuing Telemetry Transport) messages sent to an MQTT Proxy Server that feeds an Apache Kafka Server ³ and the ContextNet middleware. ContextNet has several features to deal with smart mobile objects described by Endler and Silva (ENDLER; SILVA, 2018).

The MQTT was used as the application protocol to exchange messages between the gateways, central processes, and applications. MQTT is a publish-subscribe lightweight messaging protocol well-suited for exchanging data in real-time over unreliable networks or ones with intermittent connectivity (Al-Masri et al., 2020). It has a hierarchical topic structure in which each topic can broadcast messages to its subscribers in a decoupled way. Then, once a message is published to a topic, it can be consumed by multiple processes in parallel that have subscribed to that topic.

A backend controller process running in the SH-Sens controller container, subscribed to Kafka topics, receives the messages sent by the base stations (through the MQTT Proxy), persists information in a Postgresql database, and redirects periodic data and alarms to front-end applications using a public MQTT broker. The front-end applications were developed for the Web and smartphones. The Django Web server ⁴ was utilized to develop and access an administrative interface to our application. It also acts, if necessary, as an endpoint for REST APIs. WKits, base stations, and controller modules were developed using Python 3.9, while front-end applications were developed in JavaScript and Flutter programming languages. A short description of the developed user applications can be found in Appendix C.

3.5.2

Hardware prototype: integration to the virtualized environment

Figure 3.9 depicts the hardware prototype implementation and its integration into a simplified version of the virtualized environment. A description

³<https://kafka.apache.org/>

⁴<https://www.djangoproject.com/>

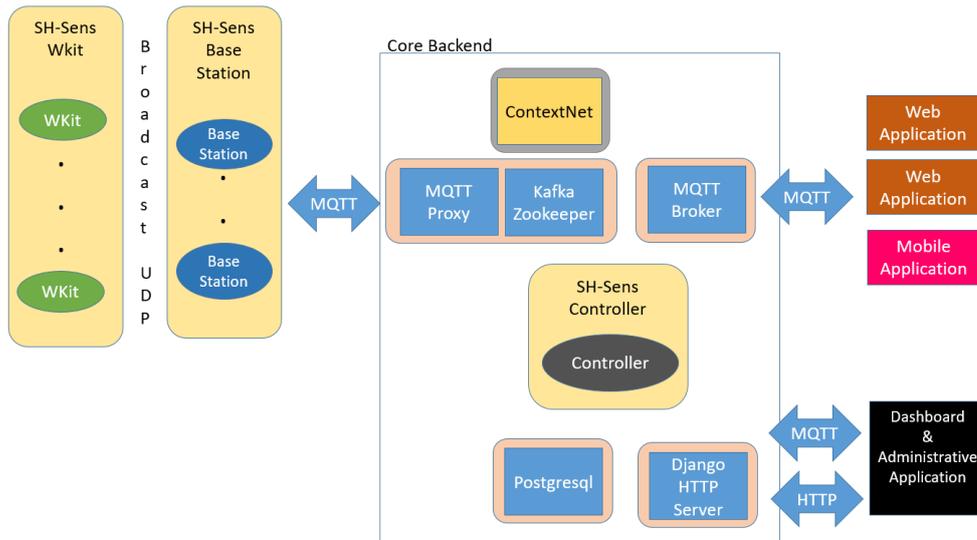


Figure 3.8: Virtualized implementation of our patient monitoring solution modules.

of the architecture is divided into four layers for clarity: (A) data acquisition, (B) wireless communication, (C) central processes, and (D) user applications.

Data acquisition comprises the hardware and software modules related to sensing, processing, and transmitting physiological parameters and alarms to internet gateways which will be detailed in subsection 3.5.3.

Wireless communication between wearable devices and the Internet is critical to the system. Using standard communication protocols, such as Bluetooth Low Energy (BLE), allows the use of existing infrastructures. For example, to communicate the WKit to a mobile phone in a domestic scenario. Nonetheless, it would not be ideal in infirmaries with several beds. A more efficient low-energy radio protocol could be used. Therefore, our wearable device has two options for wireless communication. The first uses BLE, and the other uses a proprietary Ultra-Low-Power Radio (ULPR). So, the wearable devices communicate to a proprietary radio base station or a BLE gateway like a mobile phone running our gateway agent. The base station was designed to accept connections from ULPR devices in the range of 10m. Thus, they can be positioned up to 20m apart. The base station agent was developed to attend a unique connection that was enough for our tests. It only supports single connections in the current version since it is not relevant to our solution proposal.

Moreover, data sent by gateways do not identify patients, only the source device. Among the central processes, a Web-based application (W, in Figure 3.9) manages the administrative information kept in a database schema. Figure 3.10 presents a simple data model as an example. Briefly, the model shows that

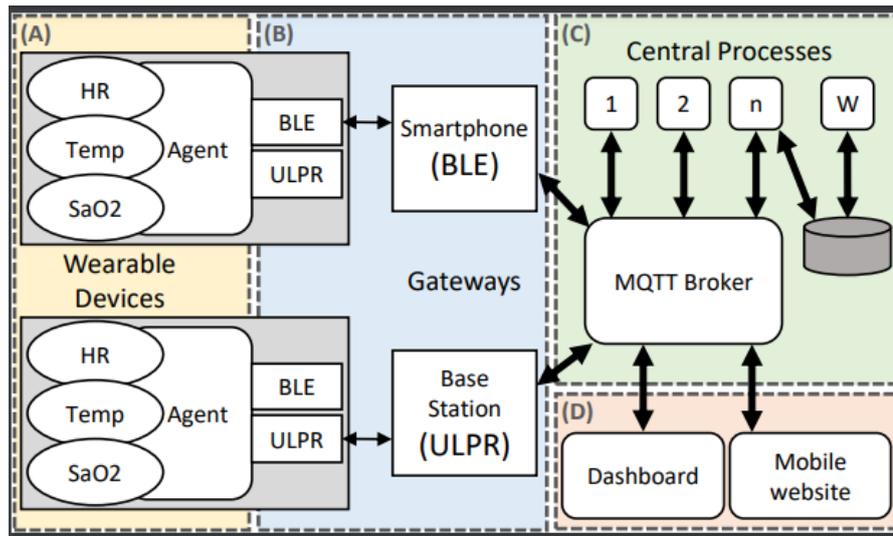


Figure 3.9: WKit hardware modules in the context of a simplified view of our conceptualized architecture.

one device is attached to a patient who occupies a bed. Beds, in turn, belong to a sector with one or more people responsible. The devices attached to patients provide periodic values of vital signs, clinical status, and trigger alarms. Other processes (1 .. n, in the figure) can run in parallel to process alarms and analyze periodic data from sensors, for instance.

In the user applications layer, after the user logs in, applications subscribe to topics in the sector(s) covered by the logged user. A specific MQTT topic structure was created for other functions, such as reading administrative information and requesting historical data. All the requests sent from the application are posted on these topics. Well-defined application programming interfaces hide implementation details. For example, it was possible to replace the Postgresql repository with a NO-SQL database, and there was no need to change the application code, only the database API.

3.5.3

Hardware prototype and the embedded agent

Our proposed approach requires computing the self-adaptations embedded in the WKit, such as implementing the scoring system, executing the behavior function to find the Bézier curve, handling alarms, and periodic transmissions. Implementing the solution in the WKit is the closest location to where data is acquired and potentially achieving the highest benefits to reduce data flows and enhance the system's efficiency regarding the use of computational resources, which is the motivation of our research.

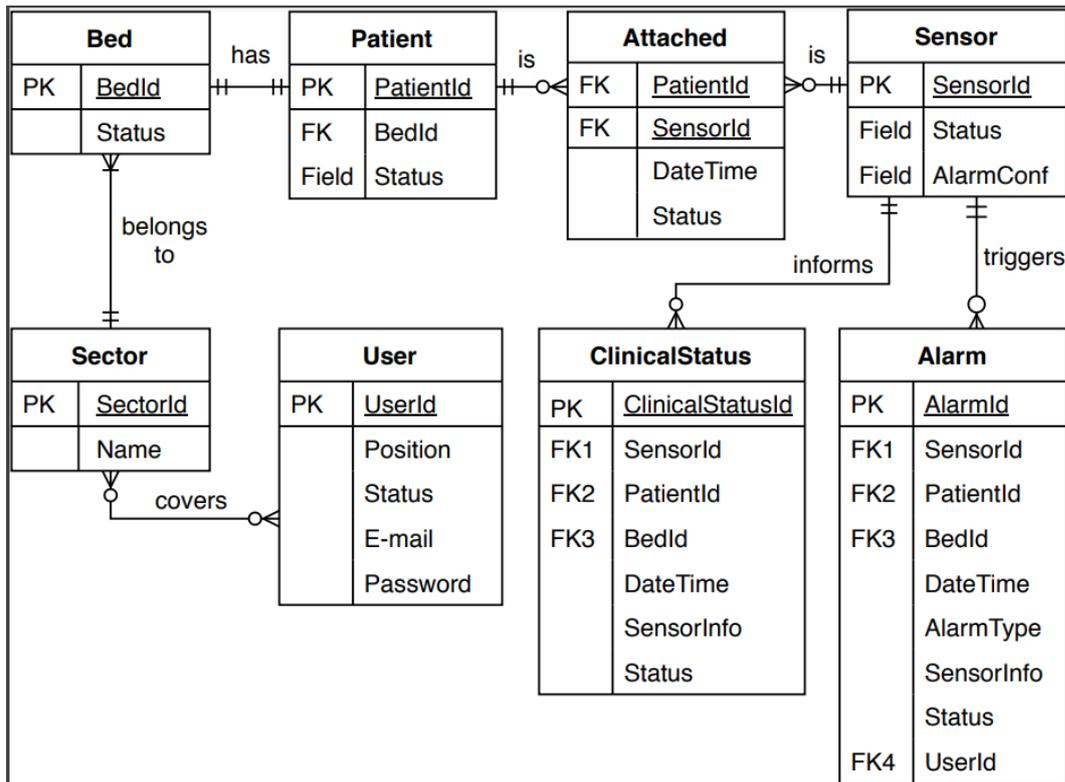


Figure 3.10: Data model summary. SensorInfo represents Heart Rate (HR), Body Temperature, and Blood-oxygen saturation (SO).

Our hardware prototype was built on the NRF52833 microcontroller ⁵, composed of an ARM Cortex M4 32-bit with FPU, 64 MHz, 512 kB flash, 128 kB RAM, and a 2.4 GHz multi-protocol transceiver. In addition, our project utilizes a photoplethysmography sensor (MAX30100) that is trustworthy for continuously measuring HR, and SO (BARTELS, 2015), and a skin temperature sensor (MLX90614) that has been widely used in research (MACRAE et al., 2018). Using these sensors, the WKit captures three vital signs: pulse rate (HR), arterial blood saturation (SO), and body temperature (BTEMP). Further, the energy module was composed of two 1.5v AA batteries. Figure 3.11 shows our WKit and the base station hardware prototypes.

Moreover, to explore the use of different radio protocols considering that the communication task is reported as the most energy-intensive operation in WBANs (ZHOU et al., 2014) (FILIPE et al., 2015) and taking into account two different use cases, one within clinical sets and the other one at home, it was decided to implement two different wireless protocols: Bluetooth Low Energy (BLE) and a 2.4 GHz ULPR embedded on the NRF52833 module. BLE is a well-known standard adopted in healthcare applications that operates at 2.4 GHz, supporting star and bus topologies (Dementyev et al., 2013) (FILIPE et

⁵<https://www.nordicsemi.com/products/nrf52833>



Figure 3.11: Images of our SH-Sens WKit prototype and base station. The wearable device is a finger clip.

al., 2015). Nordic developed the ULPR configured to operate in a star topology and its default configuration to support up to eight nodes.

The WKit sends two types of messages, periodic data with vital signs information (message type 101) and alarms (message type 102). The binary structures of these messages are described below. Note that message 101 has two versions. Version one with fixed payload and version two with variable payload to support all vital signs utilized in the NEWS-2.

Structure of Message 101

```

# Pos   Type      ID      Description
# 1     uint8_t   Type    Message type: 101 (Periodic Data)
# 2     uint8_t   Version Message structure version.
# 3     uint8_t   BsID    Base station unique address
# 4     uint32_t  Time    UNIX Epoch time
# 5     uint16_t  DevID   WKit unique identifier
# 6     uint8_t   Seq     Message sequence - circular
# If Version == 2
# 7     uint8_t   Sensors Bitmask representing payload content
# SO = 0; HR = 1; BTemp = 2; RR = 3; SBP = 4; DBP = 5; HR2 = 6;
# Mov = 7. Followed by parameter values according to Bitmask
# If Message Structure version == 1
# 7     uint8_t   SpO2    %SpO2
# 8     uint8_t   HR      Heart rate
# 9     uint8_t   Temp    int((Body Temperature -30) * 10)
# 10    uint8_t   RR      Respiratory rate
#####

```

Structure of Message 102			
# Pos	Type	ID	Description
# 1	uint8_t	Type	Message type: 103 (Alarm)
# 2	uint8_t	Version	Message structure version.
# 3	uint8_t	BsID	Base station unique address
# 4	uint32_t	Time	UNIX Epoch time
# 5	uint16_t	DevID	WKit unique identifier
# 6	uint8_t	Seq	Message sequence - circular
# 7	uint8_t	CCbef	Previous clinical condition
# 8	uint8_t	CCnow	Current clinical condition

Figure 3.12 shows the main modules of our WKit software project. The programs were developed in C++ and are linked to the Nordic nRF5 SDK library version 17.1.0. The Wearable module has a common library and customized drivers for MLX90614 (Temperature) and MAX30100 (Oximeter) sensors. It was also developed in two separate versions, one to support the BLE connection and the other to the proprietary ULPR connection. In addition, the base station module act as a gateway between the WKit and the Internet. Then, the Nordic interface communicates with the WKits using the proprietary radio. The ESP32 interface within the base station receives raw data and transforms them into MQTT packets sent to an MQTT Proxy over a TCP connection. This project organization facilitates the extension to support other sensors and hardware interfaces.

Table 3.2 demonstrates the compactness of the embedded software, confirming its feasibility to run efficiently on the selected hardware platform with a small footprint. Both versions utilized less than 15% of the available RAM. Moreover, the BLE version occupies less than 30% of available persistent memory, while the UPLR utilizes less than 5%. BLE is a multi-purpose protocol. Then, many features, such as support for mesh networks, are loaded in the standard libraries, which justifies the size difference of the generated binaries compared to the ULPR version.

Table 3.2: WKit and base station binary application size and memory requirements.

Module	Flash size (KB)	RAM (KB)
Wearable sensor / BLE	144.8	17.3
Wearable sensor / ULPR	24.2	11.2
Base Station / ULPR	10.3	10.1

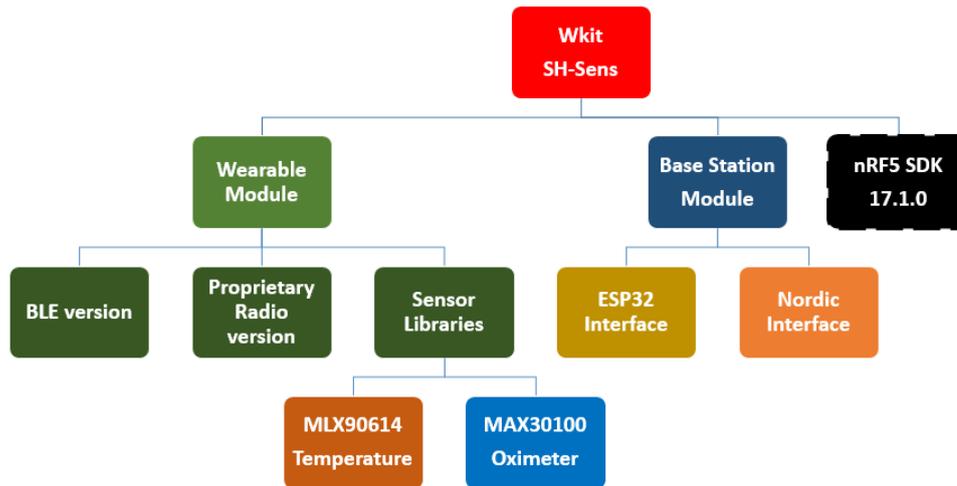


Figure 3.12: Embedded software project organization. The nRF5 SDK and the implemented modules in the Data Acquisition Layer.

3.5.4 Scheduler and power down mode

The scheduler controls the execution of tasks asynchronously on the WKit. The implementation allows the microcontroller (nRF52833) to spend most of the time in sleep mode, also called power-down mode. Such mode refers to a condition where the microcontroller has very low power consumption. The scheduler is non-preemptive, event-driven, and was based on the TinyOS project ⁶, which is an operating system created for microcontrollers, being indicated for WSNs and IoT. Once a task starts executing, it is not interrupted and does not call other tasks directly. Instead, the task puts other tasks, similar to a message passing, in the ready queue when necessary. Actually, the ready queue is implemented as an array that is traversed sequentially every time the scheduler starts executing. The last array position's state is persisted in memory, avoiding starving situations. This state indicates the next position from where the scheduler must start. Moreover, each task type has a specific array position that holds a logic bit indicating if that task should be processed or not. Then, when the bit is enabled, the correspondent procedure is called by the scheduler to be executed.

The scheduler can perform the tasks listed below:

- TASK-WAKEUP-TEMP-SENSOR
- TASK-READ-TEMP-SENSOR
- TASK-WAKEUP-OXIM-SENSOR
- TASK-READ-OXIM-SENSOR

⁶<http://www.tinyos.net/>

- TASK-SCORESYSTEM-PROCESSING
- TASK-SEND-PACKET-RADIO
- TASK-END-PERIOD
- TASK-WAKEUP-DELAYED-ALARM1
- TASK-WAKEUP-DELAYED-ALARM2
- TASK-WAKEUP-DELAYED-ALARM3
- TASK-WAKEUP-REDUNDANT-ALARM1
- TASK-WAKEUP-REDUNDANT-ALARM2
- TASK-WAKEUP-REDUNDANT-ALARM3

Sensors are responsible for measuring temperature, blood saturation, and heart rate. They are placed in sleep mode to reduce overall energy consumption when not in use. Specific tasks placed in the ready queue activate these sensors, taking them out of sleep mode so that they can perform their reading procedure.

Most of these tasks are inserted into the ready state through their respective timers or else by functions started by the scheduler. For example, the `TASK-WAKEUP-TEMP-SENSOR` activates the temperature sensor (MLX90613), taking it out of sleep mode and starting the sensor readings. When the sensor finishes the reading process, the implemented driver will place in the scheduler another task (`TASK-READ-TEMP-SENSOR`), indicating that it already has a valid temperature reading and that such reading can be processed. In the same way, `TASK-WAKEUP-OXIM-SENSOR` wakes the oximeter (MAX30100) from sleep mode. When the reading process finishes, our customized driver places another task (`TASK-READ-OXIM-SENSOR`) into the scheduler array, indicating that it already has valid SO and HR readings. All the readouts put the acquired value in a global buffer updated after each readout.

The `TASK-SCORESYSTEM-PROCESSING` is triggered to calculate the current score of the last read values periodically according to current sampling rates. This task also increments period counters of each individual score and finds the composite score (*rCombScore*) to analyze if there is an alarm. In the case of alarms, according to the alarm, the correspondent timer is started if the correspondent redundant lockout window flag is down. When this timer expires, the task `TASK-WAKEUP-DELAYED-ALARM[1-3]` takes place, so the current combined score (*rCombScore*) is analyzed and stored in a global buffer. If *rCombScore* is equal to or greater than the alarm type (1,2,3) given by the *rCombScore* which initiated the alarm, the redundant lockout window alarm flag corresponding to the type of the alarm is activated, and the alarm is sent.

In addition, a timer is started to withdraw the flag after the redundant lockout window time elapses thru the task `TASK-WAKEUP-REDUNDANT-ALARM[1-3]`. Then, our two strategies to avoid insignificant and redundant alarms can be controlled by our code.

In parallel, the timer for task `TASK-END-PERIOD` is running to wake up in `pLEN` (period length) intervals. When it expires, this task goes to the ready state and is executed to process the self-adaptive procedure, update the parameters, and start a new timer for the next period.

Finally, the `TASK-SEND-PACKET-RADIO` transmits the processed values of the sensors to the base station. In the case of alarms, it is processed when it is triggered. In the case of periodic messages, when the shortest interval to send the message elapses, the radio is wakened to send that message.

This prototype was important to demonstrate a feasible implementation of our proposal and analyze the energy requirements of the solutions.

More details of our embedded solution will be given in Chapter 4 when it is explained the application of our proposal. In Appendix B, details of the development phases of the WKit hardware prototype are presented.

Finally, in this Chapter, our solution proposal to address the problems described in Section 1.3 was presented. In the next Chapter, it will be demonstrated the application of our solution and the effects on data samplings, alarms, and energy consumption.

4

Proposed Approach Application

This Chapter presents the experiments performed to support our proposed approach for reducing excessive sensor data generation (DG), non-actionable alarms (AL), and energy consumption (EC) in IoT-PMA wearable sensors using combined early-warning scores.

The experiments aimed at demonstrating the potential reductions and their cost in terms of monitoring integrity (MINT), alarm accuracy (ALACC), and alarm missed detection rate (ALMDR). A testing environment was built and configured to execute the experiments, and a hardware device was designed and implemented. Data used in our experiments were selected from well-known public health datasets and energy parameters collected from our hardware device.

The following sections will present the testing environment, utilized data, experiments methodology, achieved results, and a deep discussion about our experiments and interpretation of the results.

4.1

Testing Environment and Datasets

Our testing environment is composed of a software simulator that implements the features described in sections 3.1, 3.2, 3.3 and a hardware prototype, described in sections 3.4 and 3.5, loaded with an embedded system version with the same self-adaptive and alarm reduction features of our simulator.

To perform the comparisons, a baseline was established considering actual data collected from patients by standard multiparametric monitors in ICUs at original sampling rates. The Multiparameter Intelligent Monitoring in Intensive Care I (MIMIC) and II (MIMIC-II) datasets from the Physionet project, which collected several health markers from inpatients between 2001 and 2008, in a tertiary teaching hospital, were utilized (MOODY; MARK, 1996) (SAEED et al., 2002). The original sampling rate is referenced as the "golden standard" sampling frequency because it is utilized in ICUs in real clinical cases.

Moreover, the NEWS-2 (PHYSICIANS LONDON, 2017) was selected as the early-warning scoring system (EWSS) to be configured in our experiments. NEWS-2, as explained in section 2.3, is one of the most utilized and accepted EWSS in infirmaries worldwide (GERRY et al., 2017). It is important to emphasize that our study does not want to validate or indicate NEWS-2 as

an ideal scoring system for clinical cases. NEWS-2 is used as a reference for a combined scoring system to support the results of our experiments and compare them to related work.

Therefore, to select the patients in MIMIC and MIMIC-II datasets, the records that contained all five vital signs (HR, RR, SO, BTEMP, SBP) necessary to calculate the NEWS-2 combined score for at least 24h were chosen. For simplification, the NEWS-2 parameters related to consciousness level and the chronic hypercapnic respiratory failure condition were fixed and considered at a normal level and not present, respectively. Thirty-six records were selected for our experiments, and some patients with unique characteristics for more specific experiments, which will be explained in the respective experiments' sections. A detailed description of the patient's records is presented in subsection 4.1.1.

The simulator was developed in Python 3.9 on the Linux operating system within a Docker container to facilitate reproducibility and control. It is based on the SH-Sens WKit container implemented in our proposed architecture as explained in subsection 3.5.1. The simulator is a simplified version of the WKit container that does not send messages but registers and records all operations of interest.

Patients' records were stored in a local drive repository and mounted dynamically in the container to keep a lightweight image of the simulator. The simulator starts one thread for each record. Record identification is passed as an argument to the thread. Data is loaded in memory using the waveform database (WFDB) Python module version 4.1.0 ¹, specifically designed for handling the Physionet dataset format. Each thread processes all data points for the selected patient, one by one, in a looping structure, within the selected period (24h) and calculates the NEWS-2 individual and combined scores, recording all score transitions and alarms as defined in Section 3.3.

Furthermore, the simulator code implements the proposed self-adaptive algorithm and alarm reduction strategies using pre-defined parameters in the same looping structure. It controls the data points our proposed procedures should process by filtering them according to current sampling rates determined by our adaptive sampling method and skipping data points that should not be processed.

Since the original datasets recorded data at approximately 1Hz, each full readout is considered 1 second. The procedure reads synchronously 1 data point for each vital sign and one full readout correspondent to read all five vital signs. Then, for each iteration in the main looping structure, the algorithm filtered

¹<https://wfdb.readthedocs.io/en/latest/>

Table 4.1: Evaluation metrics.

Metric	Description
Reads	Number of reads by each sensor
Messages sent	Number of sent messages
Total payload	The sum of all payloads
Total alarms	Number of alarms
Alarms by a score [1,2,3]	Number of alarms for each score
Avoided alarms delay	Alarms avoided by delay strategy
Avoided alarms redundant window	Alarms avoided by window strategy

the data point in the original datasets that should be skipped or handled by our proposed algorithm. Then, 1Hz is the maximum frequency allowed in our experiments.

In addition, using the same main looping structure, the simulator implemented the algorithms proposed by Elghers et al. (ELGHERS; MAKHOUL; LAIYMANI, 2014), Habib et al. (HABIB et al., 2016), and Harb et al. (HARB et al., 2021). So, the simulator controls each algorithm's reading time, processes data, and records interesting metrics accordingly. Table 4.1 shows the recorded metrics. And the pseudo-code below presents the general logic flow of our simulator, "iyda_" corresponds to our proposed solution, while "elghers_, habib_, harb_" are the related work algorithms.

```

# SIMULATOR PSEUDO-CODE
Time = 0
record := Load MonitoringDataTimeSeries(patient_id)
While Time < EXP_TIME {
  For each HP in (SO, HR, RR, BTEMP, SBP) {
    value := Read(record, Time, HP)
    indEWS := CalculateIndividualScore(HP, value)
    Record logs1_baseline()
    # *_proc1() - logic and logs for values and individual scores
    if Time matches iyda_sampling_rate: iyda_proc1()
    if Time matches elghers_sampling_rate: elghers_proc1()
    if Time matches habib_sampling_rate: habib_proc1()
    if Time matches harb_sampling_rate: harb_proc1()
  }
  Calculate the combined early-warning score
  Record logs2_baseline()
  # *_proc2() - logic and logs for combined scores/periods/rounds.
  if Time matches iyda_sampling_rate: iyda_proc2()
  if Time matches elghers_sampling_rate: elghers_proc2()
  if Time matches habib_sampling_rate: habib_proc2()
}

```

```

    if Time matches harb_sampling_rate: hard_proc2()
    Time++
}

```

Therefore, the simulator records operations regarding the baseline system, our proposed solution, Elghers et al., Habib et al., and Harb et al. proposals. Note that some metrics in Table 4.1 apply to all algorithms, such as the number of reads, messages sent, and total payload. In contrast, others are only for our proposal and hypothetically for the baseline, such as total alarms, alarms by a score, and avoided alarms because the baseline system does not process alarms using early-warning scores originally.

