

UNIVERSIDADE FEDERAL DO PARANÁ

CAINÃ DOS PASSOS

UM MECANISMO DE COMPRESSÃO DE DADOS LOCAL EM TEMPO REAL
ENERGETICAMENTE EFICIENTE PARA REDES WBANS

CURITIBA PR

2020

CAINÃ DOS PASSOS

UM MECANISMO DE COMPRESSÃO DE DADOS LOCAL EM TEMPO REAL
ENERGETICAMENTE EFICIENTE PARA REDES WBANS

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Informática no Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, da Universidade Federal do Paraná.

Área de concentração: *Ciência da Computação*.

Orientador: Aldri Luiz dos Santos.

CURITIBA PR

2020

Catálogo na Fonte: Sistema de Bibliotecas, UFPR
Biblioteca de Ciência e Tecnologia

P289m Passos, Cainã dos
Um mecanismo de compressão de dados local em tempo real energeticamente eficiente para
redes WBANS [recurso eletrônico] / Cainã dos Passos. – Curitiba, 2020.

Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-
Graduação em Informática, 2020.

Orientador: Aldri Luiz dos Santos.

1. Sistemas de transmissão de dados. 2. Processamento eletrônico de dados em tempo real. I.
Universidade Federal do Paraná. II. Santos, Aldri Luiz dos. III. Título.

CDD: 621.38216

Bibliotecária: Vanusa Maciel CRB- 9/1928

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em INFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **CAINA DOS PASSOS** intitulada: **UM MECANISMO DE COMPRESSÃO DE DADOS LOCAL EM TEMPO REAL ENERGETICAMENTE EFICIENTE PARA REDES WBANS**, sob orientação do Prof. Dr. ALDRI LUIZ DOS SANTOS, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 08 de Junho de 2020.

Assinatura Eletrônica

11/06/2020 10:17:17.0

ALDRI LUIZ DOS SANTOS

Presidente da Banca Examinadora (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

13/06/2020 19:29:25.0

DANIEL LUDOVICO GUIDONI

Avaliador Externo (UNIVERSIDADE FEDERAL DE SÃO JOÃO DEL REI)

Assinatura Eletrônica

16/06/2020 14:24:33.0

LUIZ EDUARDO SOARES DE OLIVEIRA

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

À minha esposa, meus pais e irmãos.

AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Aldri Luiz dos Santos, agradeço por todos os ensinamentos ao longo de todo o Mestrado. Pela sua incansável paciência, entender minhas limitações e estimular meu pensamento crítico e científico durante suas orientações em todo o curso de Mestrado em Informática da Universidade Federal do Paraná (UFPR). Mais uma vez, muito obrigado.

Ao Programa de Pós-Graduação em Informática da Universidade Federal do Paraná, pelo auxílio administrativo e pela infraestrutura computacional. Agradeço também aos funcionários do departamento pelo auxílio com processos burocráticos da instituição.

Aos amigos do NR2 e CCSC Agnaldo Batista, Andressa Vergütz, Arthur Emilio Garcete Ferreira, Benevid Felix Silva, Bruno Henrique Schwengber, Bruno Marquez Cremonezi, Carlos Alberto Pedroso Junior, Danilo Rodrigo Possati, Diego Milhomem Schmitt, Euclides Peres Farias Junior, Fernando Nakayama, Gustavo Henrique Carvalho de Oliveira, Igor Steuck Lopes, Ligia Francielle Borges, Marcos Antônio Dellazari, Mateus Pelloso, Nelson Gonçalves Prates Junior, Paulo Lenz Junior, Rafael Araújo da Silva e Yan Uehara de Moraes pela convivência, pelas argumentações e discussões técnicas e científicas, bem como pelo apoio durante o andamento das disciplinas e desenvolvimento da pesquisa.

Agradeço a toda equipe Data Stream Sistemas, André, Cassiel, Célia, Denise, Diego, Fran, Flávia, Gustavo, Jeferson, Jeison, Julio, João, Maria, Paloma, Rodrigo, Thairo, Thiago e Younan, por toda compreensão e apoio na minha ausência. Sou eternamente grato aos gestores Andre Alexandre de Aguiar e Júlio César Teixeira pelo apoio incondicional. #godatastreamers.

Aos meus pais, Cicero Jayme dos Passos e Simone Weiss dos Passos por toda educação, ensinamentos e valores prestados ao longo da minha vida. Aos meus irmãos, Ana Luiza, Fernã e Luan dos Passos pelo companheirismo, amizade e irmandade.

À minha esposa Mayra, pela sua compaixão, confiança, dedicação e especialmente paciência nos momentos de grandes dificuldades. Pela sua compreensão e por ficar ao meu lado nos momentos em que nada parecia possível.

RESUMO

A evolução das tecnologias sem fio tem possibilitado a criação de diversos tipos de redes para suportar os serviços do dia-a-dia das pessoas. Entre as redes voltadas para o monitoramento dos cuidados com a saúde, destacam-se as redes corporais sem fio (WBANs), que possibilitam um monitoramento contínuo e em tempo real dos dados fisiológicos. Sem a restrição dos fios, elas permitem mais mobilidade para as pessoas, e a possibilidade do monitoramento remoto. Em 2016, aproximadamente 7 milhões de pessoas tiveram monitoramento remoto por meio de dispositivos móveis e aplicações médicas. Isto deve-se ao fato das WBANs, juntamente com as aplicações médicas, disponibilizarem um monitoramento contínuo, ubíquo e flexível das pessoas. Porém este monitoramento acarreta excessiva transmissão de dados e eleva o consumo de energia dos dispositivos. A literatura apresenta diversas soluções para lidar com o consumo de energia, sejam baseadas em protocolos de comunicação, algoritmos de roteamento e sensoriamento e compressão de dados. Embora algumas abordagens reduzam o consumo de energia, poucas desconsideram a transmissão de informações redundantes. Desta forma, este trabalho apresenta um mecanismo de compressão de dados local em tempo real energeticamente eficiente para redes WBANs chamado GROWN (*Ener Gy Efficient Local Data Compression for TRansmission Over WBAN*). Ele gerencia a redundância da informação nos dispositivos corporais e emprega métodos de compressão de dados local encontrados nas WSNs e limiares definidos, reduzindo a transmissão das informações e o consumo de energia sem impactar de forma intensa no atraso total da transmissão da informação para uma estação base. Assim, as informações transmitidas são comprimidas sem perdas por meio de tabela fixa de codificação. Além disso, o mecanismo estabelece limiares definidos e analisa a diferença da correlação entre amostras consecutivas. Dessa forma, apenas valores superiores a este limiar serão comprimidos para transmissão. O GROWN foi avaliado através de experimentações e os resultados obtidos demonstram que o mecanismo aumentou o tempo de vida útil dos dispositivos corporais em até 53,73% com uma latência máxima de 55ms entre informações consecutivas, comprovando sua eficiência energética e emprego em aplicações WBANs de tempo real.

Palavras-chave: WBAN, Eficiência Energética, Transmissão de dados, Tempo Real

ABSTRACT

The evolution of wireless technologies has made it possible to create different types of networks to support people's day-to-day services. Among the networks focused on health care monitoring, wireless body networks (WBANs) stand out, which enable continuous and real-time monitoring of physiological data. Without the restriction of the wires, they allow more mobility for people, and the possibility of remote monitoring. In 2016, approximately 7 million people had remote monitoring through mobile devices and medical applications. This is due that WBANs together with medical applications, provide continuous, ubiquitous and flexible monitoring of people. However, this monitoring causes excessive data transmission and increases the power consumption of the devices. The literature presents several solutions to deal with energy consumption, whether based on communication protocols, routing algorithms and data sensing and data compression. While some approaches reduce energy consumption, few overlook the transmission of redundant information. Thus, this work presents an energy efficient local data compression mechanism in real time for WBAN networks called GROWN (*Ener Gy Efficient Local Data Compression for TRansmission Over WBAN*). It manages the redundancy of information on body devices and apply methods of local data compression found in WSNs and defined thresholds, reducing the transmission of information and energy consumption without having a strong impact on the total delay of the transmission of information to the body-gateway. Thus, the transmitted information is compressed without loss through a fixed coding table. In addition, the mechanism establishes defined thresholds and analyzes the difference in correlation between consecutive samples. In this way, only values higher than this threshold will be compressed for transmission. GROWN was evaluated through experiments and the results obtained demonstrate that the mechanism increased the useful life of body devices by up to 55.73% with a maximum latency of 55ms between consecutive information, proving its energy efficiency and use in real-time WBAN applications.

Keywords: WBAN, Energy efficient, Data transmission, Real-time

LISTA DE FIGURAS

2.1	Representação dos três níveis de Arquitetura de uma rede WBAN.	7
2.2	Classificação de acordo com o tipo de Arquitetura segundo Castanedo (Castanedo, 2013)	14
2.3	Classificação entre diferentes entradas de informação definida por Durrant (Durrant-Whyte, 1988).	15
2.4	Representação de compressão e descompressão segundo Sayood (Sayood, 2017).	16
2.5	Representação de Compressão de dados local sem perdas definida por Pu (Pu, 2005)	17
2.6	Representação de compressão de dados local com perdas de acordo com Pu (Pu, 2005)	17
2.7	Probabilidade dos cinco símbolos com codificação de Huffman exposto por Salomon (Salomon, 2004)	19
3.1	Diagrama de bloco do compressor e descompressor do método LEC	28
3.2	Reprodução de compressão de dados local com perda LTC (Schoellhammer et al., 2004)	31
3.3	Modelo de um quantizador codificador e decodificador DPCM (Gray e Neuhoff, 1998)	31
3.4	Diagrama de bloco híbrido proposto por Deepu	33
3.5	Diagrama de bloco híbrido do compressor e descompressor.	34
4.1	Modelo do ambiente de coleta e disseminação de sinais fisiológicos	37
4.2	Arquitetura da abordagem proposta.	37
4.3	Exemplo de funcionamento	41
5.1	Etapas do aplicativo no dispositivo <i>sink</i>	44
5.2	Placas Arduino utilizadas no <i>testbed</i>	44
5.3	Componentes essenciais para os dispositivos corporais	45
5.4	Dispositivo Temperatura	45
5.5	Dispositivo ECG	46
5.6	Dispositivo PPG.	46
5.7	Dispositivo ECGBD	47
5.8	Testbed do dispositivo Temperatura	48
5.9	(AC) e (AD) alcançado pelo GROWN nos sinais sensoriados	50

LISTA DE TABELAS

2.1	Padrão IEEE 802.15.6.	9
2.2	Aplicações de uma WBAN	10
2.3	Sinais corporais de uma WBAN e suas características	11
2.4	Categorias de acordo com o nível de abstração determinadas por Dasarathy (Dasarathy, 1997)	15
2.5	Tabela de Codificação de Golomb com $m = 2$ a 9	20
2.6	Tabela de codificação de Exponencial de Golomb para diversos valores de n com ordem $k = 0$ até $k = 3$	21
3.1	estado-da-arte dos trabalhos de compressão de dados local para WSNs e WBANs	25
3.2	Tabela de Codificação de Huffman para o método de compressão sem perdas de Marcelloni e Vecchio	27
3.3	Taxas de compressão com o método simples de Marcelloni e Vecchio	27
3.4	Taxas de compressão com o método LEC	28
3.5	Consumo de energia dos dispositivos com diferentes duty cycles apresentados por Deepu <i>et. al</i> (Deepu et al., 2017).	33
4.1	Tabela com valores dos índices n_i e grupos s_i	38
4.2	Regras de compressão dos dados	39
4.3	Representação da tabela com valores de a_i para os primeiros índices de n_i	39
4.4	Tabela de codificação proposta	40
4.5	Referências dos dispositivos da rede	41
4.6	Representação dos exemplos ao longo do tempo com limiar $T = 1$	42
5.1	Especificações Técnicas.	44
5.2	Atraso médio dos dispositivos sensorizados (AM)	51
5.3	Intervalo de Confiança do (AM)	51
5.4	Consumo energético dos dispositivos (CED).	52
5.5	Intervalo de Confiança do (CED).	52
5.6	Tempo de vida dos dispositivos.	53
5.7	Taxa de Compressão (TC)	53

LISTA DE ACRÔNIMOS

AP	<i>Access Point</i> (Ponto de Acesso)
ADC	<i>Analogic-to-Digital Converter</i> (Conversor Analógico-Digital)
BCU	<i>Body Control Unit</i> (Unidade de Controle do Corpo)
BLE	<i>Bluetooth Low Energy</i> (Bluetooth de baixo consumo)
CPS	<i>Cyber-Physical Systems</i> (Sistemas Ciber-físicos)
CSMA/CA	<i>Carrier Sense Multiple Access / Collision Avoidance</i> (Acesso Múltiplo de Detecção da Operadora com Prevenção de Colisão)
DPCM	<i>Differential Pulse Code Modulation</i> (Modulação por Codificação Diferencial de Pulso)
ECG	<i>Electrocardiogram</i> (Eletrocardiograma)
EEG	<i>Electroencephalogram</i> (Eletroencefalograma)
EMG	<i>Electromyography</i> (Eletromiografia)
GROWN	EnerGy Efficient Local Data Compression for TRansmission Over WBAN (Compressão de dados local energeticamente eficiente para transmissão em redes corporais sem fio)
IEEE	<i>Institute of Electric and Electronic Engineers</i> (Instituto de Engenheiros Eletricistas e Eletrônicos)
IoT	<i>Internet of Things</i> (Internet das Coisas)
LEC	<i>Lossless Entropy Code</i> (Código de Entropia Sem Perdas)
MAN	<i>Metropolitan Area Network</i> (Rede Metropolitana)
PPG	<i>PhotoPlethysmoGram</i> (Fotoplestígrafia)

QoS	<i>Quality of Service</i> (Qualidade de Serviço)
SMS	<i>Short Message Service</i> (Serviço de Mensagens Curtas)
TG6	<i>Task Group 6</i> (Grupo de Tarefa 6)
WBAN	<i>Wireless Body Area Networks</i> (Redes Corporais Sem Fio)
WHO	<i>World Health Organization</i> (Organização Mundial da Saúde)
Wi-Fi	<i>Wireless Fidelity</i> (Fidelidade Sem Fio)
WPAN	<i>Wireless Personal Area Network</i> (Rede Pessoais Sem Fio)
WSN	<i>Wireless Sensors Networks</i> (Redes de Sensores Sem Fio)

LISTA DE SÍMBOLOS

AM	Atraso médio
AC	Atraso de Compressão
C	Conjunto de valores
CED	Consumo de energia dos dispositivos
C_i	índice de cada valor pertencente ao Conjunto C
D	Conjunto de dispositivos
\hat{e}_{ts}	Erro de predição
Id	Id do dispositivo na rede
$Inf(ts)$	Informação no tempo
T	limiar
x_{ts}	informação convertida de analógico para digital
\hat{x}_{ts}	leitura anterior

SUMÁRIO

1	INTRODUÇÃO	1
1.1	MOTIVAÇÃO	2
1.2	PROBLEMA	3
1.3	OBJETIVO	4
1.4	CONTRIBUIÇÕES	5
1.5	ESTRUTURA DO TEXTO	5
2	FUNDAMENTOS	6
2.1	REDES CORPORAIS SEM FIO (WBANS)	6
2.1.1	Topologia e arquitetura	6
2.1.2	O Padrão 802.15.6	8
2.1.3	Aplicações de uma WBAN	9
2.2	FUSÃO DE DADOS	11
2.2.1	Entropia da informação	11
2.2.2	Terminologias da fusão de dados	12
2.2.3	Classificação	13
2.2.4	Compressão de dados local	15
2.3	RESUMO	22
3	MÉTODOS DE COMPRESSÃO DE DADOS LOCAL APLICÁVEIS ÀS WBANS	23
3.1	VISÃO GERAL	23
3.2	COMPRESSÃO DE DADOS LOCAL SEM PERDAS	26
3.3	COMPRESSÃO DE DADOS LOCAL COM PERDAS	30
3.4	COMPRESSÃO DE DADOS LOCAL HÍBRIDA	32
3.5	DISCUSSÃO	34
3.6	RESUMO	35
4	O MECANISMO GROWN PARA COMPRESSÃO DE DADOS LOCAL EM TEMPO REAL EM REDES WBANS	36
4.1	VISÃO GERAL	36
4.1.1	Modelo de rede	36
4.2	ARQUITETURA GROWN	37
4.2.1	Módulo gerência de compressão	37
4.2.2	Módulo gerência de descompressão	39
4.2.3	Funcionamento do GROWN	40
4.3	RESUMO	42

5	AVALIAÇÃO	43
5.1	IMPLEMENTAÇÃO	43
5.2	AMBIENTE DE EXPERIMENTAÇÃO E MÉTRICAS	46
5.3	RESULTADOS E ANÁLISES	49
5.3.1	Atraso de compressão e descompressão	49
5.3.2	Atraso médio	51
5.3.3	Consumo energético do dispositivo.	51
5.3.4	Taxa de compressão.	53
5.4	DISCUSSÃO	53
5.5	RESUMO	54
6	CONCLUSÕES	55
6.1	TRABALHOS FUTUROS	56
	REFERÊNCIAS	57
	APÊNDICE A – PUBLICAÇÕES	65
A.1	ARTIGOS PUBLICADOS	65
A.2	PREMIAÇÕES	65

1 INTRODUÇÃO

A conectividade e a expansão dos serviços online têm facilitado a comunicação entre as pessoas. Esta facilidade desempenha um papel eminente para a sociedade, que presencia o avanço tecnológico e seu elevado fator junto ao progresso e desenvolvimento. Este feito tornou-se possível devido aos benefícios e o avanço das tecnologias de comunicação sem fio. As tecnologias sem fio permitem a construção de diversos tipos de redes que apoiam ou modificam os serviços do nosso cotidiano. Muitos desses serviços voltam-se para soluções de problemas em ambientes urbanos como por exemplo, serviços voltados para radares, detecção de velocidade de veículos, ou até mesmo para análise de veículos parados sobre a faixa de pedestres (Vecchio et al., 2014). Neste sentido, detecta-se que os serviços necessitam de características em comum como a coleta de dados em diferentes ambientes e de diversas formas e a transmissão destes dados para uma central de acompanhamento. Assim, muitos destes serviços utilizam as redes de sensores sem fio por sua habilidade de coletar dados dos ambientes, e mais recentemente as redes IoT (Balaji et al., 2019; Lima, 2010).

Uma rede de sensores sem fio (WSN, do inglês, *Wireless Sensor Networks*) é composta por diversos sensores que podem ser inseridos em diferentes ambientes com o objetivo de coletar sinais físicos de maneira autônoma. Muitas das WSNs são capazes de medir grandes quantidades de sinais físicos como pressão, som, temperatura, entre outros. Além disso, ela tem sido aplicada em redes MAN (do inglês, *Metropolitan Area Network*), para cidades inteligentes, em locais como bares, restaurantes e teatros. Normalmente, estes sensores são acoplados a uma placa central responsável por gerenciar o processamento do sinal coletado e a transmissão da informação para uma estação base (*sink*) (Rashid e Rehmani, 2016). Assim, o conjunto de componentes como um transmissor de rádio, usado para comunicação, um microprocessador, usado para controlar o sensor e, a bateria, que alimenta todos os componentes formam um dispositivo (Yick et al., 2008). Estes dispositivos apresentam tamanhos diferenciados, sendo do tamanho de um grão de areia até o tamanho de uma caixa de sapatos. Atualmente, estes dispositivos são usados nas indústrias para o controle do estado de uma máquina, controle do processo industrial, entre outras atividades (Yick et al., 2008). Além disso, a grande evolução nas WSNs e em nanossistemas tem proporcionado o desenvolvimento de dispositivos inteligentes vestíveis ou implantáveis no corpo humano que auxiliam na tomada de decisão de um profissional de saúde. Assim, a união entre dispositivos inteligentes e as redes de sensores sem fio deu-se origem as redes corporais sem fio.

As redes corporais sem fio (WBAN, do inglês, *Wireless Body Area Networks*) através de seus dispositivos têm como propósito coletar, monitorar e transmitir sinais fisiológicos a diversas aplicações médicas e a profissionais de saúde, não interferindo na vida diária das pessoas (Movassaghi et al., 2014). As WBANs buscam oferecer um monitoramento contínuo dos pacientes de maneira remota. Sem a restrição dos fios, ela permite aos usuários uma maior mobilidade, onde em muitos casos, sendo desnecessária a sua permanência em um hospital (Resque et al., 2019). Este cuidado em domicílio (do inglês, *Home Care*) reduz o risco de infecção hospitalar e outras situações decorrentes da internação (Yang et al., 2020). Em 2021 espera-se que 22,9 milhões de pacientes serão monitorados remotamente por meio de aplicações médicas e/ou dispositivos (Mack, 2017). Isto deve-se ao fato das WBANs, juntamente com as aplicações médicas, disponibilizarem um monitoramento contínuo, ubíquo e flexível das pessoas. Logo, as duas têm recebido grande atenção de pesquisadores e empresas nos últimos anos (Cavallari et al., 2014). Diversos projetos como o *ResearchKit* da Apple, *CodeBlue* da Universidade de Harvard e o *Jordan Healthcare Initiative* da Cisco têm sido desenvolvidos,

visando melhorar a qualidade de vida e os cuidados dos pacientes. Contudo, uma WBAN vai além do contexto médico e pode ser aplicada em setores de entretenimento, esportes e militar (Crosby et al., 2012). Isto ocorre devido aos sensores corporais coletarem diferentes tipos de dados, como movimentos corporais para jogos de entretenimento, batimento cardíaco e temperatura corporal para atletas de alto desempenho e características do ambiente e sinais de fadiga de soldados em campos de batalha (Movassaghi et al., 2014).