Some parameters were necessary to configure the algorithms, for example, period size (pLen) representing the time interval in seconds to process the self-adaptive procedure in our algorithm, and it also determines the set of data points to analyze variability or stability in related work algorithms.

Table 4.2: Simulator parameters.

Parameter	Definition
Iterations	Number of full readouts (all vital signs)
Round Size	Number of periods to be considered
Period Size (sec) (*)	Time interval to process adaptive sampling
Max. Freq. (Hz) (*)	Maximum sampling rate
Min. Freq. (Hz) (*)	Minimum sampling rate
Transmission factor (*)	Delay factor to transmit period messages
r0 - risk factor (*)	Defines frequency distribution
Redundant window (sec)	Time frame of the redundant window
Alarm delay (sec)	Time frame of alarm's delay

Table 4.2 shows the utilized parameters. The parameters marked with an asterisk symbol (*) receive a list of values, one value for each combined score, from the lowest to the highest. The actual values for these parameters will be explained in each experiment in the next sections.

4.1.1 Patients Data Description

This subsection presents the utilized data and interesting characteristics for our experiments. In experiments I and III, it was selected from the MIMIC dataset, six patients ², and from MIMIC II, thirty patients ³.

²055n, 254n, 259n, 455n, 457n, 474n

³a44[002, 038, 129, 139, 159, 162, 178, 197, 200, 322, 378, 537, 601, 610, 635, 674, 694, 759, 921]n, a45[159, 222, 260, 384, 467, 519, 532]n, a46[165, 391]n, 3289943n, 3533682n

In experiment II, three patients were selected to characterize different health conditions: stable (a44601n), unstable (a44178n), and very unstable (a44332n). Stability was analyzed by the number of triggered alarms in the baseline, according to the combined scores. The patients triggered 0, 1000, and more than 2000 alarms during the experiment, respectively.

Figure 4.1 shows the total recorded time in seconds of all selected patients in Experiments I and III per each combined score in the baseline. It can be noticed that combined scores zero and one represent more than 66% of the total monitored time. Yet, only 13.3% of the monitored patients' time was in combined score three.

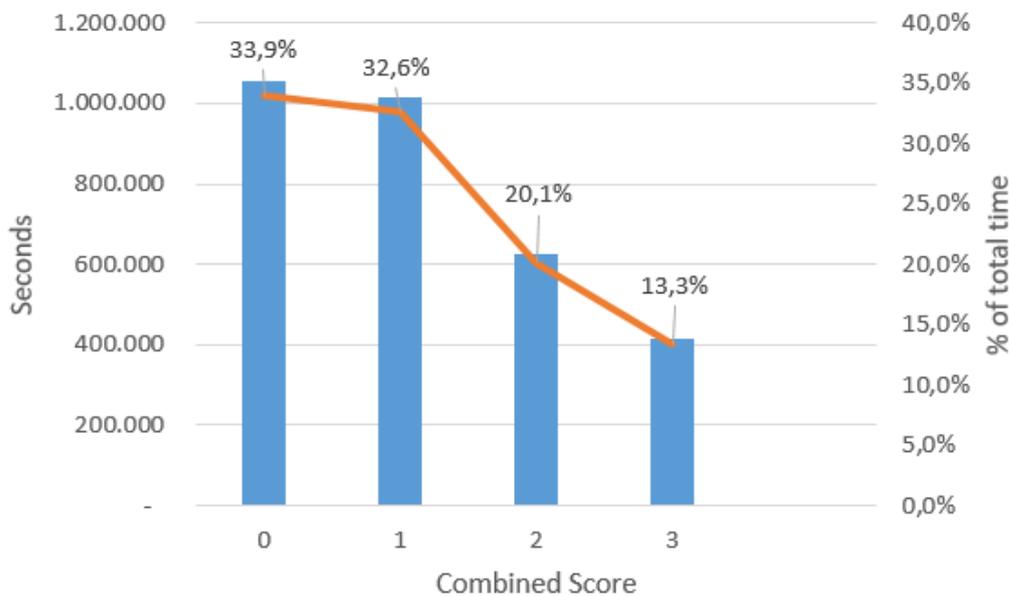


Figure 4.1: Experiments I and III - Total time per combined score recorded in the original dataset (baseline).

In experiments I and III, four patients received a combined score of zero (normal) during all the monitored times. Thirteen patients had 50% of time or more in scores two and three, as shown in Figure 4.2. It is visible the variability in the distributions of combined scores among the selected records. Thus, it allows showing different behavior patterns of our self-adaptive algorithm in experiment III.

Finally, Figure 4.3 depicts the distribution of alarms per patient from the most critical patients to healthier individuals measured by longer monitored time in higher scores. It follows the same order presented in Figure 4.2. The chart shows that even for patients that were most of the time in a critical condition (combined score 3), the number of alarms can be very low, for instance, patient a44378n.

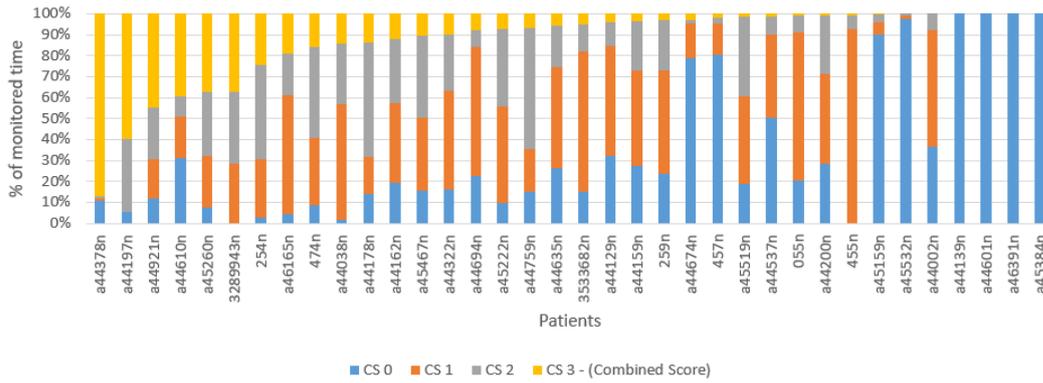


Figure 4.2: Experiment I and III - Distribution of combined scores per patients (36) in the original dataset (baseline).

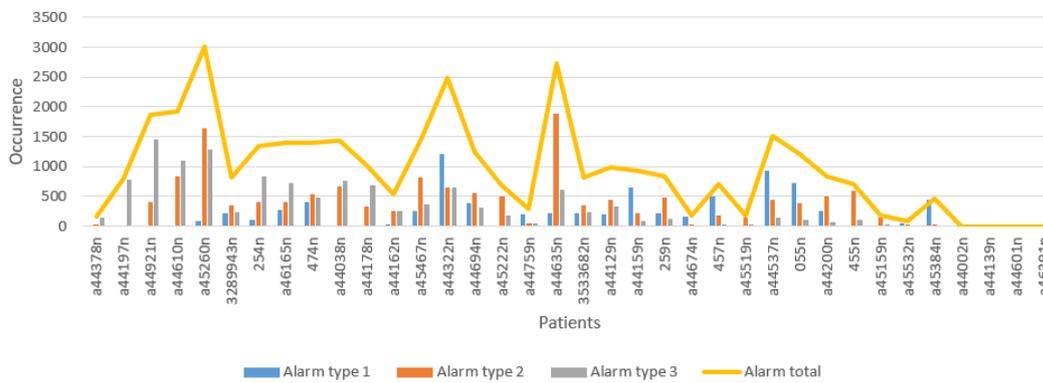


Figure 4.3: Experiments I and III - Distribution of alarms per patients (36) in the original dataset (baseline).

4.2 Experiments Organization

This section describes the performed experiments whose results can demonstrate our solution's potential benefits and drawbacks. The experiment's main objectives are to provide evidence supporting our arguments and answer our research questions.

4.2.1 Experiments Regarding the Research Questions

Experiment I, explained in Section 4.3, is based on all features described in Chapter 3 and consolidates the pieces of evidence to answer the research questions RQ1, RQ2, and RQ3. Experiment II, detailed in Section 4.4, provides evidence to answer research questions RQ4 and RQ5. Section 4.5 presents additional experiments regarding energy savings and a trial using different radio protocols, reinforcing the evidence to answer all research questions. Furthermore, the development of our hardware prototype, as illustrated in Appendix B, aimed at answering research question RQ5 and supporting

experiments II and III. Finally, Section 4.6 presents a general discussion about experiments with IoT-PMAs and summarizes the findings of our experiments.

4.2.2

Experiments Methodology Regarding the Principles of our Approach

Our experiments applied the 4 principles described in Table 3.1 that rule our proposed approach.

Experiment I combines principle P1 with principle P2, considering data redundancy and assessing the similarity of monitored data over time. Furthermore, it explores principles P3 and P4 of changing frequencies and re-assessing clinical conditions more quickly when combined scores are higher. Furthermore, it utilizes the combined scores calculated in real-time to produce alarms instead of thresholds. Moreover, the use of combined scores to trigger alarms at the sampling rate was investigated along with two strategies to avoid negligible and redundant alarms. In addition, one extra benefit of using combined scores over thresholds is the possibility of graduating alarms using well-known health guidelines.

Experiment II demonstrates the potential of energy savings in three distinct patient profiles using our proposed approach. Although our focus was on energy consumption regarding adaptive sampling and the sensor reads task, it also introduces the energy requirement of transmissions. In addition, it is possible to analyze the behavior of our algorithm and the effects on energy requirements of different combined score levels.

Experiment III entails a more comprehensive approach to the energy requirements and potential savings covering all tasks in the WKit, showing the effects of using two different radio protocols with actual energy consumption measurements and patient data to project the energy requirements during monitoring. A comparison between the self-adaptive proposal and the baseline system is also analyzed. Furthermore, this experiment also provides a deeper understanding of the effects of each proposed principle on the monitoring data with distinct profiles.

A description of the development of the SH-Sens Wearable Kit is given in Appendix B. It demonstrates the potential of embedding intelligent algorithms in wearables to promote the efficient use of resources.

4.3

Experiment I - Use of combined early-warning scores for data and alarm reductions

Based on a preliminary experiment (Appendix D) and inspired by the related works ((ELGHERS; MAKHOUL; LAIYMANI, 2014), (HARB et al., 2021)), a new algorithm was developed as described in Chapter 3 applying the principles described in Table 3.1. In addition, we defined metrics to verify the monitoring integrity and focused on sampling reductions. A new related study by Habib et al. (HABIB et al., 2016) was added to our comparisons. We also enhanced the alarm procedure, processing them at sampling instead of period rates, and implemented strategies to avoid negligible and redundant alarms. Furthermore, a more consistent method was devised to check alarm accuracy and precision. Finally, the configuration parameters of our algorithm were manipulated to achieve more considerable reductions or better alarming accuracy and precision.

The following subsection introduces the objectives of this new experiment.

4.3.1

Experiment I - Objectives

- 1) Verify the sampling and transmission reductions by comparing the results to a reference system using data from actual patients stored in public datasets;
- 2) Compare the sampling and transmission reductions to the algorithms proposed by (ELGHERS; MAKHOUL; LAIYMANI, 2014), (HABIB et al., 2016), and (HARB et al., 2021);
- 3) Verify the integrity of monitoring time regarding the NEWS-2 combined score classification after sampling reductions;
- 4) Verify the reductions in triggered alarms by comparing them to a reference system;
- 5) Verify the effects of alarm reduction strategies on alarm precision and accuracy;
- 6) Assess strategies in configuring parameters and their effects on reductions and monitoring integrity, alarm precision, and accuracy.

Algorithm 1 Modified Local Emergency Detection Algorithm *Modified LED*

Require: R_t (Instantaneous Sampling Rate).

```

while Energy > 0 do
2:   for each period do
       takes first measurement  $r_0$ 
4:     sends first measurement  $r_0$ 
       gets score  $S$  of  $r_0$ 
6:     takes measurements  $r_i$  at  $R_t$  Rate
       gets score  $S_i$  of measurement  $r_i$ 
8:     if  $S_i \neq S$  then
           sends measurement  $r_i$ 
10:     $S = S_i$ 
       end if
12:   end for
end while

```

Figure 4.4: Habib et al. (2016) Modified Local Emergency Algorithm, implemented in our simulator.

4.3.2

Experiment I - Methodology

The experiments utilized 24h hours of real data of 36 inpatients from the MIMIC (MOODY; MARK, 1996) and MIMIC II (SAEED et al., 2002) datasets as described in Section 4.1.1. These patients were chosen because their records contained the monitored information of all five vital signs (HR, SpO2, RR, BT, SBP) for 24h or more.

To verify the performance of our algorithm and compare it to the baseline and other algorithms, the metrics described in Table 4.1 were utilized. The baseline or reference system reads data at the original data rates without the self-adaptive method and alarm avoidance features. The reference system finds the individual and combined NEWS-2 scores for all readings and triggers alarms whenever the combined score is greater than the previous one.

The algorithms as described by Elghers et al. (ELGHERS; MAKHOUL; LAIYMANI, 2014), Habib et al. (HABIB et al., 2016), and Harb et al. (HARB et al., 2021) were implemented in our simulator. Subsection D.5 provides an overview of (ELGHERS; MAKHOUL; LAIYMANI, 2014), and (HARB et al., 2021) implemented algorithms. Figures 4.4 and 4.5 give an overview of the (HABIB et al., 2016) algorithm implemented in our simulator. A full description of those algorithms can be found in the original studies.

To configure those algorithms, the same round-size configuration was utilized as in our proposal. We also utilized the largest period size and our algorithm's highest and lowest sampling frequencies. For the Bézier curves utilized in (ELGHERS; MAKHOUL; LAIYMANI, 2014) and (HABIB et al., 2016) algorithms, the medium risk for all patients at 0.6 was set.

Algorithm 2 Modified Local Emergency Detection with Adaptive Sampling Algorithm
*Modified LED**

Require: m (1 round = m periods), R_{max} (maximum sampling rate).

Ensure: R_t (instantaneous sampling speed).

```

 $R_t \leftarrow R_{max}$ 
while  $Energy > 0$  do
3:   for each round do
       for each period do
           Run Modified LED (Emergency Detection)
6:   end for
       compute  $SR, SF$  and  $F$ .
       if  $N < m$  then
9:      $R_t \leftarrow R_{max}$ 
       else
           find  $F_t$ 
12:    if  $F < F_t$  then
            $R_t \leftarrow BV(F, F_t, r^0, R_{max})$  (BV behavior function).
       else
15:      $R_t \leftarrow R_{max}$ 
       end if
       end if
18: end for
end while

```

Figure 4.5: Habib et al. (2016) MLED and the self-adaptive procedure, implemented in our simulator.

It was assumed that all read values could be sent using 1 byte for the total payload calculation, even for real number values such as body temperature and the regression coefficients sent by the (HARB et al., 2021) algorithm.

All other algorithms (ELGHERS; MAKHOUL; LAIYMANI, 2014), (HABIB et al., 2016), (HARB et al., 2021) assumed that any abnormal value is critical, similar to an alarm. However, these algorithms did not aim at reducing alarms. Therefore, the number of alarms triggered by our system was compared to the reference system only. The study compares the total number of alarms and alarms classified by each combined score category (1,2,3).

To check our algorithm's integrity, the total registered time in each NEWS-2 combined score was utilized, summing up all the intervals of each NEWS-2 combined score in the reference system and our algorithm. The mean error and the absolute error percentage between both approaches were calculated.

To check the precision of alarms, the percentage of total missing alarms was verified by time differences between the alarms triggered by the reference system and those triggered by our algorithm. Therefore, this metric represents the missed detection rate of alarms. It also presents the maximum missed

detection rate of alarms for all patients by each time difference. It was considered only alarms of the baseline in which the changed interval was greater than 4s to avoid checking negligible alarms. It utilized the 5s threshold as used in (PATER et al., 2020) to reduce alarm notifications in a pediatric hospital study.

The accuracy of alarms was checked by the rate of alarms within the time interval of precision analysis that belongs to the same category as the reference system. The category of alarms is defined by the combined score that caused the alert. The higher this rate, the higher the matching and the higher the accuracy.

The efficiency of the two strategies aimed at reducing alarms was verified. One aimed at avoiding redundant alarms (redundant lockout window), and the other to avoid negligible alarms (alarm delay). It is registered the number of avoided alarms using each approach and its percentage regarding the total number of alarms in the reference system.

Finally, two configuration sets were utilized to demonstrate the effects of parameters on the reduction of samplings, transmissions, and alarms in our algorithm. Configuration (A) aims at larger reductions, while (B) aims at enhancing alarm precision and an intermediary level of reduction (Table 4.3). The configuration of the other three algorithms was adjusted to have as fair comparisons as possible according to our configurations.

Note that a transmission factor is adopted to send periodic messages in intervals longer than period sizes. For example, in setup (A), for a combined score of 0, the period size is 150 seconds, and the transmission factor is 2.0. Then, periodic transmissions will occur every 300 seconds ($150 * 2.0$). If the combined score increases, after the next assessment in 150 seconds, but before the transmission, for example, to 3, the period size will be decreased to 30 seconds, and the transmission factor will be set at 1.3. So, transmissions will occur every 39 seconds. Since it has already passed 150 seconds since the last transmission, a periodic message will be sent immediately, and the next one will be set to be sent in 39 seconds. After 30 seconds, a new re-assessment occurs, and the combined score decreases to 0 again, re-configuring the transmissions to 300 seconds. However, the next transmission will occur in 9 seconds, and the following transmission, in 291 seconds ($300 - 9$), if no increases in combined score occur again. It means that transmissions will occur at the closest scheduled time and immediately if the time elapsed from the last transmission was greater than the expected time for the next transmission after re-assessing the combined scores.

Table 4.3: Parameters of our experiments, setups (A) and (B).

Algorithms	Elghers	Habib	Harb	Ours
Round Size	4	4	4	4
Period Size (sec) (A)	150	150	150	150, 100, 60, 30
Max. Freq. (Hz) (A)	1	1	1	1/3, 1/2, 1/2, 1
Min. Freq. (Hz) (A)	1/30	1/30	1/30	1/30, 1/20, 1/10, 1/5
Period Length (sec) (B)	40	40	40	40, 30, 20, 10
Max. Freq. (Hz) (B)	1	1	1	1, 1, 1, 1
Min. Freq. (Hz) (B)	1/6	1/6	1/6	1/6, 1/5, 1/4, 1/3
Transmission factor	-	-	-	2, 1.75, 1.5, 1.3
r0 - risk factor	0.6	0.6	-	0.2, 0.6, 0.8, 0.9
Redundant window	-	-	-	(A) 90s (B) 60s
Alarm delay	-	-	-	(A) 10s (B) 5s

4.3.3

Experiment I - Results

Table 4.4 shows the performance of our algorithm with configurations (A) and (B) in comparison to a baseline and all the other three analyzed algorithms.

Our algorithm achieved the largest reduction in all analyzed variables in both configurations accordingly. Using configuration (A), it reduced the readings by 87% from baseline, while (ELGHERS; MAKHOUL; LAIYMANI, 2014) and (HABIB et al., 2016) utilized the same self-adaptive strategy, the reduction was by 65%, and for (HARB et al., 2021) by 25%. This reduction level reflects directly on the total payload, with our algorithm reaching a reduction of 47.8% compared to the second-best algorithm proposed by (HABIB et al., 2016).

All other algorithms send messages in fixed periods, while our algorithm sends messages in a variable time depending on the current inferred combined NEWS-2 score. Our algorithm sent about 66% fewer messages than the other studies supposing that all the messages from the other algorithms will be forwarded to a base station.

Configuration (B) increased the magnitude of values of all measured variables considerably. This configuration utilized shorter periods increasing the number of messages and higher frequencies, increasing the number of reads and payloads significantly. Our algorithm reduced the number of reads by 70.4% from baseline, having approximately double the readouts compared to configuration (A).

Table 4.5 shows the reductions in the number of alarms triggered by our algorithm compared to the baseline. Our algorithm reduced by 78% the number of alarms in setup (A) and 67% using (B). Negligible alarms, representing

Table 4.4: Sum of reads, payload, and number of messages using setups (A) and (B).

Algorithms	Number of Reads	Payload (bytes)	Messages Sent
Baseline	15,552,000	15,552,000	15,552,000
Elghers et al. (A)	5,448,841	3,688,326	103,680
Elghers et al. (B)	6,381,202	3,811,536	388,800
Habib et al. (A)	5,448,841	258,172	103,680
Habib et al. (B)	6,381,202	549,978	388,800
Harb et al. (A)	11,613,386	1,060,010	103,680
Harb et al. (B)	12,377,961	2,067,052	388,800
Our proposal (A)	1,964,065	134,810	35,004 (*)
Our proposal (B)	4,608,545	435,230	98,767 †
(*) 26,962 + 8,042 alarms	† 87,046 + 11,721 alarms		

Table 4.5: Comparative number of alarms to baseline.

Algorithms	Total	Score 1	Score 2	Score 3
Baseline	35,996	7,830	15,495	12,671
Our proposal (A)	8,042	1,565	3,790	2,687
Reduction (%)	78	80	76	79
Our proposal (B)	11,721	2,682	5,221	3,818
Reduction (%)	67	66	66	70
Avoided - delay	(A) 2,099	(B) 4,017		
Avoided - window	(A) 6,382	(B) 10,376		

changes in combined scores that last less than 10s captured by our algorithm, were avoided in 6.0% of total alarms, and the same alarm was triggered twice in less than the 90s in 18% of total alarms for configuration (A).

Our algorithm also increased the number of alarms by 46% from configuration (A) to (B). Note that alarms avoided by redundant lockout window and alarm delay were also increased from configuration (A) to (B). Configuration (B) avoided more alarms than (A), 11%, and 29% of total alarms for each strategy, respectively.

Table 4.6 shows that configuration (B) reduced the differences found in configuration (A), not only the total mean difference but also the total absolute time differences in comparison to the reference system. The maximum mean error of the total monitored time by combined score was -2.0% in configuration (A), which was reduced to -1.3% in (B). Yet, the maximum absolute error of the total monitored time by combined score was 15.0% in configuration (A) and 9.8% in (B).

Figure 4.6 shows the error of monitored combined scores by patients weighted by the distribution of monitored time in each combined score. For

Table 4.6: Total registered time (sec.) for each score by the baseline and our algorithm, configurations (A) and (B).

Scores	0	1	2	3
Reference	1,055,717	1,014,288	625,579	414,816
% of total time	33.9%	32.6%	20.1%	13.3%
Experimental (A)	1,102,488	1,049,179	606,008	352,725
% of total time	35.4%	33.7%	19.5%	11.3%
Difference (A)	1.5%	1.1%	-0.6%	-2.0%
Experimental (B)	1,072,599	1,043,192	620,317	374,292
% of total time	34.5%	33.5%	19.9%	12.0%
Difference (B)	0.5%	0.9%	-0.2%	-1.3%
Abs. Diff. (A)	4.5%	5.7%	7.1%	15.0%
Abs. Diff. (B)	1.6%	3.5%	3.8%	9.8%

example, a 30% of error in a combined score has different weights if the total monitored time of that combined score represents 50% of the total monitored time or only 0.5%. A non-weighted high error rate in the latter case is actually a tiny error composed of a few seconds because it occurred in 0.5% of monitored time. On the other hand, a high error in the first case represents a serious problem because it affects one-half of the monitored data. The total bar in the chart accumulates the absolute error in each combined score. It can be seen from this chart that in most patients, the weighted error rate by each CEWS was between +/- 5.0%. However, the maximum weighted CEWS error rate was 10.7% for patient a45260n. The maximum total absolute error in one patient (blue bar) was 24.5%, but the mean total absolute error was 6.8% (+/- 5.5%).

Figure 4.7 shows the error of monitored combined scores by patients weighted by the distribution of monitored time in each combined score using configuration B. It can be seen from this chart that for 35 patients out of 36, the weighted CEWS error rate was below 5.0%. The maximum weighted CEWS error rate was 5.5% for the exceptional case. The maximum total absolute error in one patient (blue bar) was 11.1%, but the mean total absolute error was 3.8% (+/- 2.8%).

Figure 4.8 shows segments' distribution by ranges of duration and combined scores. Most segments have a short duration of less than 10s. Some of these small segments may not be captured by the sampling frequencies of our algorithm, mainly in lower scores that allow very low sampling frequencies.

Table 4.7 shows the difference in the precision (missed detection rate) of alarms triggered in configurations (A) and (B) concerning the baseline. Since it is ignored transitions that last less than 5s, the total number of alarms was reduced from 35,996 (Table 4.5) to 19,433 alarms, a 46% reduction. Missed alarms represent the number of alarms not captured by our algorithm by the

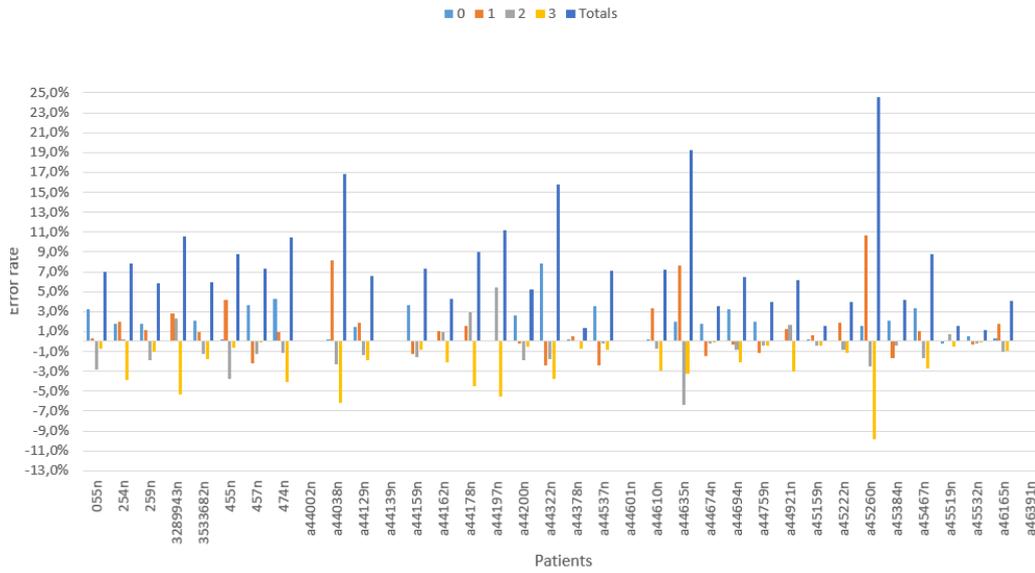


Figure 4.6: Weighted monitoring error of combined scores by patients in configuration (A). The error was weighted by the distribution of monitored combined scores.

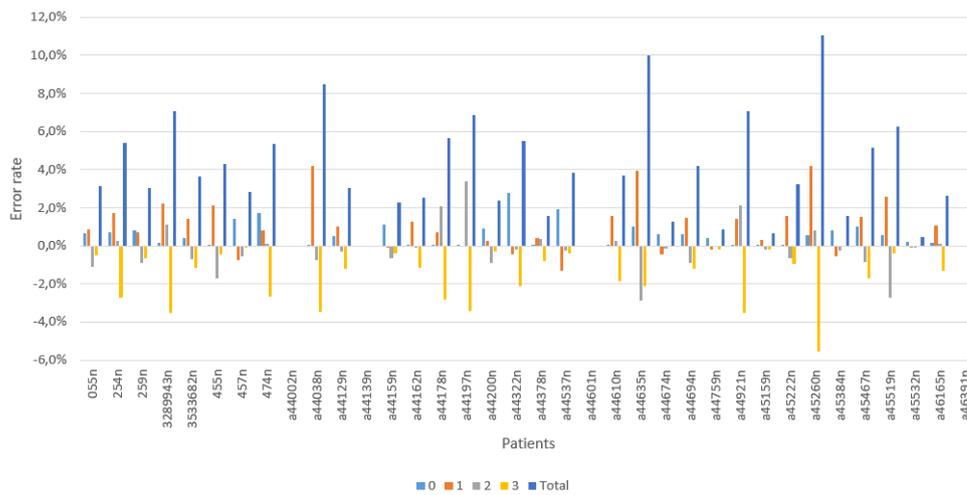


Figure 4.7: Weighted monitoring error of combined scores by patients in configuration (B). The error was weighted by the distribution of monitored combined scores.

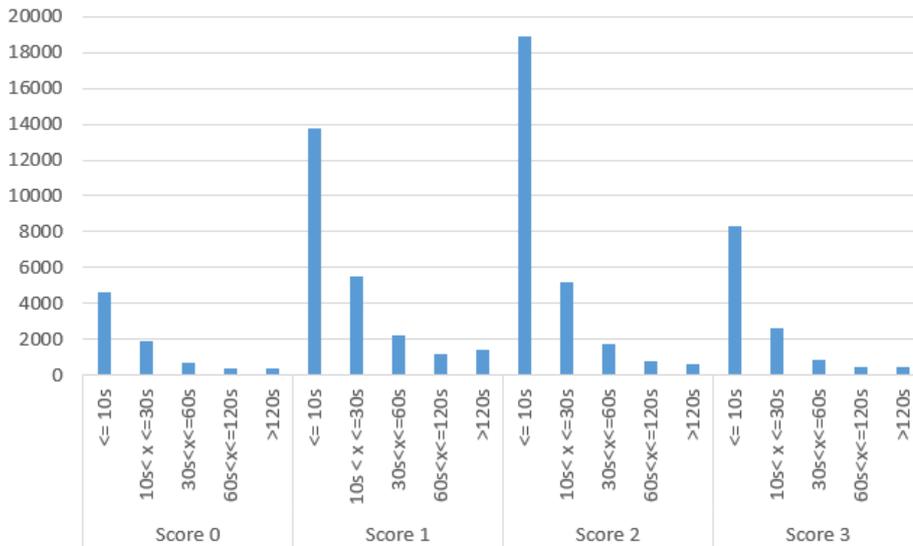


Figure 4.8: Distribution of intervals in the reference system by combined scores.

time windows. The total error is the proportionality of missed alarms to the total number of alarms. The patients w/ err variable is the number of analyzed patients who contained errors. The highest error is the highest proportion of alarms not triggered among all patients.

The same class variable is the percentage of alarms within the given interval of the same class of the reference alarm. Increasing sampling rates, reducing period sizes, and slacking the rules to avoid insignificant alarms in configuration (B) shrink errors. In a 2-minute interval, only 1.8% of alarms were not captured by the experimental configuration (B), and an absolute maximum error of 6.5% in the worst case for a specific patient. For configuration (A), the accuracy of alarms was 89%, and for (B), it was 94%. With an interval of 10 minutes, the total error was below 2.0% in both configurations, with the highest error found in one patient reduced from 12.3% to 2.8%.

Moreover, analyzing configuration (B) where the alarm delay avoidance parameter was set at 5s (Table 4.3), it was possible to avoid about 34% of negligible alarms (4,017) in relation to the total number of alarms (11,721) triggered by our algorithm using this configuration.

4.3.4

Experiment I - Discussion

In this experiment, it is proposed a new self-adaptive method aimed at monitoring vital signs remotely and reducing the waste of computational resources, reducing samplings by more than 87% from the baseline, which outperformed previous algorithms. Our algorithm reduced data reads by 64% of the second-best algorithms (ELGHERS; MAKHOUL; LAIYMANI, 2014),

Table 4.7: Alarms for transition duration $\geq 5s$ (19,433 alarms) by time distance of experimental alarms from reference’s alarms.

	(A)			(B)		
	600s	300s	120s	600s	300s	120s
Coverage						
Missed alarms	374	734	866	89	163	351
Total error	1.9%	3.8%	9.5%	0.5%	0.8%	1.8%
Patients w/ err.	32	32	32	26	28	32
Highest error	12.3%	15.9%	27.7%	2.8%	5.8%	6.5%
Same class	89%	89%	89%	94%	94%	94%

(HABIB et al., 2016), and about 48% of the total payload achieved by (HABIB et al., 2016). Our algorithm also reduced the number of alarms based on NEWS-2 combined scores by 78% when compared to the baseline.

Our method utilized combined scores allowing lower frequencies for all sensors when the clinical condition is at low risk, even when one or more vital sign has a high variation in raw data. In addition, when the clinical status is worse, it uses shorter periods to re-assess the patient’s health situation, capturing improvements faster. Thus, allowing reductions in frequencies promptly and more frequently. Moreover, when data is very similar, frequencies are also reduced to mitigate the production of redundant information. Thus, the algorithm enhanced data reductions.

Wearable sensor technology for monitoring patients’ vital signs continuously is fundamental to capturing events not realized by health teams (WEBSTER; SCHEEREN; WAN, 2022) (WEENK et al., 2017) (SAAB et al., 2021), but it also produces massive data (MICHARD; SESSLER, 2018). The baseline of our experiment produced more than 15.5 million data points. Extending the monitoring to a large number of patients, such as for the treatment of people suffering from chronic diseases or during pandemics, the amount of produced data will represent a burden in communication, processing, and storage infrastructure. Our proposal reduced the payload compared to the reference system by 97% and 99% using two distinct configurations, which is an impressive shrinkage in the volume of generated data to edge and cloud applications. Despite such a huge reduction, processing data in the wearable devices allowed a small absolute loss of alarm situations of only 0.8% using configuration (B) in a 2-minute time window.

Although our algorithm achieved a considerable reduction in sampling data, it did not significantly affect the integrity of monitoring time in each combined score (risk level) concerning the baseline. Differences in the distribution of total recorded time in each combined score were less than 2%. However, the absolute difference in the total recorded time of combined score 3 reached

15.0% in configuration (A) (Table 4.6). In this configuration, there is a larger difference between the range of frequencies of score 3 and the other scores that are parameters of the self-adaptive algorithm. The minimum frequency is half of score 2, a quarter of score 1, and one-sixth of score 0 (Table 4.3). Nonetheless, with configuration (B), matching registered times improved to a maximum absolute difference of 9.8% and recorded time distribution differences of 1.3% or less because of the more similar ranges of frequencies and period lengths for all combined scores.

Observing only the differences in the distribution of the monitored data provides a comprehensive but distant view of what might be happening in the monitoring of each patient. Thus, analyzing the error of monitored time of each combined score in each patient using configurations A and B (Figures 4.6 - 4.7) shows that the weighted CEWS error rate compared to the baseline for most patients was between +/-5% in configuration A and between +/- 3% in configuration B. Only in one case, in the patient (a45260n), the error rate reaches 10.7% in a combined score of 1 (24.7% of monitored time), -9.8% in a combined score of 3 (37% of monitored time), and only 1.6% and -2.5%, in combined scores 0 and 2 (7.7% and 30.4% of monitored time), respectively in configuration A. In the same patient using configuration B, the monitoring error was 0.6%, 4.2%, 0.8%, and 5.5% in the recorded time of combined scores 0-3, respectively, compared to the baseline. This patient also had the highest error among all patients with configuration B.

Observing the total absolute error rate that sums up each CEWS error rate for each patient, the mean among all patients in configuration A was 6.8% (+/- 5.5%) while it decreases to 3.8% (+/- 2.8%) in configuration B. Thereby, the total distribution error analysis provided only a rough picture of errors during monitoring, and the individual analysis must be considered when configuring the parameters of the algorithm to mitigate differences.

Moreover, Figure 4.8 shows the distribution of segment intervals by combined scores from the reference system. It is noticed many short intervals lasting 10s or less. Short intervals may also justify the total time differences registered by our solution and the reference system because our algorithm will filter those intervals using lower sampling frequencies and not realize small segments in different score values. For example, patient a45260n with the highest weighted error rate, 47%, 74%, 79%, and 64% of segments in scores 0-3, respectively, last less than 10s. However, using different configurations, our algorithm can potentially regulate the desired integrity level.

Furthermore, even with the huge reduction in data samplings, our algorithm could set off alarms within a time window of up to 2 minutes and

miss 9.5% of alarms in configuration (A) and only 1.8% in configuration (B). In a five-minute time difference, only 3.8% and 0.8% of alarms were skipped in configurations (A) and (B), respectively. In configuration (B), the highest alarm missed detection rate in one patient was 6.5% related to the reference system (Table 4.7). It is important to highlight that it is selected alarms arising from transitions that lasted 5s or more in the reference system. The choice of 5s is in sync with other initiatives to reduce non-actionable alarms (PATER et al., 2020). Moreover, a 2-minute or 5-minute alarm window seems a long time. However, an observational study (SAAB et al., 2021) with 782 postoperative patients reported that hypotensive and desaturation episodes are mostly missed because vital sign assessments on surgical wards are sparse in best cases in a 10-minute interval. Furthermore, compared to the bed visits performed primarily at 8 hours intervals in a general ward (LEENEN et al., 2020) (WEENK et al., 2017), it might improve the quality of provided care in those situations. Not to mention that most of the patients in general wards or at home are not monitored at all. Then, a maximum error rate between 2.8 and 6.5% (10- and 2-minute windows), as achieved by our algorithm using configuration (B) for any patient, seems an enhancement in healthcare services. A lower error can be achieved by increasing monitoring frequencies but reducing the gains in data reduction.

The number of alarms in our experimental setup using the NEWS-2 combined scores was 35,996 for 36 monitored patients, which is in sync with the literature that reports 946 alarms per patient a day in hospital units (PURBAUGH, 2014). The nursing staff can not handle this sheer number of alarms. It causes health teams to distrust the alarms' importance and to ignore them (KORNIEWICZ; CLARK; DAVID, 2008). Our proposal significantly reduced the number of alarms in configurations (A) and (B) by 78% and 67%, respectively. Addressing this problem in a solution embedded in wearable devices brings the additional benefit of reducing transmissions and extending battery life.

Our experiment also verified two strategies to reduce alarms aimed at filtering redundant and negligible alarms. The redundant lockout time window strategy better-avoided alarms than the delay approach. It avoided similar alarms within a time frame. The delay strategy in configuration (A) with 10s avoided 67% fewer alarms than the other approach, and configuration (B) with 5s also avoided about 61% fewer alarms than the redundant avoidance approach. Curiously, configuration (B) with a more permissive policy with a shorter time frame to avoid redundant alarms, the 60s instead of 90s, and a shorter delay, 5s instead of 10s, had higher reductions than configuration (A). It

can be explained because configuration (B) also increases the sampling rates for all NEWS-2 combined scores, which leads to more alarm situations. Nonetheless, the strategies effectively reduce redundant and insignificant alarms and may unload the extra burden on caregivers and health teams (FERNANDES et al., 2019).

Finally, it is checked whether the closest triggered alarms corresponded to the same combined score alarms as the reference system. The accuracy was between 89% and 94% for all scenarios (Table 4.7). Moreover, combined scores better qualify alarms compared to thresholds because they provide a degree of severity in alarms.

4.3.5

Experiment I - Conclusion

This experiment applied our approach based on reducing frequencies when the patient's clinical condition improves, or data is very similar. The algorithm mitigates the massive data production of constant monitoring of vital signs and controls the number of alarms with strategies to filter negligible alerts and avoid redundant alarms. Alarms were handled using composite scores, which also governed the adaptive sampling algorithm. Alarms missed detection rate and accuracy metrics are effective in comparing different configurations. Moreover, the data reduction results outperformed previous works in both configurations. The following experiment intends to test our algorithm in a hardware prototype and evaluate potential benefits in energy consumption.

4.4

Experiment II - Self-adaptive approach effects on energy efficiency

The first two experiments demonstrated the evolution of our approach of using combined scores to address the issues of massive data generation and excessive alarms in IoT-PMAs. This experiment focuses on the power requirements in wearable devices and the potential to save energy using an embedded smart solution with adaptive sampling.

The choice of measuring an oximeter that can acquire the arterial blood saturation and cardiac pulse and a temperature sensor is due to these health markers becoming very popular during COVID-19. Low saturation and fever were characteristic symptoms of several infected patients (KONG et al., 2020). Additionally, our self-adaptive algorithm presented the capacity to drastically reduce readouts, as demonstrated in experiments I and II. Therefore, this experiment aims at verifying the potential gains in energy savings based on measures performed using our wearable prototype.

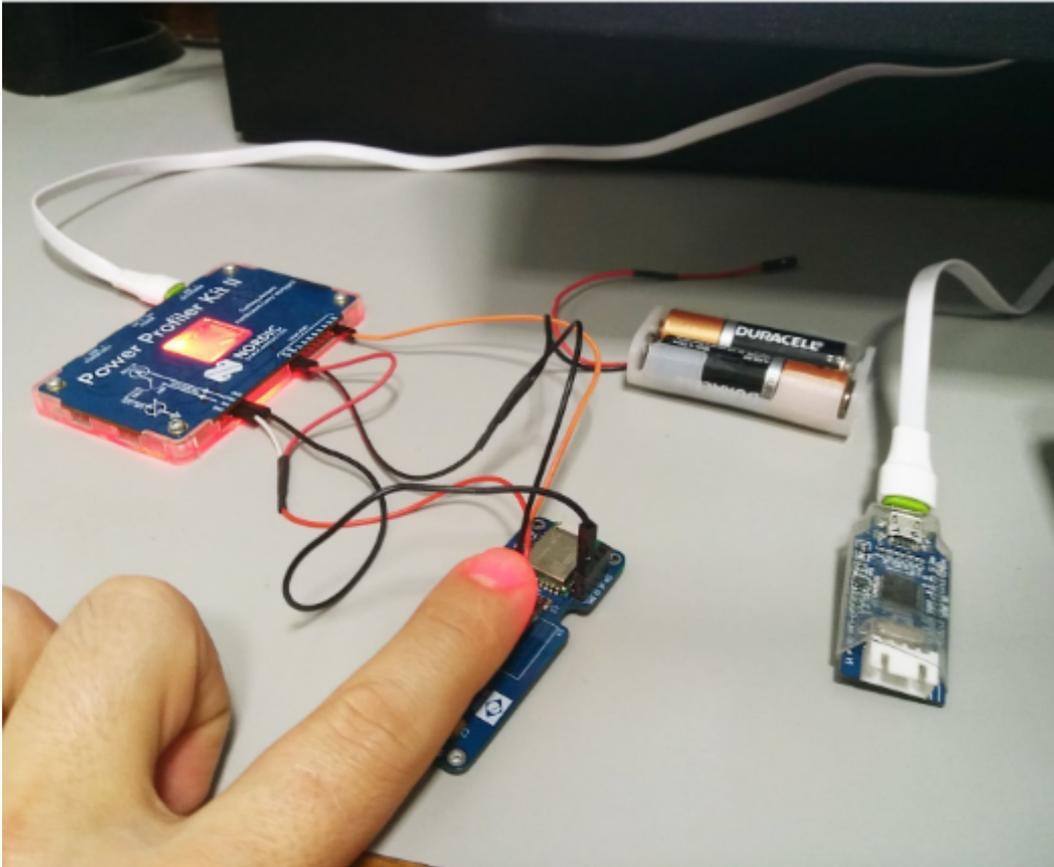


Figure 4.9: Collecting energy measurements using the PPK-II and our hardware prototype.

4.4.1 Experiment II - Objectives

The main objective of this experiment is to measure and compare the energy requirements of physiological sensors controlled by our self-adaptive algorithm and at a fixed sampling rate as utilized in infirmaries.

4.4.2 Experiment II - Methodology

A general description of our hardware prototype is given in Subsection 3.5.3 and its development in Appendix B.

It is utilized the NRF Power Profiler Kit II (PPK-II)⁴ to capture in real-time the energy consumption of components during the wearable operation. The PPK-II connects to our device via the power supply pins providing a voltage of 3v, such as two 1.5v batteries, as projected in our prototype. Then, the PPK-II can measure and record the utilized energy during the device's operation. Figure 4.9 shows the procedure to capture energy requirements.

⁴<https://www.nordicsemi.com/Products/Development-hardware/Power-Profiler-Kit-2>

Additionally, an extra connection between the PPK-II to an output pin in our microcontroller board was configured. Using this connection, it is possible to signalize specific events to the PPK-II. So, our code is instrumented to indicate the limits of each target event via this output. The pseudo-code below presents these events for the temperature sensor reading function. In line 2, it starts a signal, and in line 4, it ends it to characterize the sensor's wake-up task, while in lines 5 and 9, the same process is done for the reading task. This signalization can be viewed at the bottom line of Figure 4.10 highlighted by the red arrows.

```

1 Function Read Temperature:
2   Signal up to mark the starting point
3   Wake up TEMPERATURE_SENSOR
4   Signal down to mark the ending point
5   Signal up to mark the starting point
6   curr_temp := read TEMPERATURE_SENSOR
7   temp_ind_score := Calculate curr_temp individual score
8   Puts to sleep TEMPERATURE_SENSOR
9   Signal down to mark the ending point

```

Afterward, the segment representing the read function is selected, shown by the black arrows delimiting the gray area in the graph. For simplification, it is accounted the time between the start of wake-up and the end of reading tasks in our measurement. The peaks between the target events are from other underlying tasks, such as the BLE listening wake-ups. The PPK-II integrates the consumption within the selected area and provides the highest, average, and total consumption besides the duration of that area at the right bottom of Figure 4.10. Finally, the same procedure was executed for the oximeter.

Having the energy requirement parameters, we utilized our simulator to project the energy consumption of data acquisition during a period of monitoring.

The read operations were recorded using the original dataset sampling rate at 1Hz as a baseline (86,400 data points) for comparisons and using our self-adaptive algorithm. Then, the measures of our hardware prototype were utilized to calculate the estimated energy consumption in each scenario.

Table 4.8 shows the configuration utilized in our algorithm to run the simulation with energy parameters collected from our hardware prototype.

The experiment utilized data from patients with three distinct clinical conditions: a stable patient (a44601n) with no deterioration in clinical status, an unstable patient (a44178n) who triggered about 1,000 alarms in 24h, and a



Figure 4.10: Selection of target events of the temperature sensor in the PPK-II graph interface.

Table 4.8: Parameters of our self-adaptive algorithm.

Parameters	Configuration for each clinical status (0,1,2,3)
Round Size	4
Period Size (sec)	150, 100, 60, 30
Max. Freq. (Hz)	1/3, 1/2, 1/2, 1
Min. Freq. (Hz)	1/30, 1/20, 1/10, 1/5
r0 - risk factor	0.2, 0.6, 0.8, 0.9

highly unstable patient (a44332n) with more than 2,000 alarms a day regarding the baseline system. One run was performed for each patient simulating 24h of monitoring.

Table 4.9 presents the total time recorded in each combined score of unstable patients in the baseline system. It shows that patient a44332n (highly unstable) spent most of the monitoring time with lower combined scores, 0 or 1. Then, this patient sampled fewer data than the patient a44178n with fewer alarms.

Although the experiment only collected actual energy information from the oximeter and temperature sensors, the NEWS-2 scores were calculated using all five vital signs (HR, SO, BTEMP, RR, and SBP) in our simulator. The NEWS-2 parameter for consciousness level was considered normal, and for chronic hypercapnic respiratory failure condition was absent for all patients.

4.4.3

Experiment II - Results

Table 4.10 details the energy consumption for one full readout and the time to acquire the first valid value of our sensors. It is considered to simplify

Table 4.9: Total monitoring time (seconds) by combined scores. The highly unstable patient spent most monitoring time in lower combined scores (blue). In contrast, the unstable patient spent the most time at higher scores (red).

Combined Score	a44178n unstable	a44332n highly unstable
0	12,120	14,119
1	15,396	40,591
2	46,818	23,277
3	12,066	8,043

Table 4.10: Sensors and BLE transmission energy consumption.

Tasks	Consumption (mJ)	Time
Oximeter read	666	6s
Temperature read	1.69	116ms
Transmission (BLE)	0.036	5ms

Patients	Status	Baseline			Adaptive Algorithm			Reductions	
		Total Reads	Oximeter (J)	Temperature (J)	Total reads	Oximeter (J)	Temperature (J)	Oximeter	Temperature
a44601n	Stable	86,400	9,590	146	5,760	1,922	5	80%	97%
a44178n	Unstable				25,224	4,471	26	53%	82%
a44332n	Highly Unstable				23,486	4,039	24	58%	83%

Figure 4.11: Energy requirements after 24h of monitoring in distinct patient profiles.

the time from the wake-up to the end of acquiring one valid value from the sensor.

Figure 4.11 shows the energy required for each patient regarding the baseline and using the self-adaptive algorithm. As shown, the proposed algorithm drastically reduced energy consumption for all cases. Columns (1-3) refer to the baseline system. Note that values are equal for all patients because all data points were captured at the same sampling frequency. Column (4) presents the number of data points acquired by the adaptive sampling for each patient. Columns (5-6) show the corresponding projected total consumption of the Oximeter and Temperature sensors, respectively. Finally, columns (7-8) show the percentage reduction of consumption compared to the baseline of the Oximeter and Temperature sensors, respectively.

4.4.4

Experiment II - Discussion

Yu et al. (YU; LI; ZHAO, 2021) point out that to estimate the total energy demand of a health detection solution, it is necessary to analyze

the energy consumption of each component of the system. This experiment measured the actual energy demand of some relevant components of our proposed system, starting with the oximeter and temperature sensors. It verified the effects of using our self-adaptive algorithm on energy efficiency. The preliminary results were motivating, reaching 80% and 97% of energy reductions in a stable patient for each sensor, respectively.

A stable patient can run at very low frequencies, while an unstable patient will vary frequencies between higher values providing more reads and consequently consuming more energy. The highly unstable patient had fewer readouts than the unstable one because this patient had more monitoring time in a lower combined score. So, using lower sampling rates. Following our algorithm principles, the most determinant factor resulting in a lower number of readouts is the time the patient is classified with lower scores. Moreover, the strategy to reduce frequencies when data is very similar avoids redundant data in sampling, an important feature to save energy (HASAN et al., 2019). Therefore, the second determinant factor is the similarity of monitored data.

Several authors assumed that the higher energy consumption is due to transceiver operation (ELGHERS; MAKHOUL; LAIYMANI, 2014) (HABIB et al., 2016) (FILIPE et al., 2015). Little attention is given to sensors' data acquisition energy requirements as, typically, the experiments use low-energy sensors. However, the oximeter is based on the photoplethysmogram that uses a light-emitting diode and photodiode (YU; LI; ZHAO, 2021). In our experiments, the energy cost of the oximeter seems to be a candidate for a major contributor to our prototype energy requirements. The total consumption of the oximeter sampling data at 1Hz during 24h accounts for about 9-21% of the capacity of a typical AA battery with 2,000 mAh. At the same time, the temperature sensor at the same rate is responsible for about 0.1%.

Moreover, when the Oximeter sensor first wakes up until the first valid read, it takes about 6s and consumes 666 mJ, while the temperature sensor only 1.69 mJ and takes 116ms. For the oximeter, reads performed in intervals shorter than 6s were computed as a fraction of the consumption of the first valid read in our simulation.

Furthermore, the cost of transmitting a single packet was measured using Bluetooth Low Energy with the values read from the oximeter and temperature sensors. The energy cost was only 0.036 mJ in 5ms using an already-established connection in a one-hop direct communication to a gateway. It is about 50 times less the energy required for reading the temperature sensor. This type of connection utilizes a highly energy-efficient protocol where the radio does

not work continuously.

Although our energy measures can be used as a reference for other applications and comparisons, it presents some limitations. All measurements are tied to our software implementation and applied hardware configuration. Energy models and benchmarks can be more appropriate to utilize in generic applications. However, our main objective in this study was to compare self-adaptive features in distinct scenarios using the same baseline configuration.

4.4.5

Experiment II - Conclusion

This experiment utilized our hardware prototype and our self-adaptive proposal using combined scores to project the energy savings promoted only on data acquisition tasks.

The algorithm achieved higher energy savings in healthier patients because they spent more time with lower combined scores. Nonetheless, it was not accounted for energy requirements for the alarms. However, transmissions are unlikely to compete in energy requirement with the physiological sensors spending only a tiny fraction of energy when comparing both transactions.

Finally, the self-adaptive algorithm using combined scores demonstrated a high potential to reduce energy requirements in wearable devices for IoT-PMAs.

4.5

Experiment III - Effects of proposed self-adaptive algorithm principles on patient monitoring and a comprehensive investigation of energy efficiency

This experiment analyzes the effects of our proposed principles to rule the self-adaptive algorithm on vital-signs monitoring. Based on those effects that provide insights on energy savings according to patient profiles.

Moreover, the energy consumption of sensor acquisition, periodic and alarm transmissions, processing of alarm evaluation and avoidance alarm routines, and self-adaptation tasks were measured from our wearable kit (WKit). Additionally, sleep mode periods were also considered in our trials, covering all functionalities of our proposed solution using our hardware prototype.

All energy measurements were performed using the BLE and the Ultra-Low Power Radio (ULPR) versions aimed at monitoring patients at home and in infirmaries, respectively. Simulations were performed considering both scenarios and using data from public datasets at the original sampling rate and defined by our self-adaptive algorithm.

4.5.1

Experiment III - Objectives

The objectives of this experiment can be enumerated as follows:

1. Analyze the effect of the application of the four proposed principles on patient monitoring data.
2. Analyze and compare the energy requirements for all tasks executed by our proposed self-adaptive algorithm in the WKit.
3. Analyze and compare the project energy consumption of our wearable in two different scenarios, within infirmaries (ULPR) and at home (BLE).

4.5.2

Experiment III - Methodology

The same set of records utilized in Experiment I was chosen for this experiment. They are described in subsection 4.1.1. Further, the configuration (A) of the same Experiment described in Table 4.3 was also used to project the energy consumption during monitoring.