Em uma WBAN os dispositivos são caracterizados por sua heterogeneidade, variando suas capacidades de memória, tarefas, tamanhos, taxas de amostragem, recursos necessários, níveis de inteligência e tipos de sinais coletados (Mahapatro et al., 2011). No geral, os dispositivos são energeticamente alimentados por baterias (Bontorin et al., 2015). Em comparação com as redes de sensores sem fio, as WBANs possuem menos sensores com pequenas baterias, adicionando mais restrições energéticas (Wu et al., 2010). Além disso, deve-se considerar que o consumo de energia varia de acordo com o tipo de sensor e o tipo da aplicação. Por isso, avanços ainda precisam ser feitos para manter os serviços das WBANs em funcionamento, tal que a informação seja eficaz e possa salvar vidas (Dhanvijay e Patil, 2019).

1.1 MOTIVAÇÃO

O envelhecimento da população mundial tem ampliado rapidamente devido ao aumento da longevidade, principalmente atribuível à melhoria substancial na nutrição, medicina e saúde pública. No Brasil, o envelhecimento populacional ultrapassou a previsão do Instituto Brasileiro de Geografia e Estatística (IBGE) de 2017. Além disso, estima-se que em 2060 um terço da população brasileira será de pessoas idosas (de Geografia e Estatística, 2020)¹. Segundo Ghamari *et. al* em (Ghamari et al., 2016) no Reino Unido, a população com mais de 85 anos está prevista para quase triplicar até 2035, nos Estados Unidos, estima-se que a população com mais de 65 anos duplique até 2040, na China, a população com mais de 60 anos deverá dobrar até 2040 e, até o ano de 2050 o Japão terá a população mais velha da história da humanidade, com uma idade média de 52 anos.

Essas estimativas e estatísticas indicam o fato de que a prestação contínua de serviços de saúde a pacientes e o aumento do número de idosos com várias dificuldades de saúde aumenta significativamente o custo dos sistemas de saúde. Visto que, os sistemas atuais de saúde são amplamente baseados em redes centralizadas com fio, onde os pacientes só conseguem receber diagnóstico ou tratamento médico em hospitais. Logo, os sistemas de saúde tornam-se insustentáveis em sua forma atual (Anderson e Horvath, 2004). Segundo Islam *et. al* em (Islam et al., 2015) a crescente evolução das aplicações médicas é essencial para os sistemas de saúde do futuro. Além disso, torna-se indispensável a detecção e o diagnóstico precoce da doença, uma vez que, retardar efetivamente seu progresso reduz significativamente o custo dos sistemas de saúde (Levin e Stevens, 2011; Fouad, 2014). Assim, o acompanhamento das informações fisiológicas das pessoas por aplicações de tempo real auxilia os profissionais da saúde na tomada rápida de decisão (Vergütz et al., 2020).

As aplicações de tempo real são aquelas que exigem a necessidade de redes com atraso limitado na entrega dos pacotes. Estas aplicações são capazes de monitorar, responder imediatamente à entrada do usuário ou controlar um ambiente externo. O ambiente externo pode ser conectado ao sistema do computador através de sensores, atuadores e dispositivos de entrada e saída (Zhou et al., 2008; Akyildiz et al., 2002). Segundo Ali *et. al* em (Ali et al., 2017) as redes sem fio que permitem um atraso limitado da rede, essencial para a entrega de

¹De acordo com a Organização Mundial da Saúde WHO (*World Health Organization*), idoso é todo indivíduo com 60 anos ou mais.

pacotes, são denominadas redes sem fio de tempo real. Neste contexto, para o acompanhamento em tempo real da saúde das pessoas torna-se evidente o emprego das WBANs. Elas possibilitam o monitoramento contínuo e em tempo real com o objetivo de facilitar médicos e aplicações médicas no diagnóstico (Movassaghi et al., 2014). Assim, através de dispositivos corporais sem fio as pessoas podem receber diagnósticos de forma remota, sem a necessidade da sua presença em hospitais.

Embora as redes corporais sem fio apresentem os requisitos para o monitoramento contínuo e em tempo real das pessoas, elas possuem grandes desafios como heterogeneidade dos sinais e dos dispositivos, tamanho e consumo de energia (Movassaghi et al., 2014). Além disso, essas redes demandam de uma latência máxima de 250ms para aplicações não médicas e 125ms para aplicações médicas (of Electrical e Engineers, 2012). Já o monitoramento em tempo real faz uso excessivo de transmissão de dados, o que consequentemente afeta o consumo de energia dos dispositivos e seu tempo de vida. Logo, há a necessidade de pesquisas que tornem estes dispositivos mais eficientes energeticamente e adaptáveis as aplicações de tempo real nas WBANs.

1.2 PROBLEMA

O processo de coleta e transmissão em uma rede corporal sem fio inicia após o sensor obter um sinal fisiológico (amostra) de uma pessoa. Normalmente, estes sensores são acoplados a um dispositivo corporal que realiza o processamento do sinal fisiológico. Assim, este sinal transforma-se em uma informação (Qu et al., 2019). Após a etapa de processamento, esta informação é transmitida para centrais de monitoramento e profissionais da saúde através de uma interface de comunicação sem fio para futuros diagnósticos (Lay-Ekuakille et al., 2015). Porém, este processo de modo contínuo e em tempo real exige o uso constante dos recursos computacionais existentes nos dispositivos, o que aumenta drasticamente o consumo de energia e, por consequência, diminui seu tempo de vida (Nadeem et al., 2013).

A fim de mitigar o consumo de energia dos dispositivos nas redes corporais sem fio diversos pesquisadores (Khan e Pathan, 2018; Javaid et al., 2015, 2016; Adhikary et al., 2016; Roy et al., 2017) propuseram protocolos de comunicação e assim maximizaram o tempo de vida. Porém, estas soluções não consideram a redundância das informações coletadas e assim, todas as coletas são transmitidas. Dessa maneira, o principal problema abordado nesta dissertação consiste em **como ser mais eficiente na economia de energia de modo a oferecer uma transmissão eficaz, garantir o funcionamento dos serviços de coleta e ampliar o tempo de vida dos dispositivos corporais**. Particularmente, uma abordagem para otimizar a energia dos dispositivos é reduzir a quantidade de transmissão através de técnicas de fusão de dados (Liu et al., 2015). Essas técnicas destacam-se por agregar e/ou diminuir a redundância da informação (Ding et al., 2019).

Neste sentido diferentes técnicas de fusão de dados em WSNs têm sido utilizadas nas WBANs (Uthayakumar et al., 2018), destacando-se a compressão de dados local, que é comumente dividida em duas classes, com perdas e sem perdas (Giorgi, 2017). Na compressão com perdas uma quantidade aceitável da informação é descartada, tal que seja reduzida a quantidade de bits transmitidos, a fim de obter-se maiores taxas de compressão de dados. Em contrapartida, as técnicas de compressão sem perdas mantém a informação inalterada, e usam tabelas de codificação para diminuir a quantidade de dados na transmissão (Marcelloni e Vecchio, 2009). No entanto, as técnicas propostas para compressão com perdas e sem perdas em WSNs e WBANs (Marcelloni e Vecchio, 2009; Schoellhammer et al., 2004; Azar et al., 2018b,a; Arrabi e Lach, 2009), visam reduzir a informação e/ou aumentar a taxa de compressão dos dados, e não

consideram questões como latência máxima permitida para uma WBAN (Movassaghi et al., 2014) e a heterogeneidade dos sinais. Em um ambiente real, perdas ou atrasos de informações podem causar inúmeras consequências no diagnóstico médico (Cavallari et al., 2014). Dessa maneira, torna-se necessário pesquisar novas formas para reduzir o consumo de energia dos dispositivos sem afetar drasticamente a latência máxima permitida de uma WBAN. Assim, esta dissertação busca explorar e responder as seguintes questões de pesquisa atreladas ao problema levantado:

- **Como o emprego de métodos compressão de dados local contribuem para a eficiência energética dos dispositivos corporais nas WBANs?** O emprego de métodos de compressão de dados local contribui para redução da informação na transmissão. Além disso, é possível reduzir a redundância da informação coletada pelos sensores. Logo, a quantidade total de pacotes transmitidos é reduzida e gera-se eficiência energética nos dispositivos.
- **Em que medida o uso de compressão de dados local interfere na transmissão das informações processadas pelos dispositivos corporais em aplicações de tempo real?** As redes corporais sem fio de tempo real possuem uma latência máxima estabelecida conforme seu padrão. O uso de compressão de dados local interfere no pós-processamento do sinal fisiológico coletado pelo sensor. Tendo em vista que toda informação leva um tempo para ser comprimida. Logo, a medida altera-se conforme o método empregado.
- **A aplicação de métodos de compressão de dados local auxilia para ampliar o tempo de vida dos dispositivos corporais numa WBAN?** As abordagens tradicionais transmitem toda informação coletada pelos sensores corporais. O emprego de métodos de compressão de dados local auxilia na redução de pacotes transmitidos. Assim, diminui-se o uso do rádio para transmissão, um dos principais consumidores de energia. Logo, o aumenta-se o tempo de vida dos dispositivos.