To address the primary objective of this experiment, we will analyze the patients' monitoring data using metrics that reflect the principles governing the self-adaptive algorithm. Thus, it was verified the mean combined score, which is the sum of all applied combined scores divided by the number of periods (P1). The combined scores represent the patient's risk of worsening. It is a major force in our self-adaptive algorithm to increase monitoring frequencies and the number of periods (re-assessments). In addition, the number of periods itself was analyzed for each patient (P4).

Furthermore, the mean dissimilarity index was also analyzed (P2). It is the sum of found dissimilarity index in each period divided by the number of periods. It is also positively correlated to applied frequencies. Interesting effects can be found in the mix of combined scores and dissimilarity index on frequencies and energy savings.

Additionally, the mean frequency is the sum of all employed frequencies by the number of periods. This index theoretically would be directly related to energy consumption.

Finally, the effect related to the acceleration/deceleration of frequencies needs a specific investigation. Thus, comparisons between the utilized Bézier distribution and a Linear distribution were performed to understand the behavior in actual patient monitoring scenarios (P3).



Figure 4.12: NRF Power Profiler Kit II chart. Acquisition of energy requirement of the self-adaptive procedure in the WKit BLE version.

To address the second objective of this experiment, we collect the energy measures following the procedure outlined in Experiment II (Section 4.4) with the NRF Power Profiler Kit II (PPK-II).

Figure 4.12 shows an example of captured information. The chart is generated in real-time and saved. Then, it is possible to select the area following the signalization generated by our code in a separate channel (0), shown at the bottom of the chart. The two black arrows in the x-axis limit the interested area, and the PPK-II software generates the consumption cost automatically. All measurements will be collected using the ULPR binary version, followed by the BLE version. The measures will be equivalent to one operation of each task.

Once the measurements are collected, they will feed our simulator that will run the baseline and our proposed solution procedures using all vital signs. Each task execution will be counted and registered. After the monitoring time ends, the WKit total energy requirement will be estimated by multiplying the unit energy consumption acquired in the previous step by the correspondent task counter. In addition, the related total execution time for the task will also be found by multiplying the number of executions by the unit operation execution time. Then, the results of all tasks' execution times will be summed up. The summed time will be decreased from the total monitoring time, and the result will be multiplied by the energy expenditure of the sleep mode. A basic measurement of the sleep mode was acquired and adjusted to one second. Then, for calculating the total energy consumption in sleep mode, the

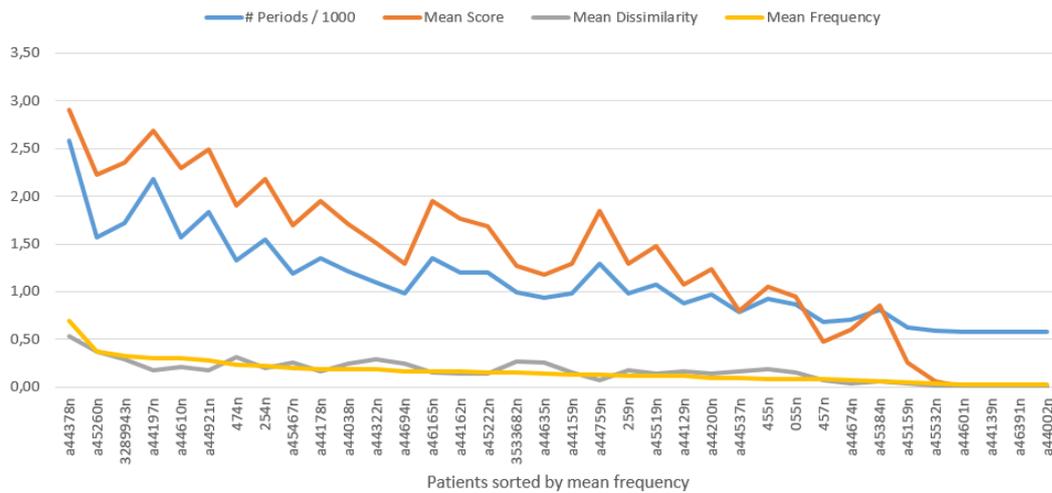


Figure 4.13: Metrics to analyze the application of the four principles' effects on monitored data. Patients are sorted by the mean frequency (yellow) from the highest to the lowest.

time units are adjusted accordingly.

The energy consumption analysis will be performed using these records, considering the reading, processing, and transmitting operations using both implemented versions (ULPR and BLE). Comparisons of these versions are performed to the baseline system and between them.

Finally, an investigation of reading, processing, and transmitting operations was executed to demonstrate the weight of main WKit operations and discuss the inclusion of our proposal embedded in the WKit.

4.5.3 Experiment III - Results

Figure 4.13 shows the mean score in orange. The number of periods (divided by 1,000 to visualize the chart better) is shown in blue. The number of periods represents how many times the self-adaptive procedure has run. It also shows the mean dissimilarity index in gray, which may represent the amount of redundancy in data. The higher the dissimilarity index, the higher the variability of vital signs, and vice-versa. Finally, the mean employed frequency in yellow was used to sort the patients from the highest to the lowest mean.

Table 4.11 shows the comparison between the Bézier and Linear distributions of frequency progression according to the values of the dissimilarity index. For combined scores greater than zero, Bézier increases frequencies more quickly than the Linear distribution. The higher the score, the higher the acceleration in frequencies. In contrast, with a combined score of zero, frequencies increase slower than the Linear progression.

Table 4.11: Progression of Bézier (Béz) and Linear (Lin) distribution of frequencies (Hz) by the dissimilarity index.

D-Idx	Score 3		Score 2		Score 1		Score 0	
	(Béz)	(Lin)	(Béz)	(Lin)	(Béz)	(Lin)	(Béz)	(Lin)
0.001	0.20	0.20	0.10	0.10	0.05	0.05	0.03	0.03
0.101	0.33	0.20	0.12	0.11	0.06	0.06	0.03	0.04
0.201	0.33	0.25	0.17	0.12	0.07	0.06	0.04	0.04
0.301	0.50	0.25	0.20	0.12	0.08	0.07	0.04	0.05
0.401	0.50	0.33	0.25	0.14	0.09	0.08	0.04	0.05
0.501	0.50	0.33	0.25	0.17	0.11	0.09	0.04	0.06
0.601	1.00	0.33	0.33	0.20	0.12	0.11	0.05	0.07
0.701	1.00	0.50	0.33	0.25	0.17	0.14	0.05	0.09
0.801	1.00	0.50	0.33	0.25	0.20	0.17	0.06	0.12
0.901	1.00	1.00	0.50	0.33	0.33	0.25	0.10	0.17
1.000	1.00	1.00	0.50	0.50	0.50	0.50	0.33	0.33

Figure 4.14 depicts the sum of applied frequency differences between simulations running the self-adaptive algorithm with the Bézier distribution and with a Linear distribution (risk factor = 0.5 for all scores). The distributions are calculated in real-time according to the previous period's combined scores which define the correspondent range of frequencies and the risk factor. Then, the applied frequency in each period was summed in each case during the simulations. In this Figure, it can be noticed the effect of the elevated risk factor of combined scores greater than zero on the sum of all frequencies utilized during each period along the trial time. In patients with high scores and high dissimilarity, using the Bézier distribution, the sum of frequencies was increased up to 58% depending on patient-monitored data, and for healthy individuals, there were no differences to the Linear distribution because the combined score remained zero during the whole monitored time.

Table 4.12 shows the energy consumption measures acquired from our hardware prototype tasks running the ULPR and BLE. These measures were utilized in our simulations to project the energy consumption during the monitored time with distinct patient profiles.

Figure 4.15 shows the total energy consumption for each record, comparing the baseline and running the ULPR version of the self-adaptive algorithm. The gray line shows the percentage of the economy, which varies from 32.1% to 81.5%. The baseline system does not consider the patient's health condition and the similarity of data, having fixed sampling and transmitting rates. Therefore, the energy requirements are the same for all records. For the baseline system, the total energy cost was 10,731 J. Yet, using our self-adaptive algorithm, it varies from 1,981 J to 7,290 J.

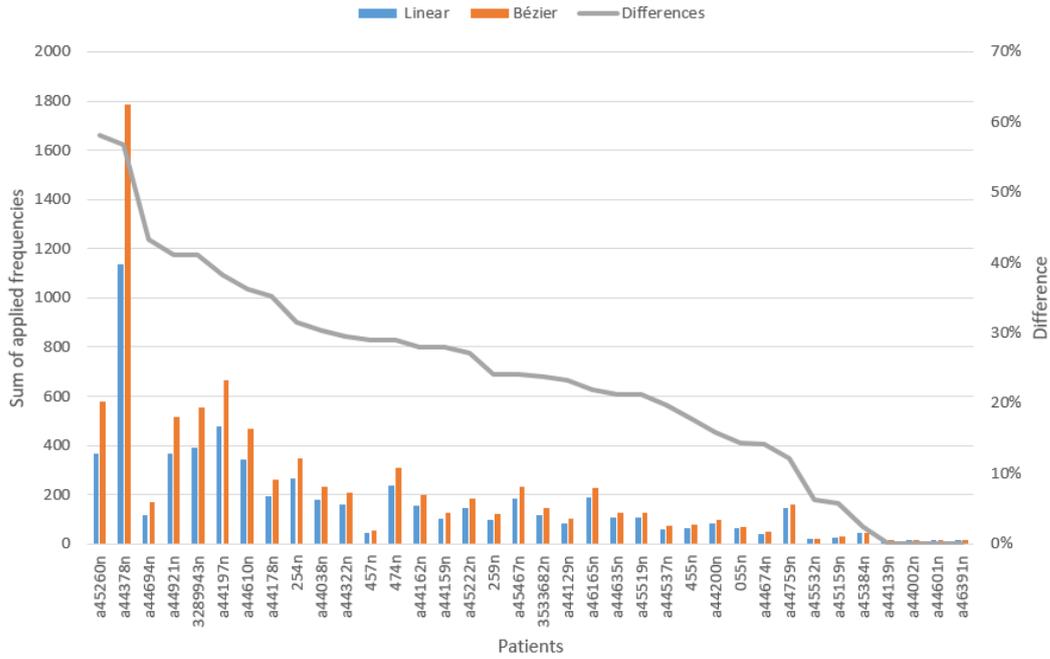


Figure 4.14: Bézier and Linear distributions, the sum of all applied frequencies, and their differences (%) for all records.

Table 4.12: Energy measures from WKit ultra-low-power radio (ULPR) and Bluetooth Low Energy radio (BLE).

Tasks	ULPR		BLE	
	mJ	time (s)	mJ	time (s)
Temperature sensor wakeup	0.555	0.035	0.861	0.03485
Temperature sensor read	0.027	0.0014	0.021	0.00127
Oximeter (wakeup + read)	618	5	557	5
Transmit periodic data	0.02	0.004	0.0006	0.0001
Process alarms	0.00004	0.00001	0.00025	0.00001
Transmit alarms	0.023	0.005	0.001	0.00008
Process self-adaptation	0.001	0.00018	0.001	0.00011
Sleep mode	2.769	1	2.304	1

Figure 4.16 shows the energy requirements for each record, comparing the baseline and running the BLE version of our self-adaptive algorithm. The gray line shows the percentage of the economy, which varies from 32.1% to 81.7%. The WKit total energy cost was 9,696 J for the baseline system, varying from 1,771 J to 6,582 J for our proposal. The BLE version was 10.7% more economical than the ULPR version in the baseline system. Moreover, using our self-adaptive algorithm, economy levels compared to the ULPR version vary from 9.7% to 10.6% in relation to the baseline.

Figure 4.17 compares the WKit total energy consumption of the ULPR and BLE embedded versions projected in our simulator and their differences for all records.

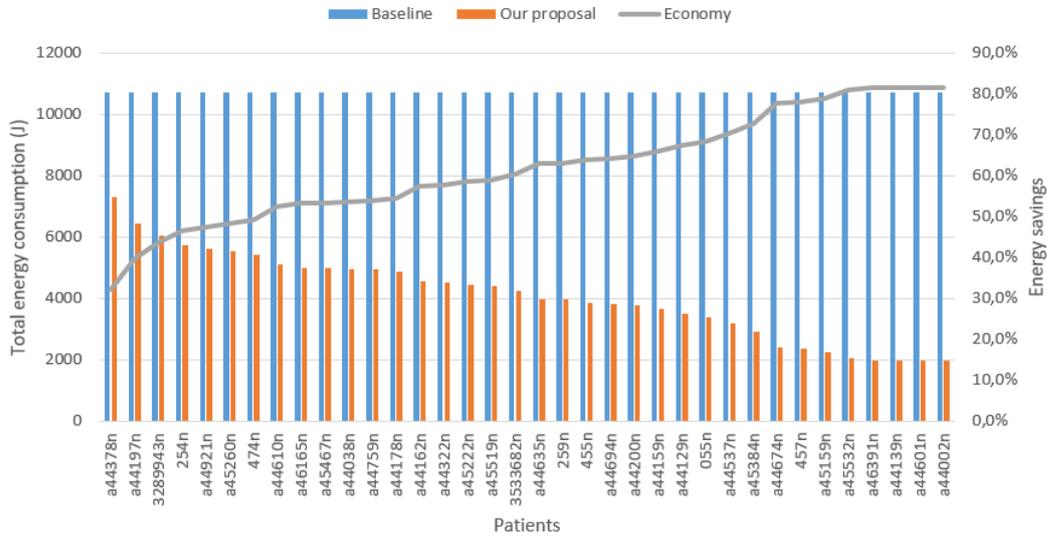


Figure 4.15: WKit total energy consumption on all records using the ULPR radio. Comparison between the baseline system and our proposal.

PUC-Rio - Certificação Digital N° 1912730/CA

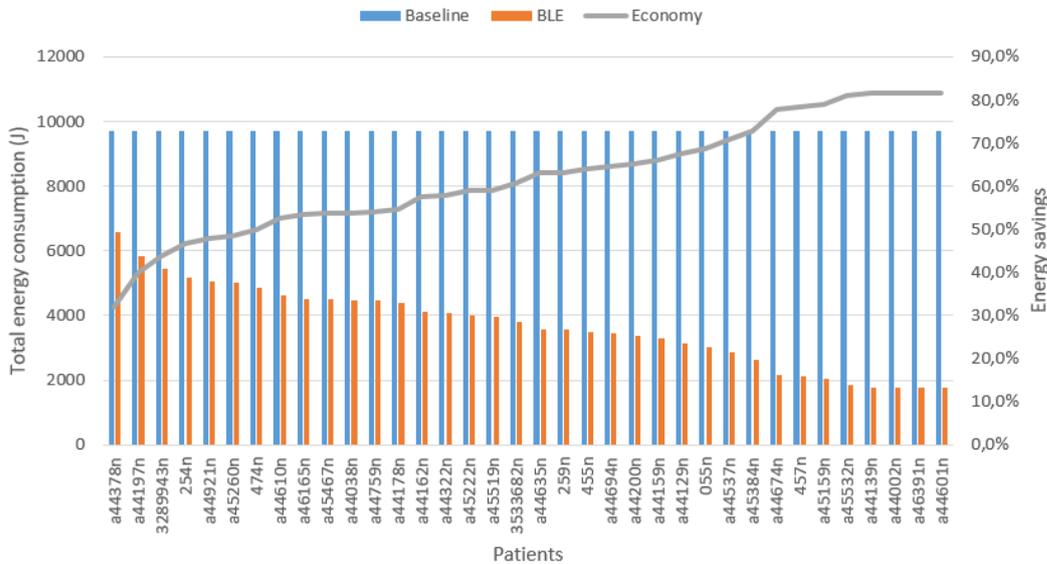


Figure 4.16: WKit total energy consumption on all records using the BLE radio. Comparison between the baseline system and our proposal.

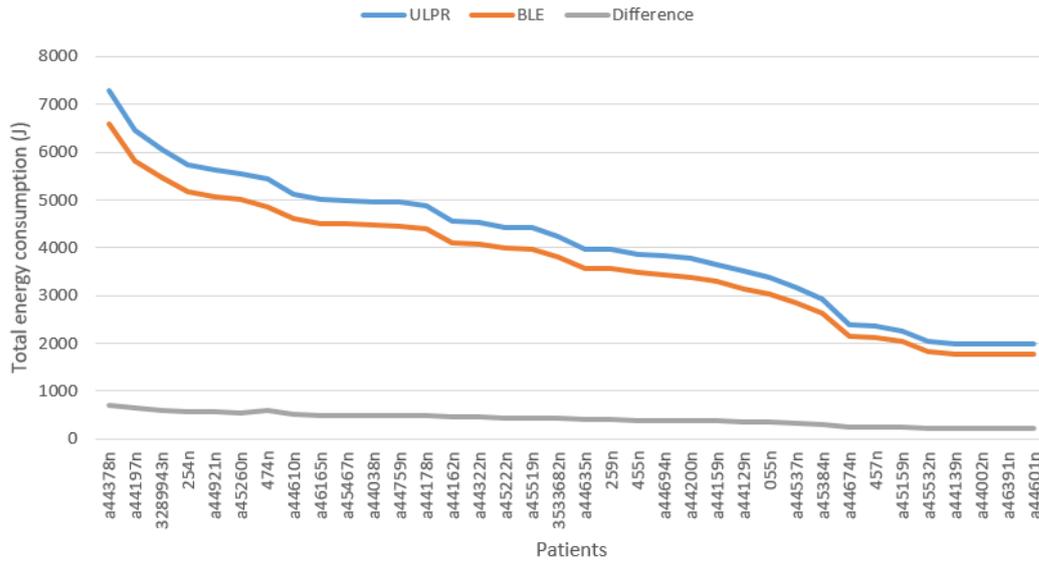


Figure 4.17: WKit total energy consumption on all records using the ULPR and BLE radios running the self-adaptive algorithm.

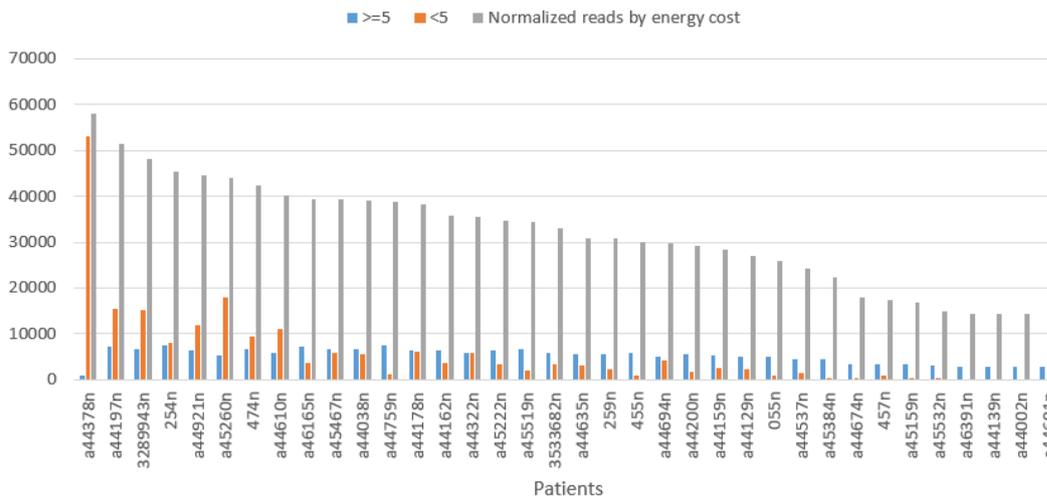


Figure 4.18: Number of oximeter reads classified by intervals between readouts $<5s$, $\geq 5s$, and reads weighted by energy cost.

Figure 4.18 depicts the number of oximeter reads that occurred in intervals between readouts shorter than 5s (orange), 5s, or longer (blue) and an index (gray). The oximeter read index is composed of the number of reads that occurred in intervals greater than or equal to 5s multiplied by 5 plus the number of reads that occurred in intervals shorter than 5s. In our simulator, the energy required to read the oximeter in intervals shorter than 5s were considered 1/5 of the required for the ones performed in 5s or longer intervals. The oximeter index ranked the patients from the highest to the lowest, as shown in the chart.

Figure 4.19 compares the WKit total energy required for transmitting messages and alarms versus processing composite scores, the self-adaptation

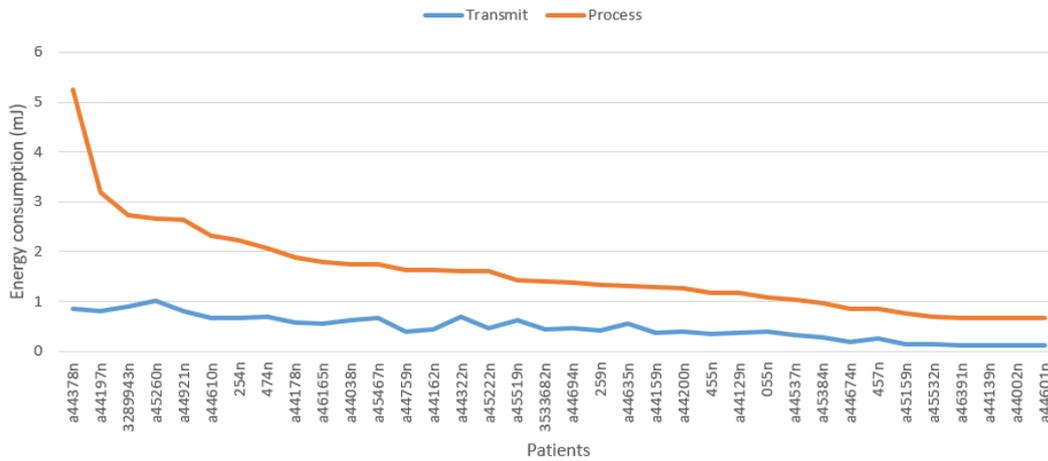


Figure 4.19: WKit total energy cost for transmitting periodic data and alarms versus processing self-adaptive and alarm assessment processes. BLE version - self-adaptive algorithm.

of the system, and managing alarm conditions. The BLE-embedded version was utilized with our self-adaptive algorithm to execute this simulation. Processing tasks required 2.3 to 6.1 times more energy than transmitting data.

4.5.4 Experiment III - Discussion

The first objective of this experiment was to verify the effects of the four principles (Table 3.1) that rule our self-adaptive algorithm on monitoring samplings, as listed below:

P1 - the highest the score, the highest the frequency, and vice-versa. In Figure 4.13, the yellow line represents the mean frequencies. It sorted the patients from the highest to the lowest mean in this chart. Patient a44378n (mean frequency of 0.69, mean score of 2.91) in the left spent almost 87% of monitoring time in combined score 3. On the other hand, the four patients on the right spent 100% of the time with a combined score of 0. Comparing the order of the patients in this figure with Figure 4.2 in Subsection 4.1.1 can be noticed the effect of combined scores on employed frequencies. Patients who spent more time with higher scores had, with few exceptions, higher frequencies. The effects of principle P2 may explain these exceptions.

P2 - The highest the dissimilarity, the highest the frequency, and vice-versa. The second patient in the right side of Figure 4.13, a45260n (mean frequency of 0.367), is not the one with the second highest mean score (orange line, mean score=2.23). There is a patient (a44197n)

with a higher mean score (2.69, mean frequency=0.304) and a patient (3289943n) with a mean score (2.35, mean frequency=0.321). However, both patients have a much lower dissimilarity index, 0.17 and 0.29, respectively, while patient a45260n has a dissimilarity index of 0.38, increasing the applied frequencies. The number of alarms of patient a45260n was 3,003, while the other patients had 812 and 790 alarms in the baseline system, respectively (Figure 4.3). This may indirectly indicate the higher number of transitions and the potential dissimilarity between periods. Another example is patient a44921n, who had the third highest mean score (2.49) but a low dissimilarity mean (0.17) and only appeared in the sixth position, with a mean frequency of 0.281. These examples demonstrate how the dissimilarity index modulates the use of frequencies offsetting cases with very high or low dissimilarity in time series.

P3 - In higher combined scores, frequencies will increase faster and vice-versa. In our experiments, the risk factor was configured as 0.2, 0.6, 0.8, and 0.9 for each combined score (0,1,2,3), respectively. Table 4.11 shows that the utilized Bézier curve increases frequencies more quickly than a linear progression, except for a combined score of zero (risk factor = 0.2). Then, as the combined score is higher, the faster the frequencies rise. The effect can be seen in Figure 4.14 showing that patients that spend more time in higher scores increased their frequencies up to 58% above the correspondent simulation running using a linear distribution (risk factor = 0.5). The combined score and the amount of dissimilarity will increase the frequencies. Consequently, acquiring more data. The risk factor is a parameter that can be configured in our solution. By reducing the risk factor for each combined score, there will be a reduction in samplings and energy consumption, but with a potential effect on patient monitoring and alarm quality.

P4 - The higher the combined score, the shorter the interval to re-assess the patient's condition and vice-versa. It can be seen that the number of periods (blue) follows the mean score (orange) line. Then, the magnitude of the number of periods is directly proportional to the mean of scores, as expected. Re-assessing combined scores more frequently in higher scores may capture further deterioration or improvements more quickly. The self-adaptive algorithm using the configuration parameters will assess the patient's condition at least 577 (1 every 2:30 minutes) and about 2,600 times (1 every 33s) in 24h of monitoring,

according to the patient's health condition using the proposed configuration settings.

Knowing the effects of the four principles on monitored patients' data provide a better understanding of the mechanisms that reduce data generation and energy consumption using our proposal. Moreover, it can be used to manipulate configuration parameters. For instance, reducing the risk factor for each combined score to reduce data generation.

In addition, this experiment utilized actual energy measurements for all operations performed in the WKit from two different embedded versions, the ULPR, and the BLE. Using these measures, we projected the WKit total energy consumption of our proposed solution for monitoring patients for 24 hours. We compared the results to a baseline system running at a standard sampling rate, such as those used in ICUs.

The records were sorted by the mean frequency of our proposed solution during the experiment in Figure (4.13). So, the records with the higher mean frequencies should be in the group with the higher energy consumption. However, comparing the order of patients in Figures 4.13 (frequency mean sorting) and Figures 4.15 and 4.16 (energy consumption sorting), it can be noticed that the records are not in the same sequence. The patients who sampled more data are not necessarily those whose devices consumed more energy.

The oximeter read is the highest energy cost (Table 4.12) and lasts 5s to acquire the first valid value. In our simulator, when oximeter reads occur in intervals between readouts shorter than 5s, the energy accounted for these reads is a fraction (1/5) of the energy for the first valid read. In Figure 4.18, the patients are ordered by the total reads, taking into account this calculation procedure. It can be realized that patients' order is the same as the WKit total energy requirement, such as in Figures 4.15 and 4.16. Therefore, for our implementation, in terms of energy requirements, not only the amount of data read accounts but also the intervals at which they were processed.