1.3 OBJETIVO

Esta dissertação tem como objetivo otimizar o desempenho energético dos dispositivos corporais na transmissão das informações provenientes das aplicações de tempo real nas WBANs de modo a ampliar o tempo de vida dos dispositivos sem afetar drasticamente a latência máxima permitida de uma WBAN. Para atingir este objetivo é proposto um mecanismo de compressão de dados local em tempo real energeticamente eficiente para redes WBANs chamado GROWN (*EnerGy-Efficient Local Data Compression for TRansmission Over WBAN*). Este mecanismo tem como intuito evitar a transmissão de dados redundantes, e assim, reduzir o consumo de energia dos dispositivos sem impactar de forma intensa no atraso total da transmissão da informação para o *sink*. Para isto, o GROWN emprega métodos de compressão de dados local encontrados nas WSNs e limiares definidos. Assim, as informações transmitidas são comprimidas sem perdas por meio de tabela fixa de codificação. Além disso, o mecanismo estabelece limiares pré-definidos e analisa a diferença da correlação entre amostras consecutivas. Dessa forma, apenas valores superiores a este limiar serão comprimidos para transmissão. Logo, foram definidos os seguintes objetivos específicos com a finalidade de atingir o objetivo geral:

- Desenvolver um mecanismo energeticamente eficiente baseado em métodos de compressão de dados local encontrados nas WSNs;
- Implementar o mecanismo proposto em dispositivos corporais sem fio;

- Investigar os métodos de compressão de dados local empregados nas redes WSNs e WBANs e elaborar uma síntese sobre eles com base em comparações e abordagens;
- Analisar os resultados obtidos nas experimentações, suas vantagens e desvantagens, bem como pontos de melhorias.

1.4 CONTRIBUIÇÕES

O desenvolvimento dessa dissertação resultou em contribuições científicas na área de computação, com ênfase em redes sem fio não estruturadas voltada à gerência de desempenho. A seguir, as contribuições deste trabalho estão descritas de forma detalhada:

- *Um levantamento sobre o estado da arte:* O estudo revisa a literatura acerca dos métodos de compressão de dados local das redes de sensores sem fio que podem ser moldados para as redes corporais sem fio, além de métodos de compressão de dados local voltados especificamente para as redes WBANs. Esses métodos foram classificados de acordo com o tipo de rede pesquisado e sua aplicação para o problema de pesquisa;
- *Elaboração do mecanismo GROWN:* Ele baseia-se em métodos encontrados nas WSNs e analisa a diferença da correlação entre amostras consecutivas coletadas por sensores fisiológicos. Assim, a compressão desta diferença ocorre através de uma tabela fixa que não impacta drasticamente no atraso para apresentação da informação. Dessa forma, o mecanismo contribui para a redução no consumo de energia dos dispositivos;
- *Avaliação e análise da eficácia do GROWN:* Uma análise e discussão sobre os resultados obtidos pelo mecanismo para redução no consumo de energia dos dispositivos corporais, bem como sua aplicação em WBANs de tempo real.

1.5 ESTRUTURA DO TEXTO

Esta dissertação de mestrado está organizada em seis capítulos. O Capítulo 2 apresenta os fundamentos relacionados aos conceitos de WSNs, WBANs, Entropia da informação e Fusão de dados, além dos requisitos necessários à compreensão do problema tratado. O Capítulo 3 discute o estado-da-arte dos trabalhos da literatura relacionados ao problema tratado. O Capítulo 4 descreve o mecanismo GROWN para redução no consumo de energia dos dispositivos corporais. O Capítulo 5 apresenta a implementação e avaliação do mecanismo. O Capítulo 6 conclui esta dissertação e apresenta direcionamentos futuros para a pesquisa.

2 FUNDAMENTOS

Este capítulo apresenta os fundamentos necessários à compreensão do contexto da pesquisa, do problema e da proposta. Inicialmente, a Seção 2.1 define as características gerais das WBANs, sua topologia e arquitetura, requisitos de qualidade de serviço, suas aplicações e características. A Seção 2.2 relata o conceito de entropia da informação e sua aplicabilidade em fusão de dados, as terminologias para fusão de dados, bem como o seu uso, suas classificações, técnicas, além de evidenciar o uso da técnica de compressão de dados.

2.1 REDES CORPORAIS SEM FIO (WBANS)

A aplicação das WSNs no sensoriamento de diferentes ambientes e com dimensões distintas tem levado naturalmente ao surgimento de novos tipos de redes de modo a atender melhor cada ambiente e os seus propósitos. Além disso, o avanço nos nanossistemas em geral têm proporcionado o desenvolvimento de dispositivos inteligentes que atuam sobre os antigos e novos ambientes monitorados de modo a colaborar para a construção dos Cyber Physical Systems (CPS). Um dos ambientes que tem recebido muita atenção trata-se do sensoriamento do corpo humano com dispositivos vestíveis ou implantáveis. Particularmente, a comunicação entre estes tipos de dispositivos deu origem as Redes Corporais sem Fio, denominadas de WBANs (do inglês, *Wireless Body Area Networks*). Essas redes têm proporcionado a facilidade de diagnósticos para profissionais de saúde, não interferindo na vida diária das pessoas (Movassaghi et al., 2014).

Uma rede WBAN dispõe de um conjunto de dispositivos com baixo poder de consumo que monitoram o estado de saúde do corpo humano (Ullah et al., 2012). Estes dispositivos são posicionados interno ou externo ao corpo e detêm de tecnologias de comunicação sem fio que operam nas proximidades do corpo (Vergutz et al., 2017). Assim, eles monitoram, coletam e transmitem as informações fisiológicas para aplicações médicas, que por sua vez, transmite estas informações para um profissional de saúde com o intuito de um diagnóstico médico (Ullah et al., 2012). As WBANs e as aplicações disponibilizam um monitoramento contínuo, em tempo real, ubíquo e flexível, de modo a oferecer praticidade e continuidade as atividades cotidianas dos pacientes (Cavallari et al., 2014).

2.1.1 Topologia e arquitetura

A topologia entre os diferentes dispositivos de uma WBAN para a transmissão dos dados acerca dos sinais vitais procede em forma de estrela de um ou dois saltos (of Electrical e Engineers, 2012). De modo geral, a topologia estrela possui um dispositivo central, representado pelo *sink*, muitas vezes colocado na cintura do corpo humano. Quando todos os dispositivos da rede comunicam-se diretamente com o *sink*, a rede é considerada uma topologia estrela de um salto. De outra forma, a topologia estrela de dois saltos faz uso de um segundo dispositivo para a comunicação, comumente posicionado mais próximo ao *sink*. Sendo assim, o *sink* faz o controle de acesso ao meio, a gerência de energia dos demais dispositivos, e transmite os dados coletados para os centros de saúde (Movassaghi et al., 2014).

Assim como seguiu a topologia estrela, dois métodos foram propostos com a finalidade da comunicação entre o *sink* e os demais dispositivos: *beacon* e *sem-beacon*. No modo *beacon*, o dispositivo *sink* controla a comunicação, ou seja, transmite *beacons* periódicos a fim de definir o início e o fim de um superquadro para ativar o controle de associação de rede e a sincronização

dos dispositivos. De modo geral, a duração destes períodos podem ser especificados pelo usuário e baseado no padrão WBAN. No modo *sem-beacon*, um dispositivo da rede é capaz de transmitir dados para o *sink* e pode utilizar o CSMA/CA (*Carrier Sense Multiple Access / Collision Avoidance*) quando necessário. Sendo assim, os dispositivos que estão em modo de espera, precisam ligar e pesquisar o *sink* para receber dados. No entanto, o *sink* não pode se comunicar com os dispositivos em todos os momentos, pois os dispositivos devem esperar até que sejam convidados a participar de uma comunicação (Sukor et al., 2008).

Neste sentido, a arquitetura de comunicação de uma WBAN adere a uma topologia em estrela de um ou dois saltos (Movassaghi et al., 2014). Em uma topologia com apenas um salto, a comunicação com o *sink* possui um menor atraso no envio dos dados, uma vez que a comunicação é direta (of Electrical e Engineers, 2012). Porém, na topologia de dois saltos há uma diminuição na potência de transmissão. Logo, existe um menor consumo de energia dos dispositivos, o que acarreta em um maior tempo de vida da rede (of Electrical e Engineers, 2012; Naganawa et al., 2015). Assim, em muitas situações, exige-se o uso de ambas as topologias em uma única arquitetura. Aplicações em tempo real, por exemplo, dispositivos com alta prioridade voltados a aplicações médicas ficam próximos ao *sink* em uma comunicação direta, enquanto dispositivos voltados a aplicações não médicas fazem uma comunicação através de dois saltos (Javaid et al., 2015).

Por fim, o modelo da arquitetura de comunicação das WBANs pode ser separado em três diferentes camadas, como ilustradas na Figura 2.1. A primeira camada, chamada *intra-body*, expressa a interação da rede entre os dispositivos dentro e ao redor do corpo humano. Em vista disso, os dispositivos monitoram, coletam e transmitem os sinais fisiológicos para o *sink* através de tecnologias sem fio de curto alcance, como Bluetooth e ZigBee (Ullah et al., 2012). Uma vez no *sink*, os dados são então processados e encaminhados para o ponto de acesso via WiFi (Vergutz et al., 2017).

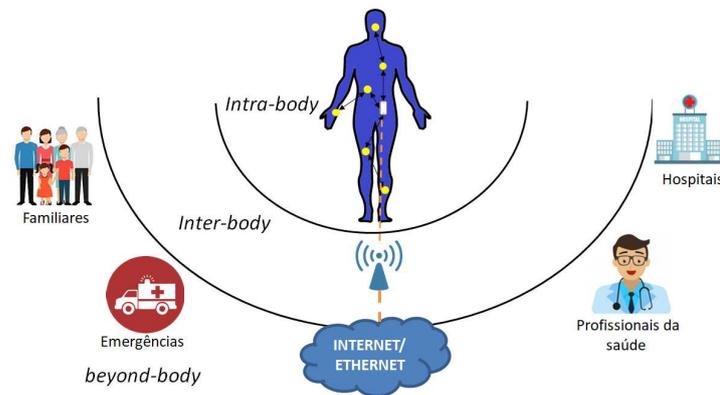


Figura 2.1: Representação dos três níveis de Arquitetura de uma rede WBAN

Denominada de comunicação *inter-body*, esta comunicação ocupa-se entre o dispositivo *sink* e o ponto de acesso AP. Os APs podem ser considerados como parte da infraestrutura, ou até mesmo colocados estrategicamente em um ambiente dinâmico para lidar com situações de emergência. Além disso, a comunicação *inter-body* visa interconectar WBANs com várias redes, as quais podem ser acessadas diariamente, bem como nas redes celulares e na Internet (Chen et al., 2011). Logo, quanto mais tecnologias suportadas por uma WBAN, mais fácil a integração com diversos aplicativos (Movassaghi et al., 2014). Por fim, a terceira camada designada *beyond-body* tem como principal foco áreas metropolitanas. Em vista disso, o dispositivo *sink* pode fazer uma ponte entre as camadas *inter-body* e *beyond-body*, e passar da Internet para um servidor médico em uma aplicação específica (Latré et al., 2011). Em essência, em um ambiente médico, um

banco de dados é um dos componentes mais importantes da *beyond-body*, pois inclui o histórico médico e o perfil do usuário. Assim, os médicos ou pacientes podem ser notificados de um estado de emergência através da Internet ou de uma mensagem de texto (SMS). Além disso, a *beyond-body* permite restaurar todas as informações necessárias de um paciente, as quais podem ser usadas para um possível tratamento (Ullah et al., 2012).

2.1.2 O Padrão 802.15.6

As WBANs foram desenvolvidas para atender as exigências das aplicações médicas e da comunicação dos dispositivos. Deste modo, requisitos como baixo consumo de energia, mínimas interferências, baixa taxa de dados, confiabilidade e Qualidade de Serviço (QoS) devem ser suportados a fim de abordar extensivamente as aplicações médicas (Vergutz et al., 2017). No entanto, as WPANs efetuam a conexão entre diferentes dispositivos próximos ao corpo humano para a troca de informações. Porém, elas não satisfazem os requisitos de comunicação médica, devido à proximidade do tecido do corpo humano. Logo, foi necessário um modelo padrão para a implementação de uma WBAN, a fim de abordar tanto dispositivos eletrônicos quanto as aplicações médicas (Movassaghi et al., 2014).

De modo a definir um padrão específico para as redes WBANs, em maio de 2008 um grupo de tarefas IEEE (*Task Group 802.15.6 - TG6*) foi criado (Vergutz et al., 2017). Denominado IEEE 802.15.6, o padrão teve seu primeiro esboço em abril de 2010, otimizado para dispositivos com poucos recursos de energia, interno ou externo ao corpo humano, para aplicações médicas e não médicas (Kwak et al., 2010). Posteriormente teve sua última versão ratificada em fevereiro de 2012, bem como definiu seu objetivo, desenvolver um padrão de comunicação para dispositivos com poucos recursos de energia e operação no corpo humano ou ao seu redor, mas não limitado apenas para humanos, a fim de atender a uma variedade de aplicações, médicas, eletrônicas de consumo, entretenimento pessoal e outros (of Electrical e Engineers, 2012). Logo, WBAN é considerada uma rede de pequena escala, com restrições de energia e curto alcance de transmissão, com poucos metros de distância para realizar a comunicação entre os dispositivos (Vergutz et al., 2017).

Assim como a definição dos objetivos na última versão ratificada, uma série de requisitos foram estabelecidos pelo TG6 conforme ilustrada na Tabela 2.1. Deste modo, as WBANs podem suportar no máximo 256 dispositivos, sendo estes, adicionados ou removidos em menos de 3 segundos. Além disso, os dispositivos devem possuir uma comunicação confiável, mesmo quando existe movimento corporal. Embora seja aceitável que a capacidade da rede seja reduzida, os dados não devem ser perdidos devido a condições instáveis do canal de comunicação. Ademais, a taxa de dados deve suportar uma velocidade de até 10Mbps, e a latência deve ser menor que 125ms para aplicações médicas e 250ms para aplicações não médicas. Por fim, as redes WBANs devem ser capazes de operar em ambientes heterogêneos, onde diferentes tecnologias de rede cooperam entre si com o intuito de receber informações (Movassaghi et al., 2014).

Um requisito importante nas WBANs consiste no baixo consumo de energia (Vergutz et al., 2017). Um dos fatores que agravam esta condição se dá pelo tamanho dos dispositivos, comumente pequenos para não atrapalhar a vida diária dos pacientes, os dispositivos são projetados para serem implantáveis ou vestíveis, o que diminui drasticamente o tamanho da bateria (Antonescu e Basagni, 2013). Dessa forma, é esperado que a vida útil de um dispositivo dure por anos, especialmente para dispositivos com aplicação em tempo real e dispositivos implantáveis, onde nem sempre é possível a troca ou a recarga da bateria. Além disso, o consumo de energia dos dispositivos acontece na coleta, no processamento e na transmissão dos dados. Por consequência, os dispositivos precisam ser extremamente cautelosos quanto ao seu consumo (Bachir et al., 2010).

Tabela 2.1: Padrão IEEE 802.15.6

Requisitos	Descrição
Topologia	Estrela de um ou dois saltos bidirecional
Tempo de Configuração	Inserção/Remoção < 3s
Núm. de Dispositivos	Escalável até 256 nós
Taxa de dados	10Kb/s - 10Mbps
Alcance	≈3m
Latência	< 125ms (médica); < 250ms (não médica)
Jitter	50ms

2.1.3 Aplicações de uma WBAN

Uma WBAN pode ser aplicada em setores como, saúde, entretenimento, esportes, militar entre outros (Crosby et al., 2012). Isto corre, devido aos sensores corporais coletarem diferentes tipos de sinais, como movimentos corporais, características do ambiente e sinais vitais (Movassaghi et al., 2014). Deste modo, segundo o padrão que rege as WBANs (of Electrical and Engineers, 2012) suas aplicações podem ser categorizadas como médica ou não médica. As aplicações médicas são comumente subdivididas em três categorias, vestíveis, implantadas e controle de dispositivos médicos (Zhou, 2017), conforme ilustra a Tabela 2.2. Uma WBAN vestível é capaz de monitorar distúrbios do sono, asma, avaliar a fadiga do soldado e prontidão de batalha, auxiliar o treinamento esportivo para evitar lesões. Para as WBANs implantadas, os dispositivos são implantados abaixo da superfície da pele ou residem na corrente sanguínea. As funções como controle de diabetes, monitoramento de doenças cardiovasculares e detecção de câncer podem ser realizadas por WBANs implantadas. Em controle de dispositivos médicos, cada WBAN troca informações com uma central médica, a fim de permitir atendimento automático aos pacientes por meio das WBANs e liberar pacientes de cuidados pessoais intensivos (Lipprandt et al., 2009).

As aplicações não médicas podem envolver *streaming* de vídeo, como capturas de vídeos por um celular, feiras comerciais de artigos esportivos e lançamentos de designs de moda, transmissões de áudio através de comunicação de voz, como em explicações de arte no museu, programação de ônibus, multicast para chamadas em conferência entre outros. Além disso, aplicações não médicas consistem de categorias para aplicações de entretenimento como jogos com realidade aumentada e redes sociais. Microfones, câmeras, podem ser usados como dispositivos integrados nas WBANs. Ademais emergências não médicas como incêndio, ou vazamento de gás podem ser comunicadas ao usuário através de dispositivos corporais (Lewis, 2008).

Indiferente do tipo da aplicação, médica ou não médica, as WBANs dispõem de três tipos de dispositivos: corporal, atuadores e dispositivos pessoais, cada qual com suas funcionalidades (Movassaghi et al., 2014). Os dispositivos corporais fazem o monitoramento, a coleta e a transmissão de parâmetros corporais como, batimentos cardíacos e temperatura, tanto internamente como externamente. Alguns tipos existentes desses dispositivos podem ser usados em relógios de pulso, celulares ou fones de ouvido e, conseqüentemente, permitir o monitoramento sem fio de uma pessoa em qualquer lugar, a qualquer momento e com qualquer pessoa. Por sua vez, os atuadores interagem com o usuário ao receber dados dos dispositivos corporais. Sua função é fornecer *feedback* na rede e agir sobre dados de sensores, por exemplo, injetar a dose correta de medicamento para o corpo em aplicações médicas. Por fim, os

Tabela 2.2: Aplicações de uma WBAN

Aplicação	WBAN	Descrição
Médica	Vestíveis	Fadiga de soldado e prontidão para batalha Treinamento para esportistas Distúrbio do sono Asma Monitoramento de dispositivos vestíveis
	Implantáveis	Doenças cardiovasculares Detecção de câncer
	Controle de Dispositivos	Monitoramento de pacientes Sistemas de tele-medicina
Não Médica		<i>Streaming</i> Aplicações de entretenimento Emergências não médicas

dispositivos pessoais são responsáveis por coletar todas as informações recebidas dos dispositivos corporais e atuadores e transmitir para o usuário. Em aplicações médicas por exemplo, os dispositivos pessoais transmitem os sinais fisiológicos, como batimento cardíaco ou quantidade de insulina injetada para os profissionais da saúde e pacientes (Hanson et al., 2009). Os dispositivos pessoais são trivialmente denominados nó *sink*, nó coordenador, nó *gateway*, ou ainda, Unidade de Controle do Corpo (BCU). Neste trabalho, o nome nó *sink* será utilizado.

Segundo Yuce em (Yuce, 2010) as redes WBANs fornecem uma visão das condições fisiológicas de um ser humano. Assim, elas monitoram as condições de saúde e enviam *feedbacks* ao usuário final ou aos médicos por meio de conectividade sem fio. Através dos dispositivos corporais, elas são usadas para monitorar parâmetros fisiológicos, como batimento cardíaco, temperatura, nível de estresse, nível de oxigênio entre outros. Além disso, elas monitoram e analisam a rotina de exercícios e a exigência do corpo humano através de aplicações de tempo real.

Aplicações de tempo real são capazes de monitorar, responder imediatamente à entrada do usuário ou controlar um ambiente externo. O ambiente externo pode ser conectado ao sistema do computador através de dispositivos corporais, atuadores e dispositivos de entrada e saída (Zhou et al., 2008; Akyildiz et al., 2002). Segundo Ali *et. al* em (Ali et al., 2017) as redes sem fio que permitem um atraso limitado da rede, essencial para a entrega de pacotes, são denominadas redes sem fio de tempo real. No contexto das WBANs, elas seguem os padrões de latência estabelecidos pelo TG6 que estabelece o máximo de 125ms para aplicações médicas e 250ms para aplicações não médicas (Movassaghi et al., 2014).

Asif e Sumra em (Asif e Sumra, 2017) afirmam que as características que envolvem estas aplicações dependem necessariamente da sua finalidade. De modo geral, estas aplicações estabelecem uma comunicação direta com o *sink* para evitar a latência de mais um salto e realizam a coleta de diferentes tipos sinais. Segundo Ullah *et. al* em (Ullah et al., 2017) estes sinais são definidos como contínuos ou discretos, de tal forma que, os sinais contínuos como (ECG, EEG, EMG) necessitam de transmissões contantes para exibir as informações, e os sinais discretos como (Glicose, Temperatura, umidade) podem encaminhar suas informações de tempos em tempos. Além disso, uma característica comum a estas aplicações é o alto consumo de energia dos dispositivos que utilizam o rádio transmissor constantemente. A Tabela 2.3 apresenta diferentes sinais corporais de uma WBAN de tempo real.