In real scenarios, reading intervals shorter than the minimal time to return the first valid value interfere with the implementation of duty cycle mode in the sensor driver. It is necessary to create a queue of requests or other signalization and return the last valid read value or an error when these reads in short intervals occur. Configuring the sensor to sleep after acquiring a valid measure can not be used without checking if new requests for reading values from the sensor have arrived.

The energy cost of the oximeter read in comparison to Experiment II is lower because the time to capture the first valid read was reduced from 6s to

5s. Therefore, it lowers the energy cost for the first readout from 666 mJ to 557 mJ. This generated a significant energy economy. For example, for the patient a44601n (Stable), the oximeter consumption during monitoring was 1,922 J (Figure 4.11) in Experiment II, it decreased to about 1,600 J, an economy of around 17% in the current experiment.

The oximeter is the primary concern about energy requirements in our solution. It represents between 98.2% to 99.99% of the energy requirement in our simulations. For simplification, we collected all vital signs using the same sampling rate in our experiments. However, it can be desirable to have specific sampling rates for each health parameter according to the clinical case as proposed by (IDA et al., 2020). To add this feature to our algorithm, the sensor read intervals must be shorter than the period length during which the combined score is calculated in order to adjust the frequencies for the next period. For the calculation of the combined score to process alarms, the algorithm already considers the last valid read value from sensors when the calculation procedure is executed. Finally, to schedule the reading tasks, it is necessary to modify the configuration structures to replicate the frequency of combined score intervals for each sensor and use them accordingly. However, without changing the logic of our algorithm.

Moreover, IoT-based remote monitoring of patients is aimed at less severe clinical cases when data updates of 5 to 10 seconds seem reasonable. In our experiments, we utilized higher frequencies to compare to the available benchmark, such as the multiparameter monitors found in ICUs.

Comparing the WKit total energy requirements for running the ULPR and the BLE versions of our embedded system to the baseline system, energy savings are very similar, from 40.5% to 81.5% (ULPR) and 40.7% to 81.7% (BLE).

Moreover, comparing the ULPR to the BLE version, the latter was more economical (Figure 4.17), with differences of about 9.7% in higher energy consumption records up to 10.6% in lower ones. We utilized the standard configuration from code examples of the manufacturer and customized only the payload of the messages in the ULPR implementation. Improvements in the ULPR version may be implemented in future work. For example, turning off the radio after successful transmissions. The wearable, to receive data, would utilize the ACK packets to notify when the base station intends to send information. Then, the wearable lets the receiver on only in this case. This procedure would turn the ULPR more economical.

Observing Table 4.12 with the measures from both versions, the additional consumption for the self-adaptive and alarm routines is very low com-

pared to other tasks. They presented the lowest individual costs among all tasks. Nonetheless, the WKit total energy needed to process data was higher than to transmit them, as shown in Figure 4.19. However, compared to the savings promoted by the inclusion of the procedures in the reduction of readings and transmissions more than justify its energy cost.

There are several limitations to our method of acquiring energy measures. We utilized manual demarcation of the interested areas in PPK-II charts. This manual process may generate small differences in actual energy costs that, multiplied by thousands of operations, may result in significant errors. Moreover, our measures are specific to our hardware and software designs being challenging to generalize to other hardware platforms. Moreover, a deeper study of Nordic SDK and microcontroller features would be necessary to customize the radios and protocols. Probably, it would be possible to turn off several features not used by our solution that decrease the general energy requirement in both implementations.

Moreover, the presented WKit total energy requirements do not represent the actual energy costs in real scenarios. In our simulator, we utilized the data for five vital signs necessary to calculate the NEWS-2 combined scores. However, our prototype generated information for three of them (HR, SO, BTEMP). Nonetheless, our primary objective was to demonstrate the potential energy savings of our approach and the feasibility of embedding it in a wearable device. Most studies only present theoretical models and utilize hypothetical energy consumption information (ELGHERS; MAKHOUL; LAIYMANI, 2014) (HABIB et al., 2016) (HARB et al., 2021) (IDA et al., 2020). Therefore, our study gives a valuable contribution by providing actual energy measures and a feasible embedded solution in a hardware prototype.

4.5.5

Experiment III - Conclusions

Firstly, this experiment demonstrated the effects of each of our proposed principles on the adaptive sampling behavior and how it affects the number of operations and, consequently, the energy requirements. Specific sensor technologies to capture data may also influence the total energy consumption, such as the long time for the oximeter read the first valid value and the behavior for subsequent reads. Although, for the use cases aimed at IoT-based patient monitoring applications, this would not be a concert because the intervals between readouts will be longer than 5s in most cases.

Experiment III confirmed the results achieved in Experiment II of the potential energy savings of our self-adaptive procedure. In addition, it analyzed

several energy aspects of our proposed self-adaptive approach in a simulated environment using actual energy parameters of all WKit-performed tasks collected from our wearable prototype. Moreover, the Experiment demonstrated the effects of the principles that govern the algorithm on the monitoring sampling.

Furthermore, our tests compared energy consumption with ULPR and BLE versions. They represent two scenarios of patient monitoring applications within clinical settings and at home. The BLE version was more energy efficient than our implementation for the ULPR one. However, both present similar gains compared to the baseline system.

Future evolution of our proposal would include different frequency parameters for each sensor instead of using the same frequencies for all sensors, given the huge difference in energy consumption captured from the oximeter. It also may spur the search for more economical sensor techniques to acquire arterial blood saturation.

Finally, our algorithm added a minimal load in the embedded system and brought a massive benefit in energy savings, extending battery lifetime.

4.6

General Discussion

In this section, we discuss topics related to all experiments and review the relevant findings of our study.

The research design to verify experimental monitoring in actual patient monitoring environments is quite complex. For example, running the experimental and ground-truth systems in parallel would be necessary to verify actual alarm accuracy, as observing the alarms and following-up outcomes in patient cases for a prolonged time and classifying them as true or false alarms. However, there is the inconvenience for patients carrying two devices and for health teams in controlling another tool and following the research procedures without compromising patients' treatment. Moreover, it is very difficult to double-blind participants to avoid bias (BHATIA; MADDOX, 2021). Therefore, alternative procedures should be applied to develop new technologies, such as the ones with simulations applied in our study.

Moreover, several applications of the IoT-PMAs are intended for non-critical patients outside the ICUs, such as for long-term conditions like chronic diseases. We hypothesize that these patients are most of the time in less severe clinical conditions. Figure 4.1 shows that even for patients in ICUs, most of the time, vital signs present normal or low combined scores. Then, using the combined scores to modulate the functioning of the wearable sensor kit can

reduce the sampling rate, data transmission, and energy consumption, in most cases, increasing the efficiency and enhancing the use of resources, but limited by the operational capacity to provide timely information.

Data reductions in WSNs can be categorized as aggregation, adaptive sampling, compressing, and network coding (NASHIRUDDIN; RAKHMAWATI, 2022). Our proposal emphasizes the former two categories. Adaptive sampling considers that sampled data could be highly redundant or negligible for monitoring. Additionally, combined early-warning scores promote aggregation, combining the assessment of individual sensor values and classifying ranges of nominal values into scores.

Moreover, the four proposed principles (Table 3.1) rule the adaptive sampling. The principles are based on real-time risk assessments according to the early-warning scoring system directly using the combined scores and indirectly inferring the similarity of monitored data through individual scores occurrences. Using our approach, data reductions achieved by previous methods ((ELGHERS; MAKHOUL; LAIYMANI, 2014), (HABIB et al., 2016), (HARB et al., 2021)) were surpassed, as shown in Experiments I and II.

Fixed frequency monitoring systems may produce highly redundant information, mainly when the monitored object is stable. Observing other application domains, Shu et al. (SHU et al., 2017) mentioned that in water quality monitoring, if the monitored properties remain stable within a period, it is likely that no significant changes are taking place. Therefore, a lower sampling rate to monitor the water properties could be applied. This idea was utilized in previous studies (ELGHERS; MAKHOUL; LAIYMANI, 2014) (HABIB et al., 2016) (HARB et al., 2021) with physiological data and incorporated into our proposal. The similarity of sets of individual early-warning scores was evaluated to produce a dissimilarity index in our proposal and acts as a second pointer to regulate sampling frequencies after the risk assessment. Differently from previous studies that utilized raw data variance (ELGHERS; MAKHOUL; LAIYMANI, 2014) (HABIB et al., 2016) and a stability index based on intersections of individual scores distribution (HARB et al., 2021) not accounting for the severity of the differences. In the contrast, our weighted index considered the amount and the magnitude of the differences in the classified data. Furthermore, in our solution, variance in raw data within the range of values of a given score does not increase the frequency of monitoring.

Additionally, following the analogy with the watering monitoring applications where the goal is to protect potable water against pollution. Then, if the water is pure and potable, it may deserve closer attention than sewer wa-

ter pipelines that have not received any treatment (SHU et al., 2017). Thus, monitoring frequencies are adjusted according to risk conditions. In patient monitoring, an individual in a healthy state does not need to be monitored as frequently as a patient with abnormal vital signs. In the water monitoring application, the risk is associated with a high loss of potable water, while for patient monitoring applications, the risk is related to worsening health conditions.

Moreover, the combined scores have a very low computational complexity facilitating the implementation in embedded solutions closer to where data are generated. Thus, enhancing data reduction benefits the whole computational and data communication infrastructures (LEQUEPEYS et al., 2021). It is reported that data reductions, in particular, adaptive sampling algorithms, can save up to 79.33% of energy (NASHIRUDDIN; RAKHMAWATI, 2022). Experiments III and IV demonstrated similar or even better results achieving up to 81.7% savings only in the device, not considering the other components of the system, such as gateways, servers, and storage devices. Additionally, our study provided more robust evidence of energy savings based on actual hardware measurements acquired in our wearable prototype, different from other studies that utilized hypothetical models (ELGHERS; MAKHOUL; LAIYMANI, 2014) (HABIB et al., 2016) (HARB et al., 2021).

In addition, combined early-warning scores could cope with alarms and replace threshold alarms with a more comprehensive risk assessment. They aggregate risk information and classify monitored data into different category levels. Then, changes in these categories may carry on more qualified alerts. So, alarm events were assessed when combined scores increased, mirroring worsening levels in the patient's clinical conditions. Moreover, alarm delays and avoiding repeating the same alarm in a time window can drastically reduce the excessive number of alarms with a small loss, as demonstrated by Experiment II 4.3.

Our method takes two main principles to rule our algorithm. The first is the patient's health risk assessed by the combined scores. The second is based on the similarity of monitored data. Additionally, two secondary principles were developed to increase the opportunities for enhancing system efficiency. The first one aims to change frequencies more quickly when risk increases with the same dissimilarity of data in different risk levels. The second one aims at realizing changes quicker in higher-risk conditions. Therefore, the system can use lower frequencies as soon as possible.

In conclusion, the presented experiments applied our approach and demonstrated the potential benefits of data and alarm reductions and energy

savings in IoT-PMAs, achieving the main objective of this study to **minimize(DG, AL, EC)** with a manageable loss in MINT, ALACC, and ALMDR according to algorithm's configuration.

5

Final remarks and Future Work

Patient monitoring applications are complex systems that involve different technologies for acquiring, transmitting, processing, alarming, protecting, and storing patient and treatment information. IoT brings many benefits to monitoring patients remotely but also more challenges to these applications, such as efficient use of resources, massive data generation (MOURA et al., 2020), increased number of alarms (NGUYEN et al., 2019), and energy autonomy of sensor devices (RAULT; BOUABDALLAH; CHALLAL, 2014). The approach presented in this thesis addresses these issues.

The first contribution of our research is the presentation of a solution abstraction that utilizes four simple principles implemented with the aid of an early-warning scoring system vastly utilized in infirmaries to tackle those issues. The low complexity of the approach using individual and combined scores allows embedding a self-adaptive algorithm ruled by those principles in wearable sensor devices, which promotes overall efficiency in the monitoring application infrastructure by reducing data transmissions and energy requirements.

As a second contribution, we designed and developed a hardware prototype with our solution embedded. Moreover, we collected energy consumption information from this prototype and projected energy requirements using different use cases with two radio protocols that could be utilized in infirmaries or outside clinical settings. Moreover, we utilized actual patient data from public datasets to simulate the monitoring of patients in several clinical conditions. From these experiments, we demonstrated that our algorithm added a tiny energy requirement compared to the achieved data savings that reached almost 82% compared to a standard monitoring frequency without our algorithm. Therefore, the implementation of our wearable sensor prototype provided a two-fold benefit for our research. Firstly, by demonstrating the feasibility of embedding our solution in an IoT wearable device. Secondly, by serving our simulations with actual energy measures from the IoT microcontroller and sensors and projecting more reliable energy requirement data.

As a third contribution, we implemented several approaches to reduce less significant alarms with individual and combined scores, in addition to two strategies that can avoid very short transitions in monitored data and repeat the same alarm in a configurable time window. In our experiments, these strategies reduced alarms by up to 78%, using combined early-warning scores.

However, the missed alarm detection rate in one patient reached almost 28% in a 2-minute in the worst case but could be reduced to 6.5% with a configuration for a smaller overall reduction of alarms by 68%. Nonetheless, increasing the time window to ten minutes, the maximum missed alarm detection rate in any patient dropped to 2.8%.

As a fourth contribution, we proposed three metrics to assess the monitoring fidelity loss. Firstly, a general metric to analyze the recorded time during monitoring in each combined score and compare the distributions between our solution and the baseline, observing the overall mismatched error and the error in each monitored patient. Secondly, a metric for checking the missed alarm detection rate, and finally, a metric for verifying the accuracy of alarm types. An advantage of our method of alarming over the thresholds is that alarms can be graded according to the clinical status given by the scoring system.

Another contribution of our work was the definition of a conceptual architecture describing an IoT-based patient monitoring end-to-end solution, including the interfaces and scenarios to connect our proposed smart wearable device with our self-adaptive algorithm.

Finally, in order to develop this research, a comprehensive systematic literature review was conducted covering recent studies regarding real-time data analysis performed in real-time in IoT-based health solutions. This review gives a broad overview of utilized sensors, datasets, properties, and algorithms to detect anomalies, perform predictions, and diagnose in real time.

Our approach demonstrated another feasible strategy to apply EWSS in health systems. It is an abstraction based on analyzing risk and data similarity through a scoring system to enhance the use of computational, networking and energy resources in IoT devices for monitoring patients. An EWSS is a suitable way to implement this abstraction because it supports aggregated weighted classification mechanisms using combined scores to classify risk and aggregates raw data to analyze the similarity of categories of health markers, such as vital signs using individual scores. Early-warning scoring systems have been scrutinized in the patient monitoring domain and evolved quickly in recent years.

The findings of our study have the potential to mitigate the problems of leveraging IoT-based solutions to monitor patients in various scenarios, from managing pandemics to providing ongoing support for chronic illnesses and post-hospital care, to name a few. By harnessing the power of these technologies, healthcare services can be significantly enhanced and optimized to meet the evolving needs of patients and providers alike.

In conclusion, a solution is considered efficient when it can surpass

existing technologies. Despite the limitations, our proposal can enhance current technology applied in patient monitoring routines supported by inflexible monitor devices, threshold-based alarms, and manual procedures to assess vital signs and patient health status.

5.1 Future Work

This work focused on data and alarm reductions that lead to energy efficiency in IoT-PMAs. Nonetheless, other approaches could be applied to promote energy savings, such as compression, radio optimization, and energy harvesting. These approaches are orthogonal to our proposal. They can be implemented to extend it.

In addition, our study can be extended and generalized. The list below presents topics for future investigation:

1. Explore the use of composite scores in other domains.

Risk assessments utilizing scoring systems are performed in other domains, such as agriculture and industrial applications. Exploring the use of our proposal in different fields can uncover benefits and restrictions for those systems and our approach.

2. Utilize our principles with other early-warning scoring systems.

As shown in Section 2.3, several EWSS could be embedded in a sensor device. As demonstrated by our study, the footprint of our implementation is tiny, which would allow the addition of other EWSS scales in the embedded system. Then, our proposal could be used in other clinical domains, providing more flexibility to investigate its use in real scenarios.

3. Develop a wearable device with other vital signs.

Our wearable design can be extended to include other vital sign sensors, such as for inferring respiratory rate and systolic blood pressure. New experimental approaches can be used to acquire these health parameters.

4. Validate sensor measurements in our wearable.

Measures of vital signs acquired in our prototype can be validated by comparing them against a multiparametric hospital monitor utilized in ICUs. Improvements in ergonomics and other sensors and drivers can be proposed to enhance the fidelity of measures.

5. Explore the customization of the ULPR communication aimed at reducing energy consumption.

Our experiments did not focus on achieving energy gains using the proprietary radio (ULPR). However, there are some opportunities to develop a communication protocol that is more efficient than what was implemented in our prototype. For instance, turning off the receiver when not in use and adjusting the power for transmissions based on the distance to base stations. A deeper understanding of radio capabilities and specific communication needs would enhance energy savings in transmissions.

6. Utilize different transmission protocols, such as ZigBee.

Our prototype platform supports natively other communication protocols that could be explored and investigated, such as ZigBee.

7. Use of adaptive delays and adaptive redundant lockout windows.

Use of variable delays and window sizes according to combined scores. For example, the higher the combined score, the shorter the delay and window size, and vice-versa. This approach could better regulate alarm missed detection rate and accuracy, mainly when patients are in more severe health conditions.

8. Utilize distinct frequency ranges for each sensor.

To simplify and demonstrate the main principles in our approach, we did not explore the use of different sampling rates for each vital signs sensor. However, other criteria to modulate the frequencies could be added, such as the sensor's energy consumption and available energy. Thus, configuring different data acquisition frequencies for each sensor.

9. Configuration management for IoT-based patient monitoring applications.

Managing and administrating the configuration of IoT devices is challenging. Issues related to consistency, compatibility, scalability, availability, latency, and security must be addressed in health applications. Privacy and accountability are essential requirements. Our proposal is very flexible, with many parameters that can be adjusted for each monitored patient. Providing a reliable configuration management model is still a challenge to overcome.

This work also presents limitations regarding using patients' public datasets instead of actual data from monitored patients. Additionally, our studies did not address other challenges of IoT applications, such as communication

reliability, privacy and security, and validity of measurements. Each one has a comprehensive and vivid field of study that can be combined or extended to encompass our proposal.

Furthermore, health assessments need the validation of clinical trials, and a multidisciplinary research team would be necessary to apply our proposed approach in real scenarios. EWSS predictive models are still under scrutiny. EWSS based on vital signs should probably be combined with other health records to provide safer and more assertive monitoring for less severe clinical cases. In addition, health infrastructure and teams have been overloaded with the high demand for health services making it more challenging to validate new healthcare tools and strategies. However, our proposal provides a new approach and tools to facilitate the investigation of EWSS in clinical cases.

Moreover, our proposal, in other domains, may not reflect the benefits achieved in our study because patient's IoT monitoring is based on wireless body networks with one or two hops which facilitates the composition of combined scores. In contrast, applications in other domains may utilize wireless sensor networks with several hops, such as environmental control. On the other hand, the benefits of data and alarm reductions could be applied to several IoT applications based on wired networks.

However, our work opens up the opportunity to explore several issues and challenges in IoT-based monitoring applications using early-warning scoring systems.

6

Bibliography

FILIFE, L. et al. Wireless body area networks for healthcare applications: Protocol stack review. **International Journal of Distributed Sensor Networks**, SAGE Publications, v. 2015, p. 1–23, 2015. Disponível em: <<https://doi.org/10.1155/2015/213705>>.

RAULT, T.; BOUABDALLAH, A.; CHALLAL, Y. Energy efficiency in wireless sensor networks: A top-down survey. **Computer Networks**, v. 67, p. 104–122, 2014. ISSN 1389-1286. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1389128614001418>>.

LAIYMANI, D.; MAKHOUL, A. Adaptive data collection approach for periodic sensor networks. In: **2013 9th International Wireless Communications and Mobile Computing Conference**. [S.l.: s.n.], 2013. p. 1448–1453.

PAGANELLI, A. I. et al. A conceptual IoT-based early-warning architecture for remote monitoring of COVID-19 patients in wards and at home. **Internet of Things**, Elsevier BV, v. 18, p. 100399, maio 2022. Disponível em: <<https://doi.org/10.1016/j.iot.2021.100399>>.

ATZORI, L.; IERA, A.; MORABITO, G. The internet of things: A survey. **Computer Networks**, v. 54, n. 15, p. 2787 – 2805, 2010. ISSN 1389-1286. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1389128610001568>>.

PERERA, C. et al. Sensing as a service model for smart cities supported by internet of things. **Transactions on Emerging Telecommunications Technologies**, Wiley, v. 25, n. 1, p. 81–93, Sep 2013.

PERERA, C. et al. Context aware computing for the internet of things: A survey. **IEEE Communications Surveys & Tutorials**, Institute of Electrical and Electronics Engineers (IEEE), v. 16, n. 1, p. 414–454, 2014. Disponível em: <<https://doi.org/10.1109/surv.2013.042313.00197>>.

GUPTA, M. et al. Design and evaluation of an adaptive sampling strategy for a wireless air pollution sensor network. In: **2011 IEEE 36th Conference on Local Computer Networks**. [S.l.: s.n.], 2011. p. 1003–1010.

SHU, T. et al. An energy efficient adaptive sampling algorithm in a sensor network for automated water quality monitoring. **Sensors**, MDPI AG, v. 17, n. 11, p. 2551, nov. 2017. Disponível em: <<https://doi.org/10.3390/s17112551>>.

AKKAŞ, M.; SOKULLU, R.; ÇETIN, H. E. Healthcare and patient monitoring using IoT. **Internet of Things**, Elsevier BV, v. 11, p. 100173, set. 2020. Disponível em: <<https://doi.org/10.1016/j.iot.2020.100173>>.

KHAN, R. A.; PATHAN, A.-S. K. The state-of-the-art wireless body area sensor networks: A survey. **International Journal of Distributed Sensor Networks**, SAGE Publications, v. 14, n. 4, p. 155014771876899, abr. 2018. Disponível em: <<https://doi.org/10.1177/1550147718768994>>.

AL-FUQAHA, A. et al. Internet of things: A survey on enabling technologies, protocols and applications. **IEEE Communications Surveys & Tutorials**, v. 17, p. Fourthquarter 2015, 11 2015.

LEQUEPEYS, J.-R. et al. Overcoming the data deluge challenges with greener electronics. In: **ESSCIRC 2021 - IEEE 47th European Solid State Circuits Conference (ESSCIRC)**. IEEE, 2021. p. 7–14. Disponível em: <<https://doi.org/10.1109/esscirc53450.2021.9567836>>.

NASHIRUDDIN, A. H. N.; RAKHMAWATI, L. Literature review: Energy efficiency mechanisms using data reduction in wireless sensor networks (WSNs) applications. **INAJEEE Indonesian Journal of Electrical and Eletronic Engineering**, Universitas Negeri Surabaya, v. 5, n. 1, p. 5–13, mar. 2022.

MOURA, H. et al. Fog computing in health: A systematic literature review. **Health and Technology**, v. 10, 05 2020.

MOODY, G.; MARK, R. A database to support development and evaluation of intelligent intensive care monitoring. In: **Computers in Cardiology 1996**. [S.l.: s.n.], 1996. p. 657–660.

SAEED, M. et al. MIMIC II: a massive temporal ICU patient database to support research in intelligent patient monitoring. **Computers in Cardiology**, v. 29, p. 641–644, 2002.

LIU, D.; GÖRGES, M.; JENKINS, S. A. University of queensland vital signs dataset. **Anesthesia & Analgesia**, Ovid Technologies (Wolters Kluwer Health), v. 114, n. 3, p. 584–589, mar. 2012. Disponível em: <<https://doi.org/10.1213/ane.0b013e318241f7c0>>.

FERNANDES, C. O. et al. Artificial intelligence technologies for coping with alarm fatigue in hospital environments because of sensory overload: Algorithm development and validation. **Journal of Medical Internet Research**, v. 21, n. 11, p. e15406, Nov 2019. ISSN 1438-8871. Disponível em: <<http://www.jmir.org/2019/11/e15406/>>.

NGUYEN, J. et al. Combating alarm fatigue: The quest for more accurate and safer clinical monitoring equipment. In: STAWICKI, S. P.; FIRSTENBERG, M. S. (Ed.). **Vignettes in Patient Safety**. Rijeka: IntechOpen, 2019. cap. 6. Disponível em: <<https://doi.org/10.5772/intechopen.84783>>.

INDU, S. D. et al. Wireless sensor networks: Issues & challenges. **International Journal of Computer Science and Mobile Computing (IJCSMC)**, v. 3, p. 681–85, 2014.

MISHRA, S. S.; RASOOL, A. lot health care monitoring and tracking: A survey. In: **2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)**. [S.l.: s.n.], 2019. p. 1052–1057.

KIMURA, N.; LATIFI, S. A survey on data compression in wireless sensor networks. In: **International Conference on Information Technology: Coding and Computing (ITCC'05) - Volume II**. IEEE, 2005. Disponível em: <<https://doi.org/10.1109/itcc.2005.43>>.

SANISLAV, T. et al. Energy harvesting techniques for internet of things (IoT). **IEEE Access**, Institute of Electrical and Electronics Engineers (IEEE), v. 9, p. 39530–39549, 2021. Disponível em: <<https://doi.org/10.1109/access.2021.3064066>>.

RAGHUNATHAN, V.; GANERIWAL, S.; SRIVASTAVA, M. Emerging techniques for long lived wireless sensor networks. **IEEE Communications Magazine**, v. 44, n. 4, p. 108–114, 2006.

GERRY, S. et al. Early warning scores for detecting deterioration in adult hospital patients: a systematic review protocol. **BMJ Open**, BMJ, v. 7, n. 12, p. e019268, dez. 2017. Disponível em: <<https://doi.org/10.1136/bmjopen-2017-019268>>.

PHYSICIANS LONDON, U. Royal College of. **Royal College of Physicians. National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity in the NHS**. [S.l.], 2017. Accessed: Mar. 02, 2020. Disponível em: <<https://www.rcplondon.ac.uk/projects/outputs/national-early-warning-score-news-2>>.

WONG, D. et al. SEND: a system for electronic notification and documentation of vital sign observations. **BMC Medical Informatics and Decision Making**, Springer Science and Business Media LLC, v. 15, n. 1, ago. 2015. Disponível em: <<https://doi.org/10.1186/s12911-015-0186-y>>.

RAGHUNATHAN, V. et al. Energy-aware wireless microsensor networks. **IEEE Signal Processing Magazine**, v. 19, n. 2, p. 40–50, 2002.

PAGANELLI, A. I. et al. Real-time data analysis in health monitoring systems: A comprehensive systematic literature review. **Journal of Biomedical Informatics**, Elsevier BV, v. 127, p. 104009, mar. 2022.

HARB, H. et al. A sensor-based data analytics for patient monitoring in connected healthcare applications. **IEEE Sensors Journal**, v. 21, n. 2, p. 974–984, 2021.