Tabela 2.3: Sinais corporais de uma WBAN e suas características

Sinal	Categoria	Taxa de dados	Definição do sinal	Topologia
Pressão Arterial	Vestível	1000 bps	Discreto	Estrela
EEG/ECG/EMG		≈192 kbps	Contínuo	
Glicose		1600 bps	Discreto	
Temperatura		120 bps	Contínuo	
Retina artificial	Implantável	50-700 kbps	Contínuo	
Cóclea artificial		100 kbps	Contínuo	
Pílula microcâmera		1 Mbps	Contínuo	

2.2 FUSÃO DE DADOS

Esta seção descreve os principais conceitos sobre fusão de dados. Primeiramente, uma sucinta explanação sobre entropia da informação é exposta. Esta explanação corrobora o uso da entropia para arquitetar a fusão de dados. Em seguida, são conceituadas as diferentes terminologias para fusão de dados e suas diferentes classificações. Por fim, é dissertado sobre a abordagem de compressão de dados local e sua relevância para as redes WBANs e WSNs.

2.2.1 Entropia da informação

A Teoria da Informação, proposta em 1948 por Claude E. Shannon (Shannon, 1948), está relacionada em diversas áreas de conhecimento como Criptoanálise e Criptografia, Inteligência Artificial, Teoria de Probabilidade, Teoria da Comunicação entre outras (Almeida-Santos et al., 2014). Na época, a finalidade de Shannon era quantificar a informação que transita por um sistema composto de transmissor, receptor e canal (Campani, 2005). Assim, a Teoria da Informação visa estudar questões teóricas a fim de codificar mensagens otimamente conforme suas estruturas estatísticas, como selecionar taxas de transmissão de acordo com o nível de ruído no canal e avaliar o mínimo de distorção nas mensagens. Logo, uma das formas de analisar tais estatísticas consiste na entropia da informação (Araújo, 2013).

Inicialmente no contexto da termodinâmica, o conceito de entropia tende a fornecer uma medida de irreversibilidade em um sistema físico. Posteriormente, com a inserção de métodos estatísticos, o conceito de entropia estatística visa fornecer uma quantidade do grau de desordem de um determinado sistema. Na Teoria da Informação, o conceito de entropia provê a quantidade de incerteza de uma determinada variável aleatória. Assim, denominada entropia da informação ou entropia de Shannon (Nascimento e Prudente, 2016).

Deste modo, suponha que exista uma determinada variável aleatória X , onde a esta variável possam ser atribuídos valores discretos $x \in X$ e com distribuição de probabilidade $p(x)$. A entropia da informação ou entropia de Shannon $H(X)$ da variável aleatória pode ser dada pela Equação 2.1. Na qual, $p(x)$ indica a probabilidade de evento da distribuição de probabilidade da variável aleatória discreta X . Por definição, o valor da entropia de $H(X)$ será sempre maior ou igual a zero, assim como sua unidade de medida é dada em *bits* se a base logarítmica for igual a 2 e em *nats* se a base do logaritmo for igual ao número de Euler e (Nali, 2012).

$$H(X) = - \sum_{x \in X} p(x) \log_2 p(x) \quad (2.1)$$

A entropia da informação $H(X)$ caracteriza-se como a medida da quantidade média de informação transmitida por determinada mensagem. Uma propriedade acerca desta medida se da pela combinação das incertezas serem ponderadas por suas probabilidades, o que determina o conceito chave da entropia (Araújo, 2013). Deste modo, eventos atípicos, que dispõem de uma grande quantidade de informação, possuem maior o grau de entropia. Por outro lado, eventos corriqueiros, que apresentam pouco conteúdo de informação, apresentam um menor grau de entropia (WEAVER e SHANNON, 1971).

A aplicabilidade da entropia da informação pode ser encontrada em diversas áreas como, processamento de imagens, processamento de sinais, aplicações genéticas, entre outras (Raja et al., 2019; Li et al., 2018; Farach et al., 1995). Além disso, a utilização da entropia da informação possui grande correlação com a fusão de dados, onde diversas técnicas fazem o seu uso para diminuir a redundância da informação com a análise do valor da entropia (Chen et al., 2018). Deste modo, a utilização do valor de entropia em conjunto com a fusão de dados pode oferecer diversas melhorias. Nas WSNs, por exemplo, diminuir a redundância da informação e utilizar técnicas de fusão de dados podem reduzir significativamente o consumo de energia nos dispositivos (Marcelloni e Vecchio, 2009).

2.2.2 Terminologias da fusão de dados

A grande evolução da internet nos últimos anos tornou possível e útil o acesso a múltiplas informações (Bleiholder e Naumann, 2009). Contudo, acessar diferentes fontes para se obter uma única informação pode se tornar um trabalho custoso, exigindo processamento e grandes recursos de energia (Marcelloni e Vecchio, 2009). A fusão de dados é o processo que visa unificar várias fontes de informação para produzir informações mais consistentes, precisas e úteis (Haghighat et al., 2016). Embora o conceito de fusão de dados ser de fácil entendimento, seu exato significado é amplamente diversificado. Diversos autores utilizam termos, como combinação, sinergia, integração, mesclagem, entre outros. Além disso, em muitas situações a palavra fusão é utilizada, enquanto classificação seria o termo mais apropriado (Wald, 1999).

De modo geral, muitas definições para fusão de dados foram estabelecidas ao longo dos anos. Franklin e White em (White, 1991) definiram fusão de dados como um processo multinível que lida com a detecção automática, associação, correlação, estimativa, combinação de dados e informações de múltiplas fontes. Klein em (Klein, 1999) generalizou esta definição dizendo que a fusão de dados pode ser obtida através de uma ou múltiplas fontes de informação. De mesmo modo, Hall e Llinas em (Hall e Llinas, 1997) elucidaram que a fusão de dados é a combinação de dados de múltiplos dispositivos, e informações relacionadas fornecidas por bancos de dados associados, com o intuito de obter melhor precisão e inferências mais específicas que não poderiam ser alcançadas pelo uso de um único dispositivo. Contudo, neste contexto a fusão de dados é realizada com um único objetivo, a melhoria da acurácia, e sua definição se restringe apenas ao uso de sensores, sem prover o uso de dados de uma única fonte.

Esta confusão de termos e conceitos foi altamente discutida por Dasarathy em (Dasarathy, 2001), o qual adaptou o termo como fusão da informação. Dasarathy afirmou que no contexto de seu uso na sociedade, engloba a teoria, técnicas e ferramentas criadas e aplicadas para explorar a sinergia na informação adquirida de múltiplas fontes (banco de dados, sensores, informações coletadas por humano, entre outras) de maneira pela qual a decisão ou ação resultante é, em certo sentido, melhor (qualitativa ou quantitativamente, em termos de precisão, robustez, etc) do que seria possível se qualquer uma dessas fontes fosse usada individualmente sem tal exploração de sinergia (Dasarathy, 2001). Possivelmente, esta é a definição mais ampla que abrange qualquer tipo de fonte, conhecimento e recursos usado para fundir diferentes informações. Logo, o termo fusão da informação e a definição de Dasarathy em (Dasarathy, 2001) foi adotado pela Sociedade

Internacional de Fusão da Informação (*International Society of Information Fusion*). No entanto, Castanedo em (Castanedo, 2013) esclareceu que os termos fusão da informação e fusão de dados são tipicamente empregados como sinônimos, de modo que, em alguns cenários o termo fusão de dados é utilizado para dados obtidos diretamente dos dispositivos (*raw data*) e o termo fusão da informação é empregado para definir dados já processados. Neste sentido, o termo fusão da informação implica um nível semântico mais elevado do que a fusão de dados.

Apesar da ambiguidade entre diversos conceitos e termos, a fusão de dados detém objetivos bem definidos. Sendo possível destacar, a qualidade dos dados, de modo a reduzir potencialmente as incertezas associadas às medições, aumentar a confiabilidade, através de diversas fontes de informação (Villanueva, 2009) e diminuir a redundância da informação e economizar energia, pelo meio de técnicas de entropia, compressão e agregação de dados (Marcelloni e Vecchio, 2009). Sua aplicabilidade compreende diversas áreas como, redes de sensores sem fio (WSNs), redes corporais sem fio (WBANs), bioinformática, oceanografia, sistemas inteligentes de transportes, entre outras. Nesta perspectiva, a complexidade da classificação de fusão de dados é evidente, visto que, cada área detém de regras e condições específicas.

2.2.3 Classificação

A fusão de dados corresponde uma área multidisciplinar, de modo que, sua aplicabilidade abrange diversos campos de conhecimento. Conforme este contexto, sua complexidade de classificação é evidente, visto que, cada campo de conhecimento detém de regras e condições específicas. Logo, é difícil estabelecer uma classificação clara e rigorosa. Deste modo, ao longo do tempo, diversas classificações para fusão de dados foram desenvolvidas, entre elas, classificação de acordo com o tipo de arquitetura, entre diferentes entradas de informação e de acordo com o nível de abstração da informação (Castanedo, 2013).

Classificação de acordo com o tipo de arquitetura leva em consideração onde o processo de fusão de dados será executado. Com base neste critério, os tipos de arquitetura podem ser identificados conforme a Figura 2.2, onde (a) representa a arquitetura centralizada, (b) ilustra a arquitetura descentralizada e (c) especifica a arquitetura distribuída. Na arquitetura centralizada, todos os dispositivos efetuam apenas as coletas das medições, a responsabilidade pela operação de fusão de dados é realizado por uma unidade central.

Uma arquitetura descentralizada é composta de uma rede de dispositivos, na qual, cada dispositivo possui suas próprias capacidades de processamento e não existe um ponto único de fusão de dados. Portanto, cada dispositivo funde suas informações locais com as informações recebidas de seus pares. Além disso, a fusão de dados é realizada de forma autônoma, e cada dispositivo responde por suas informações locais e as informações recebidas de seus pares (Durrant-Whyte et al., 2001).