GARDNER, R. M.; SHABOT, M. M. Patient-monitoring systems. In: _____. **Medical Informatics: Computer Applications in Health Care and Biomedicine**. New York, NY: Springer New York, 2001. p. 443–484. ISBN 978-0-387-21721-5. Disponível em: <https://doi.org/10.1007/978-0-387-21721-5_13>.

PRUTSACHAINIMIT, K.; JANTHONG, T.; TUANROPI, W. A development of inpatient monitoring system: A case study of patong hospital, phuket, thailand. In: **2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)**. [S.l.: s.n.], 2020. p. 25–28.

LEENEN, J. P. L. et al. Current evidence for continuous vital signs monitoring by wearable wireless devices in hospitalized adults: Systematic review. **Journal of Medical Internet Research**, v. 22, n. 6, p. e18636, Jun 2020. ISSN 1438-8871. Disponível em: <<http://www.jmir.org/2020/6/e18636/>>.

KYRIACOS, U.; JELSMA, J.; JORDAN, S. Monitoring vital signs using early warning scoring systems: a review of the literature. **Journal of Nursing Management**, Wiley Online Library, v. 19, n. 3, p. 311–330, 2011.

ELANGO, K.; MUNIANDI, K. A low-cost wearable remote healthcare monitoring system. In: _____. **Role of Edge Analytics in Sustainable Smart City Development**. [S.l.]: John Wiley & Sons, Ltd, 2020. cap. 11, p. 219–242. ISBN 9781119681328.

LU, L. et al. Wearable health devices in health care: Narrative systematic review. **JMIR Mhealth Uhealth**, v. 8, n. 11, p. e18907, Nov 2020. ISSN 2291-5222. Disponível em: <<http://mhealth.jmir.org/2020/11/e18907/>>.

PHILIP, N. Y. et al. Internet of things for in-home health monitoring systems: Current advances, challenges and future directions. **IEEE Journal on Selected Areas in Communications**, v. 39, n. 2, p. 300–310, 2021.

DIAS, D.; CUNHA, J. P. S. Wearable health devices-vital sign monitoring, systems and technologies. **Sensors (Basel, Switzerland)**, MDPI, v. 18, n. 8, p. 2414, Jul 2018. ISSN 1424-8220. PMC6111409[pmcid]. Disponível em: <<https://doi.org/10.3390/s18082414>>.

QI, J. et al. Advanced internet of things for personalised healthcare systems: A survey. **Pervasive and Mobile Computing**, v. 41, p. 132 – 149, 2017. ISSN 1574-1192. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1574119217303255>>.

THILAKARATHNE, N.; KAGITA, M. K.; GADEKALLU, D. T. R. The role of the internet of things in health care: A systematic and comprehensive study. **International Journal of Engineering and Management Research**, v. 10, p. 145–159, 2020.

BANAEE, H.; AHMED, M. U.; LOUTFI, A. Data mining for wearable sensors in health monitoring systems: a review of recent trends and challenges. **Sensors (Basel, Switzerland)**, Molecular Diversity Preservation International (MDPI), v. 13, n. 12, p. 17472–17500, Dec 2013. ISSN 1424-8220. 24351646[pmid]. Disponível em: <<https://doi.org/10.3390/s131217472>>.

SILVA, C.; JUNIOR, G. de A. Fog computing in healthcare: A review. In: **2018 IEEE Symposium on Computers and Communications (ISCC)**. Los Alamitos, CA, USA: IEEE Computer Society, 2018. p. 01126–01131. ISSN 1530-1346. Disponível em: <<https://doi.ieeecomputersociety.org/10.1109/ISCC.2018.8538671>>.

MUTLAG, A. A. et al. Enabling technologies for fog computing in healthcare iot systems. **Future Generation Computer Systems**, v. 90, p. 62 – 78, 2019. ISSN 0167-739X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167739X18314006>>.

HARTMANN, M.; HASHMI, U. S.; IMRAN, A. Edge computing in smart health care systems: Review, challenges, and research directions. **Transactions on Emerging Telecommunications Technologies**, n/a, n. n/a, p. e3710, 2019. E3710 ett.3710. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/ett.3710>>.

MOHAMMADZADEH, N. et al. The application of wearable smart sensors for monitoring the vital signs of patients in epidemics: a systematic literature review. **Journal of ambient intelligence and humanized computing**, Springer Berlin Heidelberg, p. 1–15, Nov 2020. ISSN 1868-5137. 33224305[pmid]. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/33224305>>.

KRISTOFFERSSON, A.; LINDÉN, M. Wearable sensors for monitoring and preventing noncommunicable diseases: A systematic review. **Information (Switzerland)**, v. 11, p. 521, 11 2020.

HASAN, M. N. U.; NEGULESCU, I. Wearable technology for baby monitoring: A review. v. 6, p. 112–120, 07 2020.

G.S, K.; P.B, P. A review on human healthcare internet of things: A technical perspective. **SN Computer Science**, v. 1, 06 2020.

USAK, M. et al. Health care service delivery based on the internet of things: A systematic and comprehensive study. **International Journal of Communication Systems**, v. 33, 09 2019.

GÓMEZ, J.; OVIEDO, B.; ZHUMA, E. Patient monitoring system based on internet of things. **Procedia Computer Science**, Elsevier BV, v. 83, p. 90–97, 2016. Disponível em: <<https://doi.org/10.1016/j.procs.2016.04.103>>.

Zainol, M. F. et al. A new iot patient monitoring system for hemodialysis treatment. In: **2019 IEEE Conference on Open Systems (ICOS)**. [S.l.: s.n.], 2019. p. 46–50.

Nubenthan, S.; Ravichelvan, K. A wireless continuous patient monitoring system for dengue; wi-mon. In: **2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)**. [S.l.: s.n.], 2017. p. 2201–2205.

KRISTOFFERSSON, A.; LINDÉN, M. A systematic review on the use of wearable body sensors for health monitoring: A qualitative synthesis. **Sensors (Basel, Switzerland)**, MDPI, v. 20, n. 5, p. 1502, Mar 2020. ISSN 1424-8220. 32182907[pmid]. Disponível em: <<https://doi.org/10.3390/s20051502>>.

RGHIOUI, A. et al. A smart architecture for diabetic patient monitoring using machine learning algorithms. **Healthcare (Basel, Switzerland)**, MDPI, v. 8, n. 3, p. 348, Sep 2020. ISSN 2227-9032. 32961757[pmid].

BAIG, M. M.; GHOLAMHOSSEINI, H. Smart health monitoring systems: An overview of design and modeling. **Journal of Medical Systems**, v. 37, n. 2, p. 9898, Jan 2013. ISSN 1573-689X. Disponível em: <<https://doi.org/10.1007/s10916-012-9898-z>>.

ESCAMILLA-AMBROSIO, P. J. et al. Distributing computing in the internet of things: Cloud, fog and edge computing overview. In: _____. [S.l.: s.n.], 2018. v. 731, p. 87–115. ISBN 978-3-319-64062-4.

YOUSEFPOUR, A. et al. All one needs to know about fog computing and related edge computing paradigms: A complete survey. **Journal of Systems Architecture**, v. 98, p. 289 – 330, 2019. ISSN 1383-7621. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1383762118306349>>.

BARIK, R. K. et al. Leveraging machine learning in mist computing telemonitoring system for diabetes prediction. In: **Advances in Data and Information Sciences**. [S.l.]: Springer, 2018. p. 95–104.

ANI, R. et al. Lot based patient monitoring and diagnostic prediction tool using ensemble classifier. In: **2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)**. [S.l.: s.n.], 2017. p. 1588–1593.

FOUAD, H. et al. Analyzing patient health information based on iot sensor with ai for improving patient assistance in the future direction. **Measurement**, v. 159, p. 107757, 2020. ISSN 0263-2241. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0263224120302955>>.

Darcini S., V. P.; Isravel, D. P.; Silas, S. A comprehensive review on the emerging iot-cloud based technologies for smart healthcare. In: **2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)**. [S.l.: s.n.], 2020. p. 606–611.

VISHNUPRASAD, K. et al. Towards building low cost multi-parameter patient monitors. In: **Conference: International Conference on Communication and Computing (ICC- 2014)**. [S.l.: s.n.], 2014.

EDDAHCHOURI, Y. et al. Effect of continuous wireless vital sign monitoring on unplanned ICU admissions and rapid response team calls: a before-and-after study. **British Journal of Anaesthesia**, Elsevier BV, v. 128, n. 5, p. 857–863, maio 2022. Disponível em: <<https://doi.org/10.1016/j.bja.2022.01.036>>.

MICHARD, F.; SESSLER, D. Ward monitoring 3.0. **British Journal of Anaesthesia**, Elsevier BV, v. 121, n. 5, p. 999–1001, nov. 2018. Disponível em: <<https://doi.org/10.1016/j.bja.2018.07.032>>.

YU, H.; LI, N.; ZHAO, N. How far are we from achieving self-powered flexible health monitoring systems: An energy perspective. **Advanced Energy Materials**, v. 11, n. 9, p. 2002646, 2021.

ZHOU, X. et al. Energy efficiency optimization by resource allocation in wireless body area networks. In: **2014 IEEE 79th Vehicular Technology Conference (VTC Spring)**. [S.l.: s.n.], 2014. p. 1–6.

LOUKATOS, D. et al. Revealing characteristic IoT behaviors by performing simple energy measurements via open hardware/software components. In: **Proceedings of 6th International Congress on Information and Communication Tech.** [S.l.]: Springer Singapore, 2021. p. 1045–1053.

HABIB, C. et al. Self-adaptive data collection and fusion for health monitoring based on body sensor networks. **IEEE Transactions on Industrial Informatics**, v. 12, n. 6, p. 2342–2352, 2016.

RAJAGOPALAN, R.; VARSHNEY, P. K. Data-aggregation techniques in sensor networks: A survey. **IEEE Communications Surveys & Tutorials**, v. 8, n. 4, p. 48–63, 2006.

FU, L.-H. et al. Development and validation of early warning score system: A systematic literature review. **Journal of Biomedical Informatics**, Elsevier BV, v. 105, p. 103410, maio 2020. Disponível em: <<https://doi.org/10.1016/j.jbi.2020.103410>>.

WEENK, M. et al. Continuous monitoring of vital signs using wearable devices on the general ward: Pilot study. **JMIR mHealth and uHealth**, JMIR Publications Inc., v. 5, n. 7, p. e91, jul. 2017. Disponível em: <<https://doi.org/10.2196/mhealth.7208>>.

VARSHNEY, U. A framework for supporting emergency messages in wireless patient monitoring. **Decision Support Systems**, Elsevier BV, v. 45, n. 4, p. 981–996, nov. 2008. Disponível em: <<https://doi.org/10.1016/j.dss.2008.03.006>>.

MORGAN, R. J. et al. An early warning scoring system for detecting developing critical illness. In: . [S.l.: s.n.], 1997.

SUBBE, C. Validation of a modified early warning score in medical admissions. **QJM**, Oxford University Press (OUP), v. 94, n. 10, p. 521–526, out. 2001. Disponível em: <<https://doi.org/10.1093/qjmed/94.10.521>>.

GOLDHILL, D. R. et al. A physiologically-based early warning score for ward patients: the association between score and outcome. **Anaesthesia**, Wiley, v. 60, n. 6, p. 547–553, jun. 2005. Disponível em: <<https://doi.org/10.1111/j.1365-2044.2005.04186.x>>.

QUARTERMAN, C. P. J. et al. Use of a patient information system to audit the introduction of modified early warning scoring. **Journal of Evaluation in Clinical Practice**, Wiley, v. 11, n. 2, p. 133–138, abr. 2005. Disponível em: <<https://doi.org/10.1111/j.1365-2753.2005.00513.x>>.

SMITH, M. E. B. et al. Early warning system scores for clinical deterioration in hospitalized patients: A systematic review. **Annals of the American Thoracic Society**, American Thoracic Society, v. 11, n. 9, p. 1454–1465, nov. 2014. Disponível em: <<https://doi.org/10.1513/annalsats.201403-102oc>>.

FANG, A. H. S.; LIM, W. T.; BALAKRISHNAN, T. Early warning score validation methodologies and performance metrics: a systematic review. **BMC Medical Informatics and Decision Making**, Springer Science and Business Media LLC, v. 20, n. 1, jun. 2020. Disponível em: <<https://doi.org/10.1186/s12911-020-01144-8>>.

BADRINATH, K. et al. Comparison of various severity assessment scoring systems in patients with sepsis in a tertiary care teaching hospital. **Indian Journal of Critical Care Medicine**, Jaypee Brothers Medical Publishing, v. 22, n. 12, p. 842–845, 2018. Disponível em: <https://doi.org/10.4103/ijccm.ijccm_322_18>.

CHERANAKHORN, C. et al. Use of modified search out severity (m-sos) score for early sepsis detection in neutropenic patients. In: . [S.l.: s.n.], 2016.

ALBRIGHT, C. M. et al. The sepsis in obstetrics score: a model to identify risk of morbidity from sepsis in pregnancy. **American Journal of Obstetrics and Gynecology**, Elsevier, v. 211, n. 1, p. 39–e1, 2014.

P, M. T. et al. Evaluation of the overall accuracy of the combined early warning scoring systems in the prediction of in-hospital mortality. **Cureus**, Cureus, Inc., abr. 2022. Disponível em: <<https://doi.org/10.7759/cureus.24486>>.

WILLIAMS, B. The national early warning score 2 (NEWS2) in patients with hypercapnic respiratory failure. **Clinical Medicine**, Royal College of Physicians, v. 19, n. 1, p. 94–95, jan. 2019. Disponível em: <<https://doi.org/10.7861/clinmedicine.19-1-94>>.

YE, C. et al. A real-time early warning system for monitoring inpatient mortality risk: Prospective study using electronic medical record data. **Journal of Medical Internet Research**, JMIR Publications, v. 21, n. 7, p. e13719–e13719, Jul 2019. ISSN 1438-8871. 31278734 pmid. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/31278734>>.

NICE. Acutely ill patients in hospital: Recognition of and response to acute illness in adults in hospital [internet]. 2007 July. London: National Institute for Health and Clinical Excellence. (NICE Clinical Guidelines, No. 50.). Disponível em: <<https://www.ncbi.nlm.nih.gov/books/NBK45947/>>.

SAAB, R. et al. Failure to detect ward hypoxaemia and hypotension: contributions of insufficient assessment frequency and patient arousal during nursing assessments. **British Journal of Anaesthesia**, Elsevier BV, v. 127, n. 5, p. 760–768, nov. 2021. Disponível em: <<https://doi.org/10.1016/j.bja.2021.06.014>>.

GAO, H. et al. Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. **Intensive care medicine**, Springer, v. 33, n. 4, p. 667–679, 2007.

SUBBE, C. P.; DULLER, B.; BELLOMO, R. Effect of an automated notification system for deteriorating ward patients on clinical outcomes. **Critical Care**, Springer Science and Business Media LLC, v. 21, n. 1, mar. 2017. Disponível em: <<https://doi.org/10.1186/s13054-017-1635-z>>.

ELGHERS, S.; MAKHOUL, A.; LAIYMANI, D. Local emergency detection approach for saving energy in wireless body sensor networks. In: **2014 IEEE 10th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)**. [S.l.: s.n.], 2014. p. 585–591.

MANN, K. D. et al. Predicting patient deterioration: A review of tools in the digital hospital setting. **Journal of Medical Internet Research**, v. 23, n. 9, p. e28209, Sep 2021. ISSN 1438-8871. Disponível em: <<https://www.jmir.org/2021/9/e28209>>.

GERRY, S. et al. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. **BMJ**, BMJ, p. m1501, maio 2020. Disponível em: <<https://doi.org/10.1136/bmj.m1501>>.

HRAVNAK, M. et al. A call to alarms: Current state and future directions in the battle against alarm fatigue. **Journal of Electrocardiology**, v. 51, n. 6, Supplement, p. S44–S48, 2018. ISSN 0022-0736. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0022073618304722>>.

RUPPEL, H.; FUNK, M.; WHITTEMORE, R. Measurement of physiological monitor alarm accuracy and clinical relevance in intensive care units. **American Journal of Critical Care**, AACN Publishing, v. 27, n. 1, p. 11–21, jan. 2018. Disponível em: <<https://doi.org/10.4037/ajcc2018385>>.

CVACH, M. Monitor alarm fatigue: An integrative review. **Biomedical Instrumentation & Technology**, Association for the Advancement of Medical Instrumentation (AAMI), v. 46, n. 4, p. 268–277, jul. 2012. Disponível em: <<https://doi.org/10.2345/0899-8205-46.4.268>>.

JONES, K. Alarm fatigue a top patient safety hazard. **Canadian Medical Association Journal**, CMA Joule Inc., v. 186, n. 3, p. 178–178, jan. 2014. Disponível em: <<https://doi.org/10.1503/cmaj.109-4696>>.

BONAFIDE, C. P. et al. Association between exposure to nonactionable physiologic monitor alarms and response time in a children's hospital. **Journal of Hospital Medicine**, Wiley, v. 10, n. 6, p. 345–351, abr. 2015. Disponível em: <<https://doi.org/10.1002/jhm.2331>>.

KANE-GILL, S. L. et al. Technologic distractions (part 1). **Critical Care Medicine**, Ovid Technologies (Wolters Kluwer Health), v. 45, n. 9, p. 1481–1488, set. 2017. Disponível em: <<https://doi.org/10.1097/ccm.0000000000002580>>.

CHEN, L. et al. Using supervised machine learning to classify real alerts and artifact in online multisignal vital sign monitoring data. **Critical Care Medicine**, Ovid Technologies (Wolters Kluwer Health), v. 44, n. 7, p. e456–e463, jul. 2016. Disponível em: <<https://doi.org/10.1097/ccm.0000000000001660>>.

SCHMID, F. et al. Reduction of clinically irrelevant alarms in patient monitoring by adaptive time delays. **Journal of Clinical Monitoring and Computing**, Springer Science and Business Media LLC, v. 31, n. 1, p. 213–219, nov. 2015. Disponível em: <<https://doi.org/10.1007/s10877-015-9808-2>>.

PAINE, C. W. et al. Systematic review of physiologic monitor alarm characteristics and pragmatic interventions to reduce alarm frequency. **Journal of Hospital Medicine**, Wiley, v. 11, n. 2, p. 136–144, dez. 2015. Disponível em: <<https://doi.org/10.1002/jhm.2520>>.

IMHOFF, M. et al. Smart alarms from medical devices in the OR and ICU. **Best Practice & Research Clinical Anaesthesiology**, Elsevier BV, v. 23, n. 1, p. 39–50, mar. 2009. Disponível em: <<https://doi.org/10.1016/j.bpa.2008.07.008>>.

WINTERS, B. D. et al. Technological distractions (part 2). **Critical Care Medicine**, Ovid Technologies (Wolters Kluwer Health), v. 46, n. 1, p. 130–137, jan. 2018. Disponível em: <<https://doi.org/10.1097/ccm.0000000000002803>>.

OREIZY, P. et al. An architecture-based approach to self-adaptive software. **IEEE Intelligent Systems and Their Applications**, IEEE, v. 14, n. 3, p. 54–62, 1999.

RAHIMI, M.; SAFABAKHSH, R. Adaptation of sampling in target tracking sensor networks. In: **2010 IEEE International Conference on Wireless Communications, Networking and Information Security**. [S.l.: s.n.], 2010. p. 301–305.

IDA, I. B. et al. Self-adaptative early warning scoring system for smart hospital. In: **Lecture Notes in Computer Science**. Springer International Publishing, 2020. p. 16–27. Disponível em: <https://doi.org/10.1007/978-3-030-51517-1_2>.

GOLDBERGER, A. L. et al. PhysioBank, PhysioToolkit, and PhysioNet. **Circulation**, Ovid Technologies (Wolters Kluwer Health), v. 101, n. 23, jun. 2000. Disponível em: <<https://doi.org/10.1161/01.cir.101.23.e215>>.

ENDLER, M.; SILVA, F. S. e. Past, present and future of the contextnet iomt middleware. **OJIOT**, RonPub, v. 4, n. 1, p. 7–23, 2018. ISSN 2364-7108. Special Issue: Proc. VLloT in conjunction with VLDB.

Al-Masri, E. et al. Investigating messaging protocols for the internet of things (iot). **IEEE Access**, v. 8, p. 94880–94911, 2020.

BARTELS, K. Advances in photoplethysmography: beyond arterial oxygen saturation. **Canadian Journal of Anaesthesia**, v. 62, n. 12, p. 1313–28, dez. 2015.

MACRAE, B. A. et al. Skin temperature measurement using contact thermometry: A systematic review of setup variables and their effects on measured values. **Frontiers in Physiology**, v. 9, p. 29, 2018. ISSN 1664-042X.

Dementyev, A. et al. Power consumption analysis of bluetooth low energy, zigbee and ant sensor nodes in a cyclic sleep scenario. In: **Proc. IWS**. [S.l.: s.n.], 2013. p. 1–4.

PATER, C. M. et al. Time series evaluation of improvement interventions to reduce alarm notifications in a paediatric hospital. **BMJ Quality & Safety**, BMJ, v. 29, n. 9, p. 717–726, jan. 2020.

WEBSTER, C. S.; SCHEEREN, T. W.; WAN, Y. I. Patient monitoring, wearable devices, and the healthcare information ecosystem. **British Journal of Anaesthesia**, v. 128, n. 5, p. 756–758, 2022. ISSN 0007-0912.

MICHARD, F.; SESSLER, D. Ward monitoring 3.0. **British Journal of Anaesthesia**, Elsevier BV, v. 121, n. 5, p. 999–1001, nov. 2018.

PURBAUGH, T. Alarm fatigue. **Dimensions of Critical Care Nursing**, Ovid Technologies (Wolters Kluwer Health), v. 33, n. 1, p. 4–7, 2014. Disponível em: <<https://doi.org/10.1097/dcc.000000000000014>>.

KORNIWICZ, D. M.; CLARK, T.; DAVID, Y. A national online survey on the effectiveness of clinical alarms. **Am J Crit Care**, United States, v. 17, n. 1, p. 36–41, jan. 2008.

KONG, J. et al. Analysis of hematological indexes of COVID-19 patients from fever clinics in suzhou, china. **International Journal of Laboratory Hematology**, Wiley, v. 42, n. 5, jul. 2020. Disponível em: <<https://doi.org/10.1111/ijlh.13290>>.

HASAN, K. et al. A comprehensive review of wireless body area network. **Journal of Network and Computer Applications**, v. 143, p. 178–198, 2019. ISSN 1084-8045.

BHATIA, A.; MADDOX, T. M. Remote patient monitoring in heart failure: Factors for clinical efficacy. **International Journal of Heart Failure**, Korean Society of Heart Failure, v. 3, n. 1, p. 31, 2021. Disponível em: <<https://doi.org/10.36628/ijhf.2020.0023>>.

COUTINHO, J. P. A. **Sistema para Monitoramento Contínuo, Remoto e não-invasivo de pacientes**. Faculdades Catolicas, 2021. Disponível em: <<https://doi.org/10.17771/pucRio.acad.54470>>.

ANDRADE, M. M. de. **Uma Graphical User Interface em Flutter para Monitoramento de Dados Fisiológicos e a Geração de Alertas para Pacientes**. Faculdades Catolicas, 2022. Disponível em: <<https://doi.org/10.17771/pucRio.acad.57943>>.

A

Research Methodology

This appendix describes the methodology applied for the development of this Thesis. Figure A.1 depicts the applied methodology in this research. The next sections detail the procedures in each phase and the pointers to their development along the chapters.

A.1 Research Topic Definition

The objective of this step is to define the research topic to be investigated. At the end of this step, the main topic of the research must be defined.

The report of the activities performed to accomplish the goal of this step can be found in subsection 1.2.

A.2 Related Work Investigation

The objective of this step is to investigate the existing literature regarding the chosen research topic defined in the previous step.

The outcomes of this step are: (1) a summary of underlying technologies and studies related to our research approach; (2) a list of literature gaps, limitations, and inconsistencies in previous work; They can be found in Chapter 2

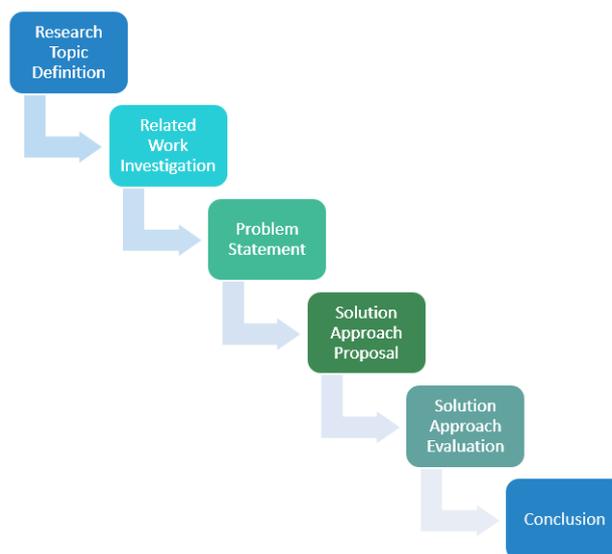


Figure A.1: Research Methodology phases

A.3

Problem Statement Definition

In this step, the scope of the problems addressed by this research is defined. The products are the problem definition, the research goals, the requirements to be met, and the statement of the research questions.

The products can be found in sections 1.3 and 1.4.

A.4

Solution Approach Proposal

In this step, it is defined the theoretical basis and drivers for solving the problems and answers the research questions formulated in the previous step.

A specific experimental methodology is also drawn to handle the independent variables and find the relationships and results as pieces of evidence that the proposal entails in the expected outcomes.

The solution approach is fully described in Chapter 3 while a short introduction is given in section 1.5.

A.5

Solution Approach Evaluation

In this step, the solution is applied to solve the target problems based on the procedures and metrics defined in the experimental methodology. Specific implementation choices and assumptions are also described here.

Experiment results are assessed based on the research questions, and explanations about the reasons for achieving the results are given. A discussion on strengths and weaknesses is drawn regarding those questions.

In addition, findings are analyzed from the perspective of other techniques presented in the literature. Furthermore, inferences for the application of our method in other use cases are made.

Finally, the limitations of our experiments and approach are discussed, and possibilities for future work are indicated.

Chapter 4 presents the topics of this step.

A.6

Research Final Remarks

Finally, the last task is to summarize the found answers to the research questions and point out the main findings and contributions of the research. Future work and limitations of the study are also summarized. These topics are presented in Chapter 5.

B

SH-Sens Wearable Kit development

In our approach, embedding the solution in the wearable device is essential to achieve the highest benefits for the efficient use of resources in IoT-PMAs. As demonstrated in our experiments, the solution has the potential to reduce drastically the readings and transmissions, which leads to extending the battery life of wearable sensors.

Therefore, it is desirable to demonstrate the feasibility of developing a hardware prototype that could embed our solution using the constrained environment of IoT microcontrollers.

This Section describes the development process of our hardware prototype.