Em uma arquitetura distribuída, as medidas de cada dispositivo de origem são processadas de modo independente, ou seja, antes que as informações sejam enviadas para o dispositivo que efetua a fusão. Deste modo, o dispositivo que realiza a fusão não é responsável pelas informações recebidas dos outros dispositivos. Logo, a associação dos dados e a estimativa de estado ocorrem no dispositivo de origem antes que as informações sejam comunicadas ao dispositivo de fusão. Portanto, cada dispositivo fornece uma estimativa do estado do objeto com base apenas em suas visualizações locais, e essa informação é a entrada para o processo de fusão. Esse tipo de arquitetura fornece diferentes opções e variam de apenas um dispositivo de fusão a vários dispositivos de fusão intermediários (Castanedo, 2013). Em princípio, uma arquitetura descentralizada de fusão de dados é mais difícil de implementar devido aos requisitos de computação e comunicação. Contudo, na prática, não há uma arquitetura única, ou melhor, e a seleção da arquitetura mais apropriada deve ser feita de acordo com os requisitos, demanda,

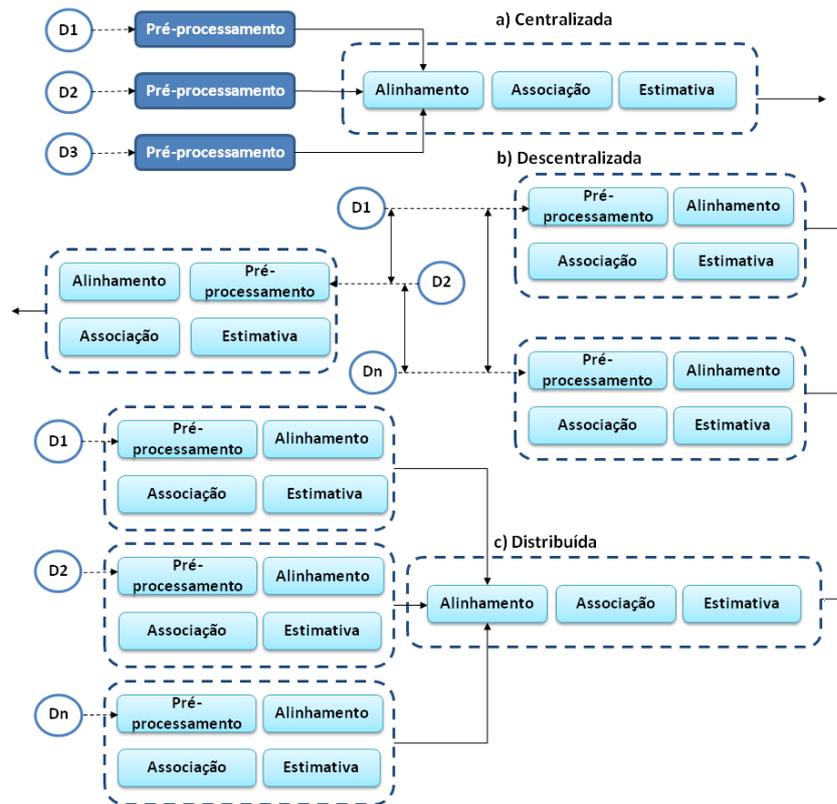


Figura 2.2: Classificação de acordo com o tipo de Arquitetura segundo Castanedo (Castanedo, 2013)

redes existentes, disponibilidade de dados, capacidade de processamento dos dispositivos e organização do sistema de fusão de dados.

Com relação entre diferentes entradas de informação, Durrant em (Durrant-Whyte, 1988) classificou as técnicas de fusão de dados em três diferentes critérios, complementar, redundante e dados cooperativos. No que diz respeito a técnica de fusão de dados complementar, Durrant afirmou que, são as informações fornecidas pelos dispositivos, e que representam diferentes partes do cenário e, portanto, poderiam ser usadas para obter informações globais mais completas. Por exemplo, redes de sensores visuais, onde as informações sobre o mesmo alvo são fornecidas por duas câmeras com diferentes campos de visão. Em redundante, dois ou mais dispositivos fornecem informações sobre o mesmo dado, deste modo podem ser fundidas para aumentar a confiança. Neste caso, em uma rede de sensores visuais os dados provenientes de áreas sobrepostas podem ser considerados redundantes. Por fim, na classificação cooperativa, as informações são combinadas em novas informações, que são tipicamente mais complexas do que as informações originais. A Figura 2.3 ilustra as três diferentes classificações definidas por (Durrant-Whyte, 1988), onde A, B e C representam as informações e $\{D1, D2, \dots, D5\}$ demonstram os dispositivos.

De modo semelhante, Dasarathy em (Dasarathy, 1997) classificou fusão de dados de acordo com o nível de abstração da informação de entrada e saída (*input - output*). Como consequência, cinco diferentes categorias foram criadas, conforme detalhadas na Tabela 2.4. A primeira categoria diz respeito aos dados de entrada e saída. Neste nível, a fusão de dados é realizada imediatamente após os dados serem coletados dos dispositivos (*raw data*). Deste modo, os resultados são tipicamente mais confiáveis e precisos. Exemplos de algoritmos empregados neste nível são processamento de sinal e imagem. Em seguida, a segunda categoria utiliza a fusão de dados brutos (*raw data*) de diversos dispositivos para extrair recursos ou atributos que

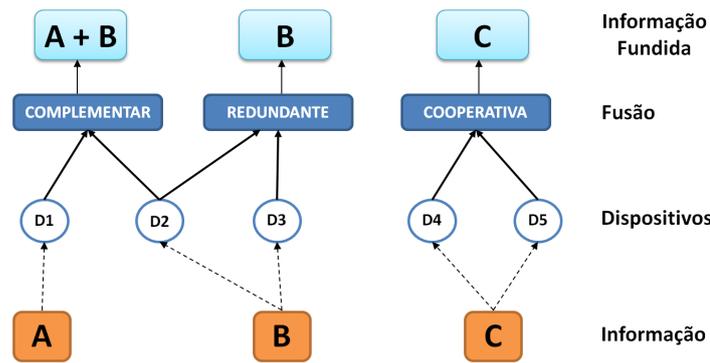


Figura 2.3: Classificação entre diferentes entradas de informação definida por Durrant (Durrant-Whyte, 1988)

descrevem uma entidade. Neste caso, entidade pode significar qualquer objeto, situação ou abstração. Na próxima categoria, tanto a entrada quanto a saída do processo de fusão de dados são recursos. Assim, o processo aborda um conjunto de recursos para melhorar, refinar ou obter novos recursos.

Tabela 2.4: Categorias de acordo com o nível de abstração determinadas por Dasarathy (Dasarathy, 1997)

Categoria	Nome	Descrição
1	DAI-DAO	Data In - Data Out
2	DAI-FEO	Data In - Feature Out
3	FEI-FEO	Feature In - Feature - Out
4	FEI-DEO	Feature In - Decision Out
5	DEI-DEO	Decision In - Decision Out

A quarta categoria obtém um conjunto de recursos como entrada e fornece um conjunto de decisões como saída. Nesta categoria é possível enquadrar a maioria dos sistemas de classificação que executam uma decisão com base nas informações dos dispositivos. Por conseguinte, a quinta categoria é altamente classificada como fusão de decisão, onde funde decisões de entrada para obter melhores ou novas decisões. Ademais, a principal contribuição da classificação de (Dasarathy, 1997) pode ser vista como a especificação do nível de abstração como entrada ou saída, de forma a fornecer uma estrutura para classificar diferentes métodos ou técnicas.

Conforme as categorias definidas por Dasarathy em (Dasarathy, 1997), onde os processos de fusão de dados são categorizados com base no nível de abstração da informação de entrada e saída, diversas técnicas de fusão de dados foram desenvolvidas. Essas técnicas são aplicadas em áreas como processamento de texto, processamento de sinais, processamento de imagens, entre outros (Castanedo, 2013). Comumente, a categoria 1 faz o processo de fusão de dados logo após a coleta do dispositivo, neste caso, uma fusão de dados de uma única fonte (Klein, 1999). Logo, a fusão de dados local ou fusão de dados de uma única fonte pode favorecer na redução da redundância da informação e na diminuição do consumo de energia dos dispositivos através de técnicas de compressão de dados local.

2.2.4 Compressão de dados local

A Compressão de Dados Local consiste de, no contexto da ciência da computação, uma técnica para representar informações de forma compacta. Ela tem sido uma capacitadora crítica para a revolução contínua da multimídia digital por décadas, de modo que, ela é utilizada

frequentemente por softwares de compactação de dados, como `zip`, `gzip`, `WinZip` entre outros, para reduzir o tamanho do arquivo antes de armazená-lo ou transferi-lo para a mídia. Deste modo, a compressão de dados local tornou-se um requisito comum para a maioria dos softwares aplicativos, bem como uma área de pesquisa importante e ativa em ciência da computação. Deste modo, diversas soluções como TV digital, comunicação móvel, técnicas de comunicação de vídeo tiveram ampla utilização de compressão de dados local para sua evolução (Pu, 2005).

Segundo Srisooksai em (Srisooksai et al., 2012) nas WSNs a compressão de dados é realizada localmente em cada dispositivo sem a colaboração distribuída entre outros dispositivos sensores. Deste modo, a compressão de dados local explora apenas a correlação temporal dos dados e não depende de topologias específicas. Salomon em (Salomon, 2004) afirma que existem muitos métodos para compressão de dados local, todos com finalidades e ideias distintas, para diferentes tipos de dados e que produzem resultados diferentes. Contudo, todos os métodos são baseados no mesmo princípio, compactar e remover a redundância da informação original no dispositivo de origem. Portanto, para tal princípio, faz-se necessário uma análise de toda informação coletada pelo dispositivo, com o objetivo de verificar a sua não aleatoriedade. Qualquer informação não aleatória possui alguma estrutura, e essa estrutura pode ser explorada para obter uma menor representação da informação.

Ainda assim, Sayood em (Sayood, 2017) informou que a estrutura da informação não é a única propriedade que pode ser explorada para obter compressão. Segundo ele, também é possível fazer o uso das características do usuário da informação. Muitas vezes, por exemplo, ao transmitir ou armazenar áudio e imagens, as informações devem ser percebidas por um humano, e os humanos têm habilidades perceptivas limitadas. Assim, humanos não conseguem ouvir os sons de frequências muito altas, se algo é representado em informações que não podem ser percebidos pelo usuário, é viável fazer uso de limitações perceptivas dos seres humanos para obter a compressão, deste modo, informações irrelevantes podem ser descartadas.

Nessa conformidade, quando um método de compressão de dados é utilizado, ele exige operações de compressão e descompressão da informação, como ilustra a Figura 2.4. Ele é comumente dividido em dois algoritmos. O primeiro algoritmo de compressão recebe uma entrada X e gera uma representação X_c que requer menos *bits*. De outro modo, o segundo algoritmo realiza a descompressão, o qual recebe a representação comprimida X_c para gerar a descompressão Y . Com base nos requisitos de descompressão, os algoritmos de compressão de dados local podem ser divididos em duas classes, compressão com perdas (*lossy compression*), a qual geralmente oferece uma maior taxa de compressão, mas permite que Y seja diferente de X , e compressão sem perdas (*lossless compression*) em que Y é idêntico a X .