B.1

WKit development - Objectives

The main objectives of developing a wearable sensor prototype are to demonstrate the feasibility of embedding our proposed algorithm and to obtain actual parameters to demonstrate the potential benefits of energy savings.

It is not an objective to develop a product or a device to be utilized in the market. The developed device is only a platform to be utilized as a proof of concept for our experiments.

B.2

WKit Design - Version 1

Motivated by the surge of COVID-19 and the deterioration of healthcare services, our group of study joint efforts to develop solutions for a smart wearable device to monitor patients remotely. COVID-19 virus transmission was very high, and the recommendation was to keep patients as isolated as possible and reduce unnecessary interactions with healthcare personnel.

As shown in Figure B.1, the plan was to develop a wearable device loaded with two sensors, a temperature sensor and an oximeter. One objective was to investigate the wearable and gateway communication protocol. Then, a proprietary radio option (ULPR) was chosen because it allowed the customization of this communication. The proprietary radio has to connect to a base station built with the same microcontroller (nRF52832) and an ESP8266 to provide a WIFI connection to the Internet.

The next step, as shown in Figure B.2, was the design of the boards and circuits for the components. At this stage, it was essential to decide which power

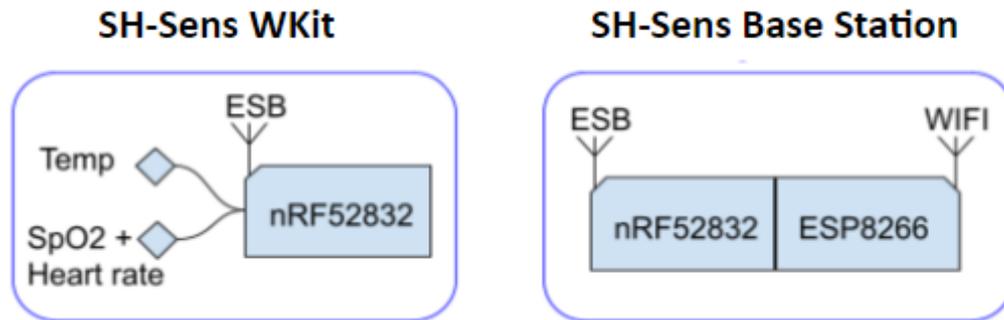


Figure B.1: First studies of SH-Sens Wearable Kit and Base Station with NRF-52832 microcontroller. Aimed at infirmaries and COVID-19 patients.

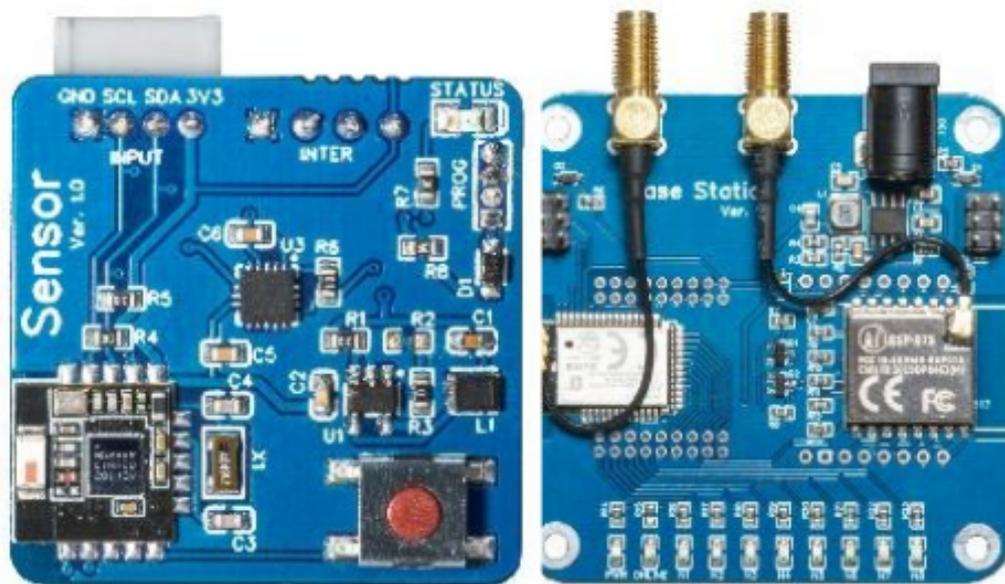


Figure B.2: First studies of board designs. On the left is the WKit. On the right is the base station.

supply type and battery case should be used, the ergonomics, and the shape of the wearable. The final design of the first version is shown in Figure B.3. João Pedro Coutinho developed this design as his final graduation project (COUTINHO, 2021) in collaboration with the Laboratory of Advanced Collaboration (LAC).

At this stage, there were no smart features embedded in the device. The wearable only sends periodic information to the base station that publishes them to be distributed using an MQTT broker.

B.3 WKit Design - Version 2

Based on the experience obtained in the first version of the WKit, we submitted a request to the Instituto Nacional de Ciência e Tecnologia (INCT) of the



Figure B.3: First version of devices.

Future Internet for Smart Cities / FAPESP project ¹ requesting funds to develop a new version of our wearable prototype. The project was approved (#2019/19312-9). With the project funds, a team composed of one electronic engineer and one experienced technician was hired to help our team with the design and production of the WKit device based on our requirements. Remembering that the university laboratories were closed because of the precautionary restrictions of COVID-19.

The second WKit version was developed to support Bluetooth Low Energy communication, in addition to the proprietary radio, test new physiological sensors, and embed our self-adaptive algorithm based on early-warning scores.

The first activity was to study the sensors and microcontroller characteristics. We tested optical and contact temperature sensors. The optical temperature sensor presented more reliable measurements. It was also tested two versions of the oximeter sensor (MAX30102 and MAX30100). The MAX30102 driver presented problems in capturing measurements. However, the MAX30100 produced stable measurements. The MAX30105 was not tested. It would be desirable to be utilized as a pulse oximeter with a watch-like design, but we opted for a finger oximeter.

Regarding the microcontroller, the nRF52832 was utilized in our previous WKit version. The nRF52833 is very similar to nRF52832 but supports BLE 5.1 specification, ZigBee, and IEEE 802.15.4 communication protocols, which are not supported by the nRF52832. Although increasing memory capacity was not a main concern for our implementation with combined scores, nRF52833 supports 128 KB RAM, while nRF52832 supports up to 64 KB RAM. However, it would be possible to explore more sophisticated embedded algorithms, such as Support Vector Machines or Random Tree models trained on the cloud. These experiments were left to future work. Considering the cited advantages, the nRF52833 was chosen for our second version. Figure B.4 shows the tested components.

After choosing the components, the WKit circuit schema was projected to build the board. The WKit circuit schema is described in Figure B.5.

¹<https://interscity.org/>

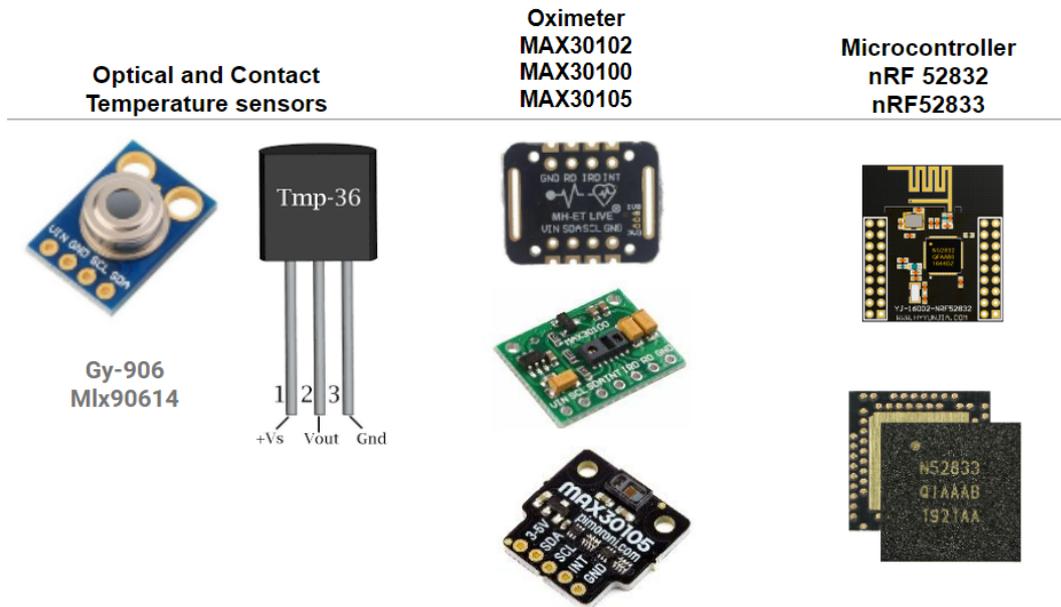


Figure B.4: Sensors and Micro-controllers utilized in our tests for the WKit version 2.

PUC-Rio - Certificação Digital N° 1912730/CA

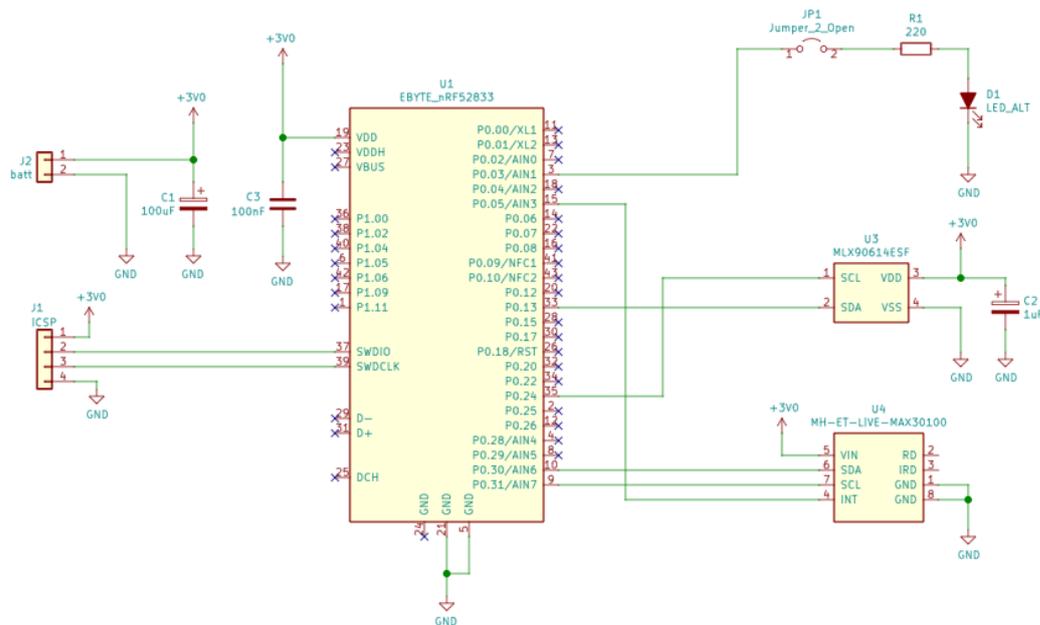


Figure B.5: WKit version 2 - circuit schema.



Figure B.6: WKit version 2 - device design.

In parallel, the study of the device's design was started as shown in Figure B.6. It was decided to use the finger clip design because we chose to supply the device with two 1.5v AA batteries, which are readily available. Initially, it was expected to operate the WKit for at least 10 days uninterruptedly. This was the recommended quarantine period for COVID-19-infected people.

In parallel, the components were mounted in breadboards as shown in Figure B.7. Basic functions provided by Nordic were first utilized to connect the WKit to the base station and adjust the software development environment, regulate radio power for transmissions, and test the pin connections of our mounted components. Then, the control looping process, the communication protocol, the drivers for the sensors to support sleep modes, and finally, each of the tasks described in Subsection 3.5.4 were developed.

In addition, the proprietary communication protocol for the ULPR was migrated to BLE using the serial mode of the driver. To test this implementation, firstly, it was utilized the MIT App Inventor and the BluetoothBLE inventor extension library ² to develop a mobile application. Afterward, a mobile version was implemented in Flutter to connect with the BLE version of WKit, which can be seen in Appendix C. Figure B.8 shows the final version of our WKit connected to the MIT App Inventor mobile application via a BLE connection.

²<https://iot.appinventor.mit.edu/#/bluetoothle/bluetoothleintro>

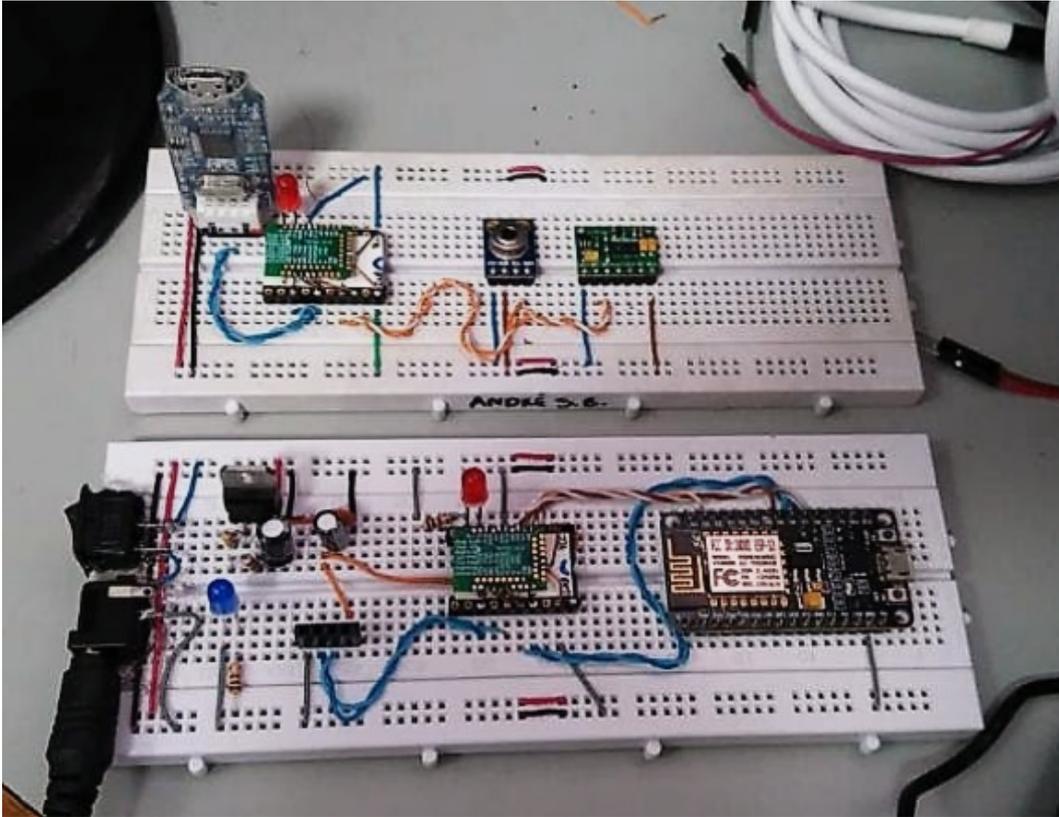


Figure B.7: WKit and base station version 2. Breadboard tests.

B.4 WKit development - Discussion

The development of the WKit was essential to collect energy requirement information from actual hardware components, such as those provided in Experiments II and III. Moreover, it demonstrated the feasibility of embedding our proposed self-adaptive algorithm in a low-cost platform.

The experiments also required a reasonable engineering effort to select the sensors and design and build the boards, components, and the wearable itself. Debugging in such an environment is quite complex because of asynchronous processes, and tracking radio communication in a distributed environment is not trivial. However, most of this work was performed by the hired team.

It also raised other questions regarding the validity of acquired data and usability for prolonged periods regarding comfort and reliability. Although the WKit allows the mobility of users, reliability can be compromised due to instability during data acquisition and transient transmissions.

However, these issues were not focused on in our research. We assume that they will eventually be solved with the evolution of sensors and the development of new platforms for IoT devices. We also did not perform extensive connective tests regarding radio interference and transceiver power configurations, for example.

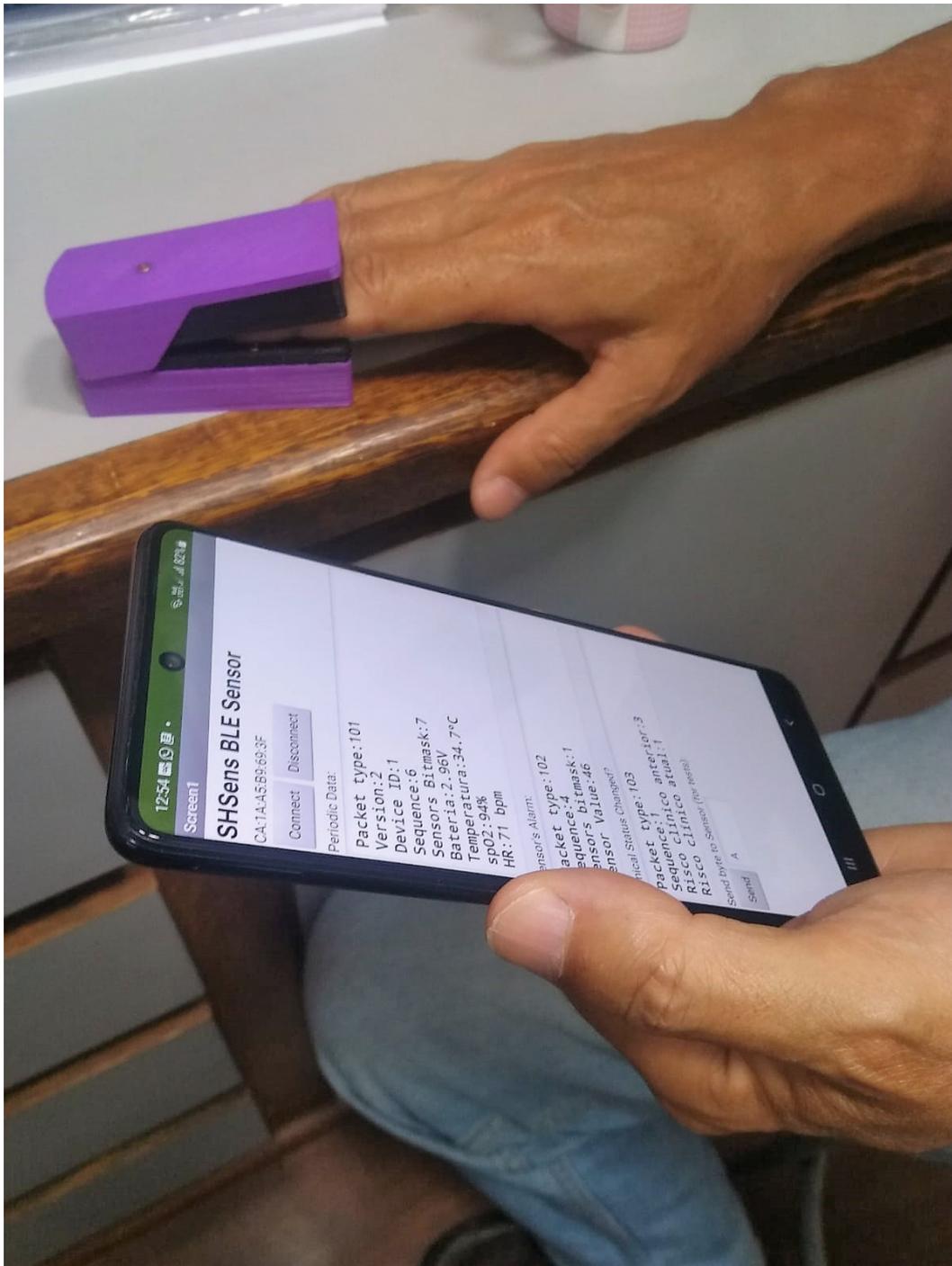


Figure B.8: WKit in use with the MIT App Inventor mobile version and the BLE connection.

In conclusion, our WKit could comfortably run our proposed algorithms and provide actual energy measurements to feed our simulator and project the benefits of the proposed approach of reducing sampling and transmission tasks.

C Web and Mobile Application Prototypes

This appendix presents the main screens and user applications developed to illustrate the use cases of our proposed IoT-based patient monitoring application.

Figure C.1 shows a dashboard developed in Javascript receiving data from our simulator using an MQTT transport layer via Websockets. In this interface, triggered alarms can be recognized and canceled. Recognition of alarms happens when a health professional attends to the alarm. Alarms can be automatically canceled when patients' health improves. Assessments are performed based on NEWS-2 combined scores. Colors represent the four possible clinical statuses: blue (normal, combined score=0), green (worrisome, combined score=1), yellow (severe, combined score=2), and red (critical, combined score=3). This was the first application prototype to show the simulated data and investigate the alarm's life cycle based on composite early-warning scores.

Figure C.2 shows some designed screens of the conceptualized mobile application for a general patient monitoring application. The screen on the left replicates the dashboard view of the web application. It aims at monitoring a group of individuals. The central screen is an individual view of a single patient. Finally, the screen on the right shows the selected time series of health parameters in the last few hours. These screens were designed by Raquel Correa Cordeiro ¹

¹Ph.D. Candidate from the Department of Arts & Design (DAD) - Laboratório de

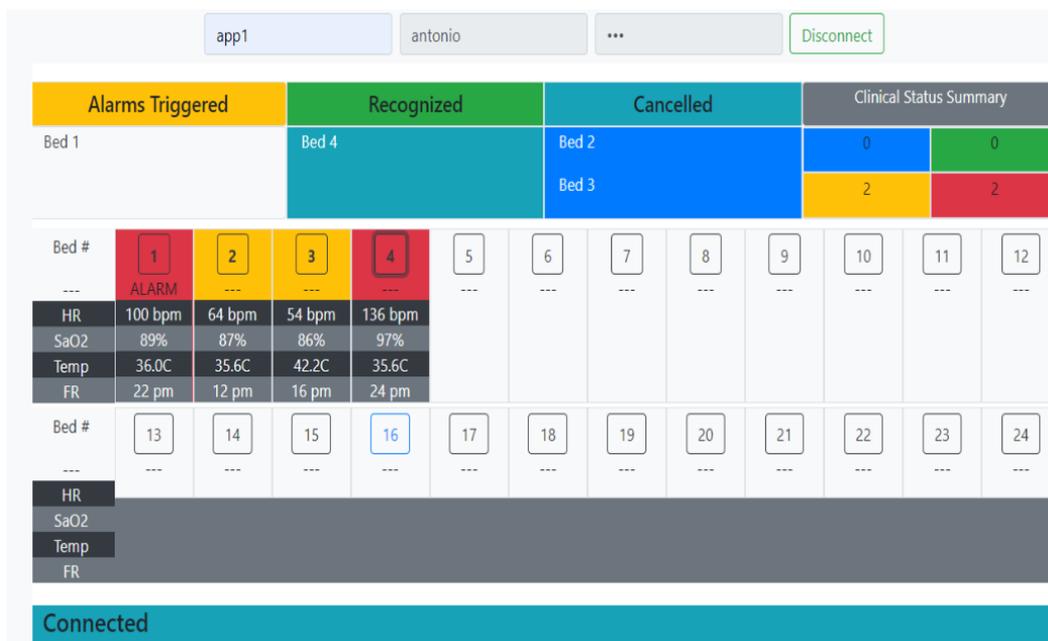


Figure C.1: Infirmery use case - Dashboard of monitored patients

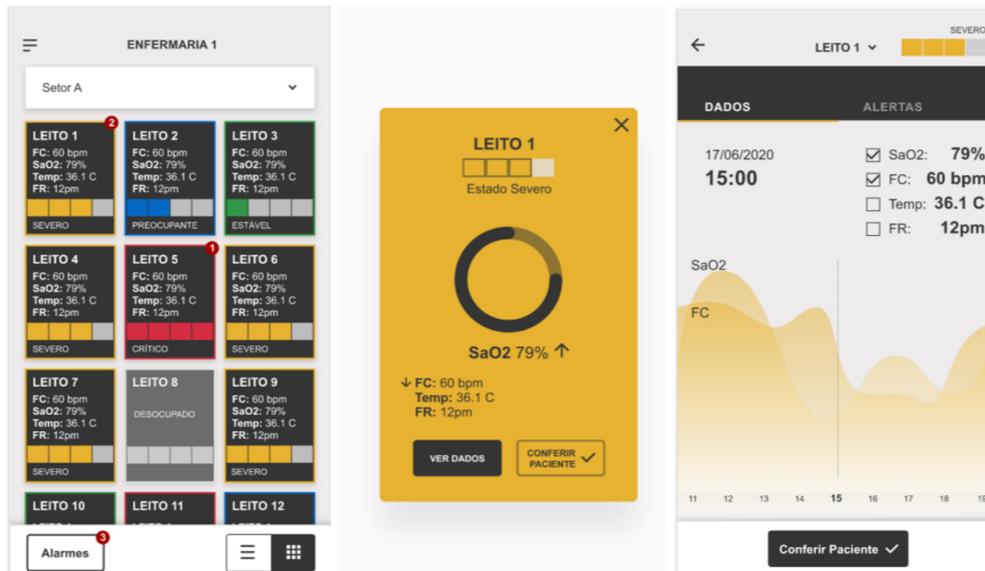


Figure C.2: Mobile application logged user screens design

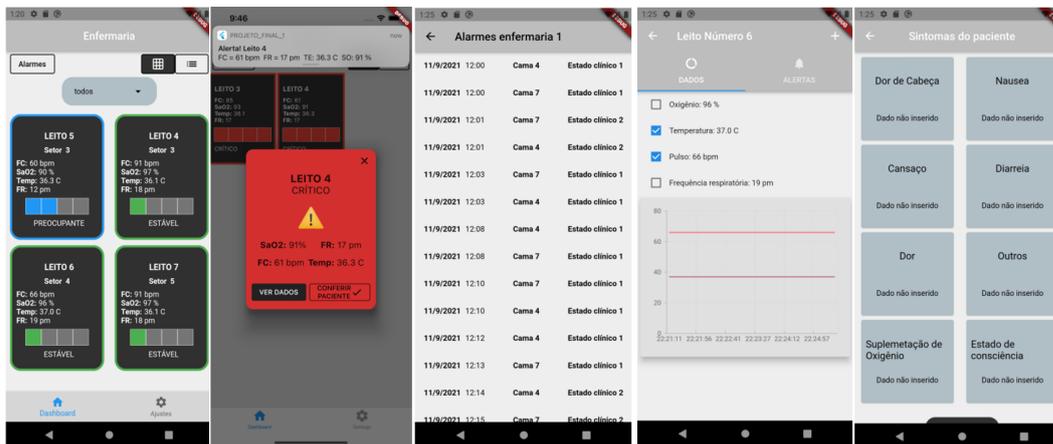


Figure C.3: Flutter Mobile application logged user screens

during the development of our system prototype.

Figure C.3 shows the implemented mobile application screens by Mariela Mendonça de Andrade (ANDRADE, 2022) during her bachelor's final project inspired by the previous design. The project was based on our conceptual architecture and extended some features, such as supporting queries on historical data and patient symptom evaluation. This mobile application did not support a BLE connection. It was aimed at providing a friendly interface for health teams and caregivers to monitor patients. It receives data from the SH-Sens backend produced by our simulations using data from public databases.

Figure C.4 depicts the base architecture view for the development of the Web and Mobile applications. The Kit (WKit) in yellow represents the infirmary use case, where multiple patients send information to a base station acting as a

Ergodesign e Usabilidade de Interfaces (LEUI) - PUC-Rio

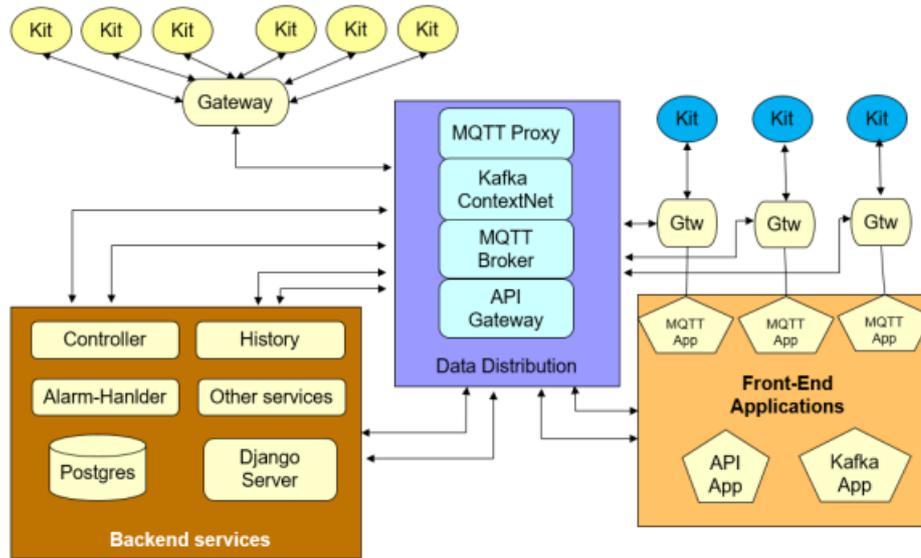


Figure C.4: Designed Architecture for supporting the user applications

gateway. The Kit in blue represents the patients monitored outside clinical settings, at home, for instance. They connect to a mobile phone that will work as a gateway. The user applications within the Front-End Applications box can support different communication mechanisms to exchange information, such as MQTT, RestAPI (API), and Kafka API ². This environment was created in Docker containers to support developed applications.

Figure C.5 shows the mobile application developed in Flutter to test and debug the BLE version of our embedded solution. The interface allowed to start the search for BLE devices manually and select the found device to establish a connection. It also keeps track and shows every exchange message. Although not explored in our proposed solution, the BLE connection allows the connected device to receive data from our mobile application. For example, configuration parameters of the self-adaptive algorithm could be sent through this link. Distributed configuration management in IoT applications is a challenge because it is necessary to provide consistency, privacy, security, availability, compatibility, and other challenging requirements. This application was a first effort to integrate the mobile application developed by (ANDRADE, 2022) with our hardware prototype.

Figure C.6 shows the login screen of the new mobile application supporting a BLE connection and the configuration screen of the wearable device. We have to adapt our application because our hardware prototype only supports three out of the five vital signs (HR, SO, and BTEMP). This information is sent regularly to the SH-Sens backend using the MQTT topics. In addition, the mobile application can also be used by a caregiver who will not have a wearable connected. Therefore,

²<https://kafka.apache.org/>

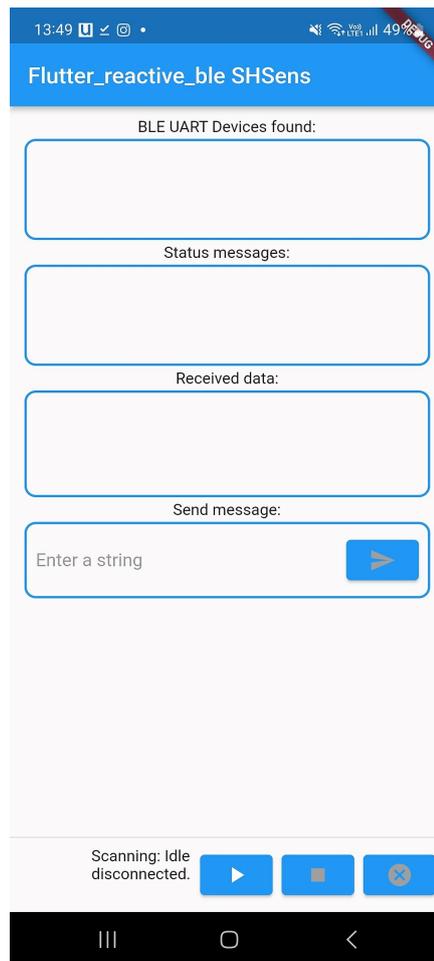


Figure C.5: Mobile application for testing and debugging the BLE connection

the "BLE Enabled" checkbox can be enabled or disabled. The caregiver will receive information from the SH-Sens backend through the MQTT connection.

In addition, the internals of this mobile application was altered to support sending received data from the BLE connection to the MQTT proxy. The previous mobile application only supported receiving information. Moreover, the sent information should also be displayed locally without the need for an internet connection. However, to send information to backend connections from a mobile phone using the public internet, it was configured a public MQTT broker in the Cloud, and a private cloud environment to process the other elements of our patient monitoring system. Different from previous application environments where it was possible to run all the simulated environments in local networks.

These applications were not the focus of this Thesis, but they were important to check the usability of generated data and potential integration problems that embedding a solution in the hardware would arise. It was also important to motivate other students and researchers to collaborate with our work.

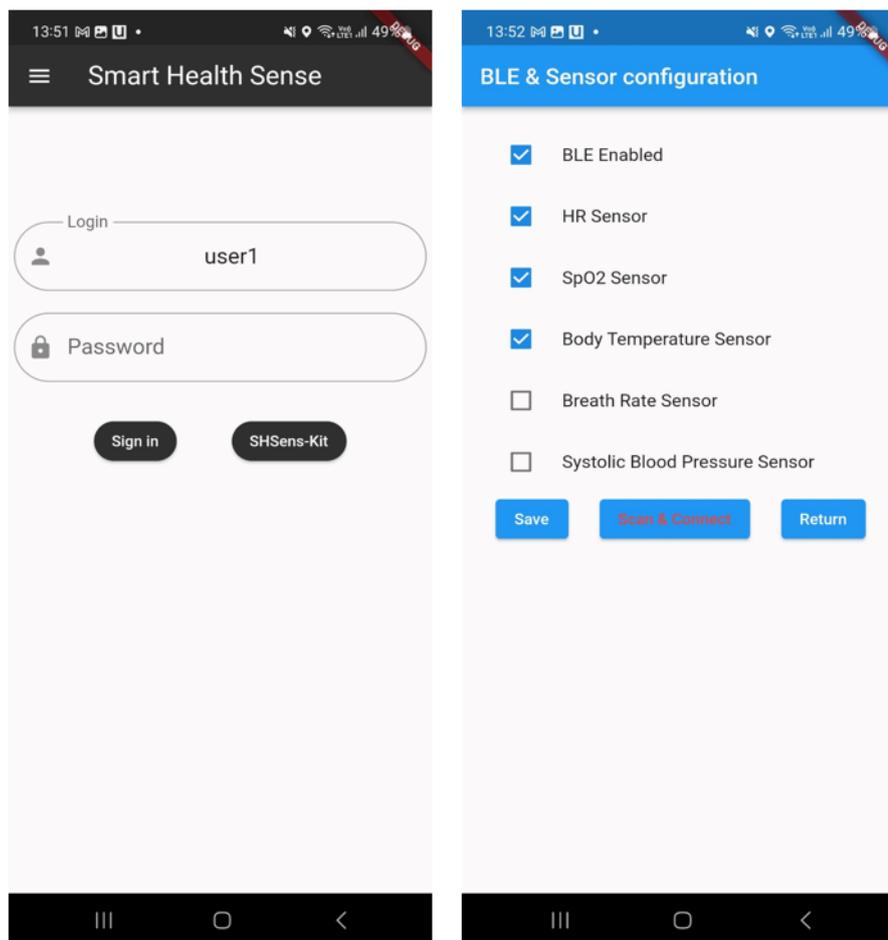


Figure C.6: Flutter Mobile application - login and sensor configuration screens.

D Preliminary Experiment

D.1 Introduction

This experiment implements the first version of our self-adaptive algorithm, based on the principle that frequencies are proportional to the magnitude of the composite early-warning scores (P1), to assess the viability of reducing transmissions and explore the use of combined scores to alarm risk situations.

The reductions were compared to a baseline system that utilizes data acquired from ICU patients using the original sampling rate of a multiparametric hospital monitor and the algorithms proposed by previous studies.

D.2 Objectives

The experiment's main objective is to verify the use of combined early-warning scores and their potential to adapt sampling, processing, and transmission frequencies and handle alarms in order to reduce transmissions compared to the baseline and the algorithms proposed by (ELGHERS; MAKHOUL; LAIYMANI, 2014) and (HARB et al., 2021) and keep the operational fidelity of alarms.

D.3 Methodology

Our simulator was utilized to run the experiment but with a preliminary version of our proposed algorithm described in the next subsection. Thirty-six records from MIMIC and MIMIC-II public databases were used to provide actual physiological parameters for our experiments. Records detailed description can be found in subsection 4.1.1. The physiological markers heart rate (HR), arterial blood saturation (SpO₂), body temperature (BTEMP), and respiratory rate (RR) were collected because they compose the NEWS-2 scoring system.

It was selected from the MIMIC dataset, seven patients ¹, and from MIMIC II, twenty-nine patients ².

Figure D.1 shows the total recorded time in seconds of all selected patients per each combined score in the baseline system. It can be noticed that combined

¹055n, 254n, 259n, 455n, 457n, 474n, 476n

²a44[002, 038, 129, 159, 162, 178, 197, 200, 322, 378, 537, 610, 635, 694, 759, 921]n, a45[222, 260, 384, 467, 519, 532, 580]n, a46[165, 297, 391]n, 3182539n, 3289943n, 3533682n

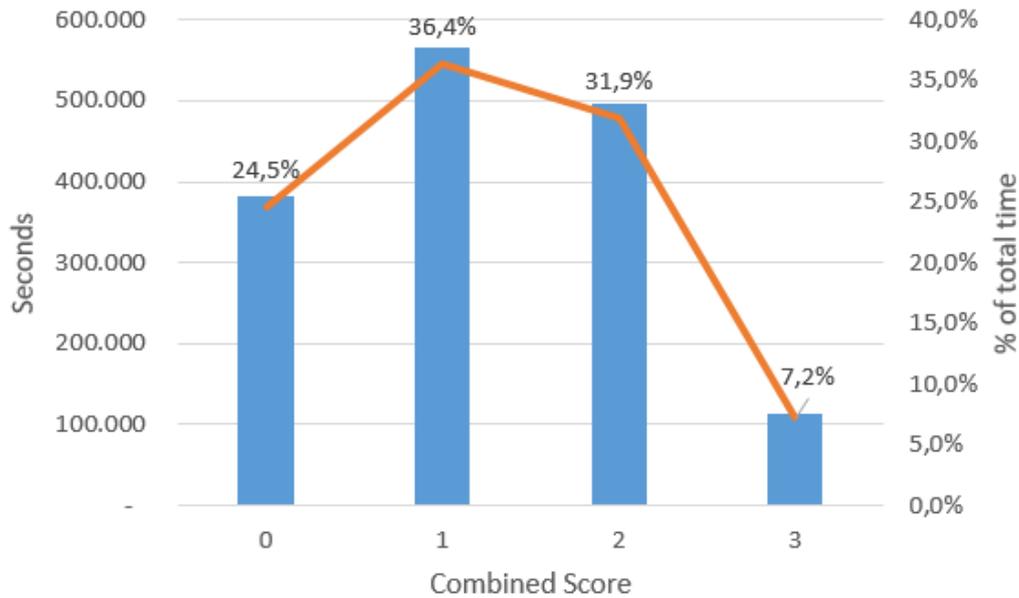


Figure D.1: Total time per combined score recorded in the original dataset (baseline).

Table D.1: Proposed self-configuration parameters based on NEWS-2 combined scores.

Combined Score	Sampling rate (sec)	Processing rate (sec)	Transmission rate (sec)
0	9	18	90
1	6	12	60
2	3	6	30
3	1	2	10

scores zero and one represent more than 60% of the total monitored time. Only 7.2% of the patients' monitored time was in combined score three.

The simulator reads data from all records for 12 hours. It emulates the behavior of algorithms capturing statistical information such as messages sent, payloads, individual NEWS-2 scores, combined NEWS-2 scores, and transitions between combined scores. The adaptive behavior of sampling rates was emulated. Table D.1 presents the proposed algorithm's configuration parameters for the experiment. Each line of Table D.1 represents the assumed configuration related to the calculated combined NEWS-2 score.

D.4

Self-adaptive features

The implemented score mechanism allows the patient's vital sign values to be evaluated. Scores were calculated based on the NEWS-2 guidelines as described in Figures 2.3 2.4. The individual scores related to each vital sign are added to

produce a summed score that is grouped and categorized to form the combined score reflecting the patient's clinical risk. According to the classification of scores calculated periodically, self-adaptive features are implemented in real-time. The procedures performed by our algorithm can be described as follows:

- i. reading vital signs values at the current sampling rate and storing these values in a buffer;
- ii. processing the buffer of sensor values using a statistical method, such as mean and median;
- iii. storing results in a global data structure;
- iv. calculating the individual early-warning scores based on the values on the global data structure, finding the combined score, and storing these values for checking the improvement or deterioration of clinical status;
- v. triggering alarms in the case of patient health deterioration;
- vi. using the combined score to update the self-adaptive function that dynamically changes the sensors' sampling rate, combined score computation (processing rate), and periodic data transmission time intervals (transmission rate).

Let's take the example of a patient being monitored using the WKit. If the combined early-warning score detects a drastic change, an alarm is sent, and in parallel, the self-adaptive parameters can be reconfigured. The reconfiguration procedure adapts the frequency of updates for shorter periods to capture any further deterioration. If the patient's clinical status improves, the frequencies are reconfigured to more extended periods to save battery life and improve efficiency by reducing transmissions, redundancy of data, and storage space.

The following pseudo-code represents the logic of our first proposed algorithm.

```
# Pseudo-code Version (1).
1. combinedScore := 0
2. samplingRate := refSamplingRate[combinedScore]
3. processingRate := refProcessingRate[combineScore]
4. transmissionRate := refTransmissionRate[combinedScore]
5. prevCombineScore := combinedScore
6. time :=0
7. while time <= ExperimentTime do {
8.   when samplingRate {
```

```

9     for each sensor: read data into a circularBuffer[sensor]
10.  }
11.  when processingRate {
12.    for each sensor: proc[sensor] :=
        statMethod(circularBuffer[sensor])
13.    summedScore := sumScores(for each sensor getProc(sensor))
14.    combinedScore := calCombScore(summedScore)
15.    if combinedScore != PrevCombinedScore {
16.      if combinedScore > PrevCombinedScore {
17.        send Alarm
18.      }
19.      samplingRate := refSamplingRate[combinedScore]
20.      processingRate := refProcessingRate[combineScore]
21.      transmissionRate := refTransmissionRate[combinedScore]
22.      PrevCombineScore := CombinedScore
23.    }
24.  }
25.  when transmissionRate {
26.    payload := mountMsg(for each sensor get(proc[sensor]))
27.    sendPeriodicMsg(payload)
28.  }
29.  time++
30.}

```

At the beginning (lines 2-4) of the pseudo-code version (1), the sampling, processing, and transmission rates from a reference table for the Normal condition (score zero) are loaded. The current score (line 5) is stored (PrevCombinedScore) for future comparison. Then, the event scheduler handles asynchronous processes during the experiment (lines 7-30). When the sampling rate timer fires, the values from the corresponding sensors are stored in a circular buffer, one for each sensor (line 9). When the processing rate fires, a pre-defined statistical method is applied to the circular buffer storing the result value into the proc[sensor] vector for the corresponding sensor (line 12). Next, the early warning score is calculated for each sensor (line 13) and combined (line 14), followed by a comparison between the combined score and the previous score. If they are different (line 15) and the clinical condition has worsened (higher combined score), an alarm is sent (line 17), causing the sampling, processing, and transmission rates to be updated to new corresponding rates according to the current score (lines 19-21). The previous score (PrevCombinedScore) variable is also updated (line 22). Finally, when it is time for transmitting, a message with the most recent values calculated by the

Algorithm 1 Local Emergency Detection Algorithm *LED***Require:** R_t (Instantaneous Sampling Rate).

```

1: while  $Energy > 0$  do
2:   for each period do
3:     takes first measure  $r_0$ 
4:     sends first measure  $r_0$ 
5:     takes measures  $r_i$  at  $R_t$  Rate
6:     gets score  $S_i$  of measure  $r_i$ 
7:     if  $S_i \neq 0$  then
8:       sends measure  $r_i$ 
9:     end if
10:  end for
11: end while

```

Figure D.2: Elghers et al. (2014) - Local Emergency Detection algorithm.

Algorithm 1 Emergency Detection Algorithm**Require:** A patient: p , A biosensor: B_v^p , A period time: t ,
Records collected during t : ${}_tR_v^p = [r_1, r_2, \dots, r_t]$.**Ensure:** A set of critical records.

```

1: for each record  $r_i \in {}_tR_v^p$  do
2:   calculate score  $s_i$  of  $r_i$  according to EWS
3:   if  $s_i > 0$  then
4:     send  $r_i$ 
5:   end if
6: end for
7: calculate  $a$  and  $b$  for  ${}_tR_v^p$  based on equation 1
8: send  $a$  and  $b$ 

```

Figure D.3: Harb et al. (2021) proposed algorithm.

processing procedure stored in `proc[sensor]` is built and sent (lines 25-28).

D.5**Comparison to other algorithms**

Figure D.2 shows the Local Emergency Detection (LED) algorithm extracted from Elghers et al. (ELGHERS; MAKHOUL; LAIYMANI, 2014), where abnormal data is sent between periods. Normality is checked using individual early-warning scores. The sampling rate is calculated based on the analysis of variance as explained in Section 2.5.

Figure D.3 shows the Emergency Detection algorithm extracted from Harb et al. (HARB et al., 2021). It is similar to the LED algorithm. However, it also performs a linear regression of vital signs values and sends the found coefficients. The sampling rate is defined using a stability index calculated over individual early-warning scores as explained in Section 2.5.

Table D.2: Sum of payloads and the number of messages considering 5 vital signs of 36 patients being monitored across 12 hours in our simulations.

Algorithms	Payload	Number of Messages	Mean Number of Messages per patient
Baseline	7,794,000	—	—
Elghers et al.	2,069,371	259,740	7,215
Harb et al.	1,112,373	259,740	7,215
Our proposal (*) + 5,726 alarms	218,760	43,752	1,215

For the algorithms of Elghers et al. (ELGHERS; MAKHOUL; LAIYMANI, 2014), and Harb et al. (HARB et al., 2021), pLEN was of 30s, and round size, 4 periods. The minimum frequency was 1/9Hz, and the maximum frequency of 1Hz. The risk factor (r_0) for the Bézier curve was fixed at 0.5.

D.6 Results

This subsection introduces the results of our preliminary experiment. Table D.2 presents the achieved results.

Although hospitals' multiparametric monitors do not process NEWS-2 scores in real time, it is important to assess whether our solution differs from actual monitoring data if they do. Figure D.4 shows the total number of alarm differences between the baseline, simulating the alarms using the original sampling rate and triggered by our algorithm. Composite score transitions for non-negligible alarms was considered in the baseline when they remained for at least 10s.

D.7 Discussion

In this first experiment, our objective was to verify the use of combined early-warning scores and their potential to adapt sampling, processing, and transmission frequencies and handle alarms, and the potential to reduce transmissions. In this version, it is established fixed sampling rates for each combined early-warning score and alarm conditions were processed when the adaptive sampling was reassessed. Thus, another focus of this experiment was on validating our simulator and the feasibility of our approach.

Table D.2 shows that our proposal has the potential to reduce the number of sent messages and network traffic drastically. A major difference in the number of messages is due to our algorithm combining scores and sending one message with all five sensors' information instead of one message per sensor utilized by the other algorithms. Even if one message were sent per sensor, our solution would

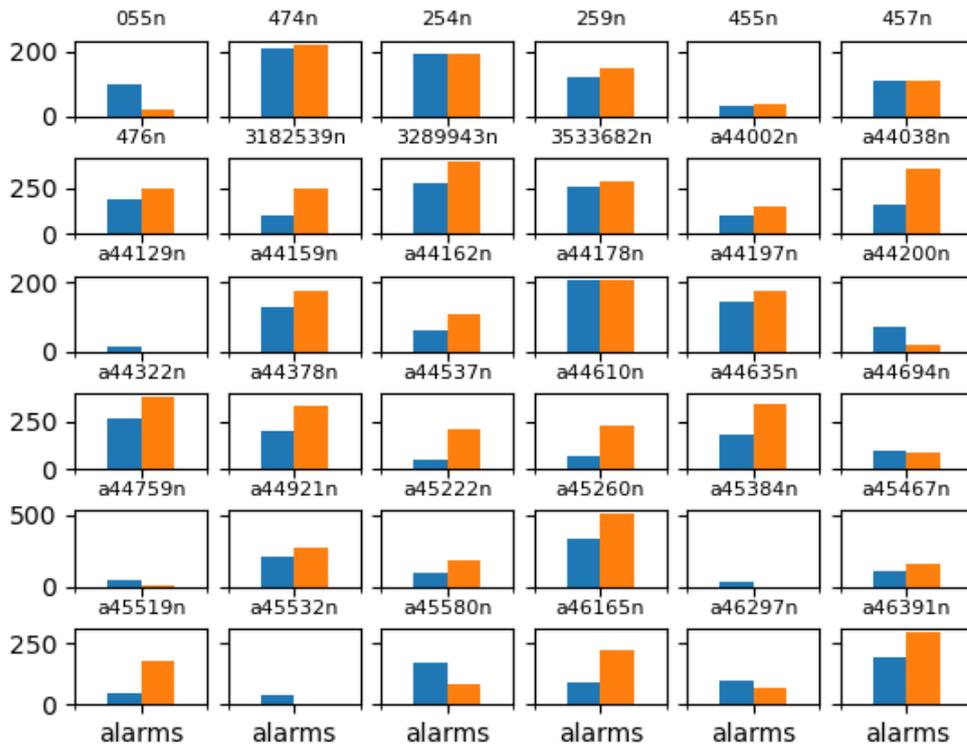


Figure D.4: Comparison of the number of transitions of increasing combined scores in original patients data (blue) and captured by our proposal (orange) using NEWS-2.

send a lower number of messages ($43,752 * 5 = 218,760 + 5,726 = 224,486$, 13,6% lower).

The huge reductions in the payload are due to reduced data sampling. Our self-adaptive algorithm is based on the patient's health condition and analyzed in real-time using the combined scores. Observing Figure D.1, it is noticeable that most of the time, patients were in the combined scores zero and one, and only about 7% of time in combined score 3, which leads to higher frequencies. Varying frequencies using patient health status seems much more efficient than data variability.

Additionally, differently from related studies, the algorithm utilized the same sampling frequency for all sensors. Then, the variability of one single vital sign would have less influence in increasing frequencies.

The dataset's total number of abnormal values (NEWS2 individual score > 0) is 2,375,730. The adaptive sampling utilized by Elghers et al. (ELGHERS; MAKHOUL; LAIYMANI, 2014) based on the analysis of variance of nominal values reduced about 13% of it, while Harb et al. (HARB et al., 2021) using the stability index of individual scores reduced about 53%. Thus, with the stability index, it was possible to utilize much lower sampling frequencies. Finally, our algorithm, using the combined scores and ignoring the variance of data, was able to reduce even

more the payload.

Moreover, our approach seems to have the potential to cope with alarms better, sending them only when an evident deterioration of health is realized based on the NEWS-2 reference. On the other hand, alarm algorithms based on threshold would trigger more than 2 million alarms when individual scores are greater than zero.

In addition, a basic assessment of monitoring quality loss was performed based on alarm events triggered by increasing combined score transitions.

Figure D.4 shows that for a few cases, such as with patients 254n, 455n, 457n, and a44178n, there is an almost perfect match in the number of alarms between our algorithm and the baseline. In contrast, for 9 cases (25% - 055n, a44129n, a44200n, a44694n, a44759n, a45384n, a45532n, a45580n, a46297n), our solution did not trigger alarms for all detected deterioration transitions using the original frequency (1Hz) for readings and processing the NEWS-2 combined scores. Finally, for 23 patients, our proposed algorithm triggered more alarms. Probably, because our algorithm considered captured transitions using our adaptive sampling, even the ones shorter than 10s. Remembering that it was filtered in the graph, alarms with transitions shorter than 10s in the baseline system. This method was not effective in observing alarm differences.

Our system could dynamically change sensor reading frequency based on a patient's clinical status: a low (high) combined score would lead to a high (low) sensor reading time interval. The same reasoning applies to periodic transmission. In this manner, the system execution varies depending on the patient's clinical status. These factors allow the system to minimize the number of alarms compared to threshold alarms while providing robust real-time monitoring and potentially optimizing resource use (e.g., energy utilization, communication channels), allowing devices such as the WKit to be used more constantly.

Furthermore, the experiment also was essential to understand subtle aspects of related work and how the advantages presented in those studies could be incorporated into our new approach.

D.8 Conclusion

In this experiment, it was possible to verify the potential of using combined scores to regulate transmissions efficiently in IoT-PMAs.

Self-adaptive features support various actions the system can autonomously execute to configure itself. They include, for example, the ability of the system to dynamically change the sensor reading, processing interval, and periodic transmis-

sion, as utilized in our proposal. The source and magnitude of adaptation drivers are fundamental to achieving the desired results.

Our approach of decreasing sampling and transmission frequencies when the combined scores are lower could achieve high reductions in the number of messages and payloads. Validating our principle P1, the higher the patient's risk, the higher the frequency, and vice-versa. Furthermore, using a method that can perceive a patient's health condition instead of using the variability or stability of monitoring values seems more appropriate for regulating wearable health monitoring sensors regarding a better use of resources. However, the effect of adapting sampling rates on monitoring quality was not clearly verified. In addition, combined scores were also utilized to handle alarms when scores increased, reflecting a worsening in patient health. Nonetheless, a more robust method to verify the accuracy and precision of such alarms related to the baseline system such be devised.

This experiment was also important to better understand the proposed methods and benefits of the related work. Our approach is distinct, but a new proposal can take advantage of techniques used in those studies and add the similarity of data to adapt sampling rates. In conclusion, this experiment provides the basis for developing the final version of our algorithm explained in Chapter 3.