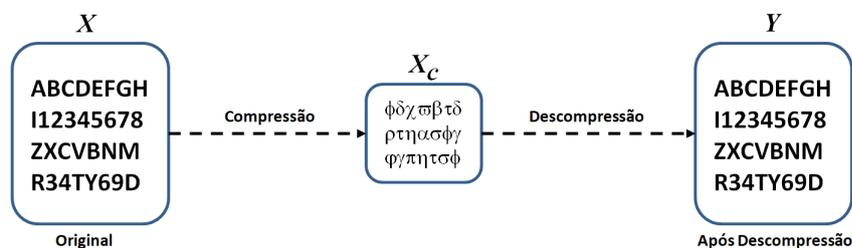


Figura 2.4: Representação de compressão e descompressão segundo Sayood (Sayood, 2017)

Os métodos de compressão sem perdas, como o próprio nome indica, não envolvem perda de informação. Se as informações foram compactadas sem perdas, esta será recuperada exatamente igual após a descompressão. A compressão sem perdas é comumente utilizada para aplicativos que não podem tolerar qualquer diferença entre as informações originais e a descompressão.

Logo, compressão de texto é uma área importante para a compressão sem perdas, onde diferenças muito pequenas podem resultar em declarações com significados diferentes.

Em vista disso, se informações de qualquer tipo forem processadas ou aprimoradas para gerar mais informações, é importante que a integridade seja preservada. Por exemplo, considere a compressão de uma imagem radiológica de maneira deficitária, e a diferença entre a descompressão Y e a informação X for visualmente indetectável. Se esta imagem foi posteriormente aprimorada, as diferenças anteriormente indetectáveis podem causar o aparecimento de artefatos. Logo, estas diferenças podem induzir imprecisão para radiologista. Assim, como o custo desta adversidade pode propiciar equívocos, e afetar drasticamente a vida de um paciente, é preciso ter muito cuidado ao usar compressões que geram uma descompressão diferente da original (Sayood, 2017). A Figura 2.5 ilustra a entrada de um texto e sua compressão, logo após, sua descompressão sem perdas.

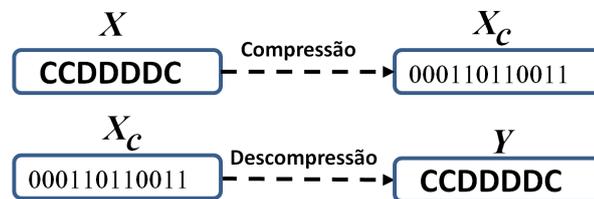


Figura 2.5: Representação de Compressão de dados local sem perdas definida por Pu (Pu, 2005)

De modo contrário, métodos de compressão de dados local com perdas envolvem alguma perda de informação. Estes métodos de compressão geralmente não podem recuperar ou descomprimir a informação exatamente. Porém, em troca de aceitar essa distorção na descompressão, é possível obter taxas de compressão mais elevadas em comparação com os métodos de compressão sem perdas. Em diversas aplicações, essa falta de descompressão exata não é um problema. Por exemplo, ao armazenar ou transmitir fala, a informação exata de cada amostra da fala não é necessária. De acordo com a qualidade exigida, quantidades de variáveis de perda de informação sobre o valor de cada amostra podem ser toleradas. Se a qualidade da descompressão da fala é semelhante ao que ouviu no telefone, uma perda significativa da informação pode ser tolerada. Contudo, se a descompressão precisa ser de qualidade, como em uma música, a quantidade de perda de informação a ser tolerada é muito menor. Da mesma forma, ao visualizar uma descompressão de uma sequência de vídeo, o fato da descompressão ser diferente da original geralmente não é importante, desde que as diferenças não resultem em elementos distorcidos. Assim, o vídeo geralmente é compactado com métodos de compressão com perdas (Sayood, 2017). A Figura 2.6 demonstra a entrada de um número decimal e sua compressão, em seguida, sua descompressão com perdas.

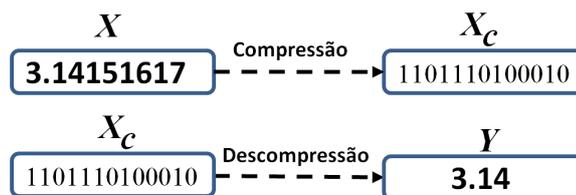


Figura 2.6: Representação de compressão de dados local com perdas de acordo com Pu (Pu, 2005)

Embora os requisitos de descompressão possam forçar a decisão sobre o método ser com ou sem perdas, a decisão exata de qual método utilizar depende de diferentes fatores. Um dos fatores mais importantes trata da característica da informação a ser comprimida. Por

exemplo, um método de compressão para texto pode não funcionar bem para vídeo. Deste modo, cada aplicação apresenta um conjunto de diferentes desafios. Contudo, Salomon em (Salomon, 2004) confirmou que, a melhor abordagem para uma aplicação específica, em grande parte, é a redundância da informação. Assim, métodos de compressão de dados consideram reduzir a redundância da informação através de duas fases. A primeira fase, referida como modelagem, é responsável por extrair informações sobre qualquer redundância existente e descrever a redundância em forma de um modelo. A segunda fase é denominada codificação, nesta fase, uma descrição do modelo e uma descrição de como as informações diferem do modelo são codificadas, onde, em muitos métodos um alfabeto binário ou um alfabeto de símbolos com determinadas probabilidades é utilizado (Sayood, 2017).

Um método comumente usado para a compressão de dados local sem perdas que faz uso de símbolos com determinadas probabilidades trata da codificação de Huffman (Huffman, 1952). Este método cria uma lista de todos os símbolos do alfabeto conforme suas probabilidades. Em seguida, constrói uma árvore com um símbolo em cada folha, de baixo para cima. Esse processo ocorre em etapas, onde em cada etapa os dois símbolos com as menores probabilidades são selecionados, adicionados ao topo da árvore parcial, excluídos da lista e substituídos por um símbolo auxiliar representado por ambos. Ao final do processo, a lista é reduzida para apenas um símbolo auxiliar, o que corresponde a árvore estar completa. Deste modo, a árvore é então atravessada para determinar os códigos dos símbolos (Salomon, 2004). A Figura 2.7a ilustra um exemplo de codificação de Huffman. Dado cinco símbolos com probabilidades conforme determina a Figura 2.7, eles são pareados na seguinte ordem:

1. a_4 é combinado com a_5 e ambos são substituídos pelo símbolo combinado a_{45} com 0.2.
2. Deste modo, existem quatro símbolos, a_1 com probabilidade 0.4, e a_2 , a_3 e a_{45} com probabilidades de 0.2 cada. Arbitrariamente a_3 e a_{45} foram combinados e substituídos com o símbolo auxiliar a_{345} com probabilidade de 0.4.
3. Nesta etapa existem três símbolos, a_1 , a_2 e a_{345} , com probabilidades 0.4, 0.2 e 0.4. Assim como a etapa anterior, a_2 e a_{345} foram combinados e substituídos pelo símbolo a_{2345} com probabilidade 0.6
4. Por fim, os dois símbolos restantes a_1 e a_{2345} são substituídos por a_{12345} com probabilidade 1.

Deste modo, a árvore é disponibilizada conforme a Figura 2.7a, com a raiz à direita e as cinco folhas à esquerda. Para atribuir os códigos, define-se como *bit* 1 à borda superior e 0 à borda inferior de cada par de arestas. Portanto, o resultado para os códigos foram (0, 10, 111, 1101, 1100). O tamanho médio da codificação de Huffman para esta árvore pode ser calculado com a seguinte expressão: $0.4 \times 1 + 0.2 \times 2 + 0.2 \times 3 + 0.1 \times 4 + 0.1 \times 4 = 2.2$ bits/símbolo. Contudo, algumas etapas para a construção da árvore ocorreram de modo arbitrário, uma vez que havia mais de dois símbolos com menores probabilidades. Sendo assim, a Figura 2.7b ilustra os mesmos cinco símbolos com uma combinação de forma diferente, a qual gerou o resultado (11, 01, 00, 101, 100), com média de codificação $0.4 \times 2 + 0.2 \times 2 + 0.2 \times 2 + 0.1 \times 3 + 0.1 \times 3 = 2.2$ bits/símbolo, a mesma média que a árvore anterior.

A ideia básica da codificação de Huffman é mapear o alfabeto, e atribuir menos *bits* para os símbolos que ocorrem com maior frequência. De modo contrário, mais *bits* são atribuídos para símbolos com menor repetição. A codificação de Huffman apresenta grande eficiência e pode alcançar altas taxas de compressão conforme descrito por Marcelloni em (Marcelloni e Vecchio, 2008). Porém, a codificação de Huffman necessita da construção de um modelo estatístico dos

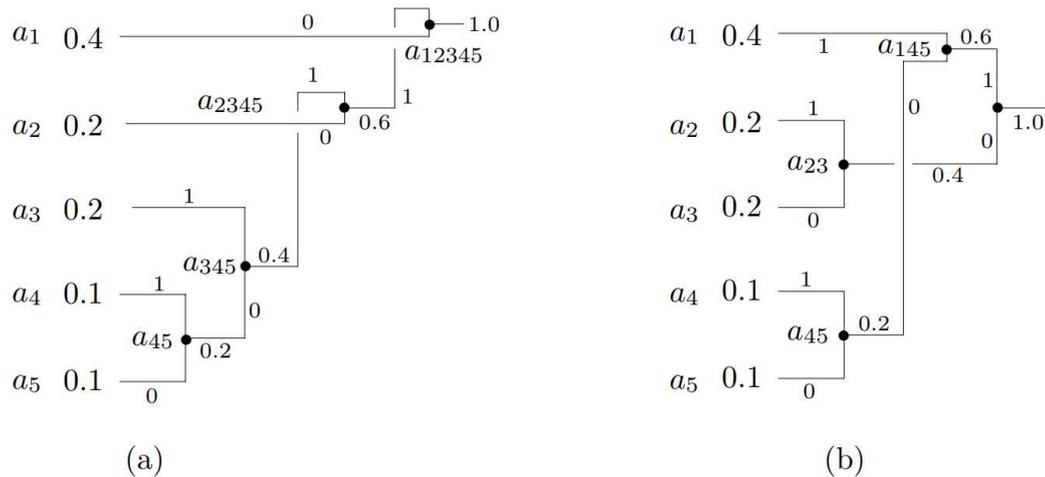


Figura 2.7: Probabilidade dos cinco símbolos com codificação de Huffman exposto por Salomon (Salomon, 2004)

dados, sendo este modelo responsável por gerar a tabela de codificação. Além disso, armazenar a tabela de codificação exige uma quantidade significativa de memória, o que impede o seu uso em dispositivos com poucos recursos de armazenamento. Logo, redes como, WSNs e WBANs, onde grande parte dos dispositivos dispõem de poucos recursos, a codificação de Huffman torna-se imprópria.

De modo semelhante, o método de codificação local sem perdas de Golomb efetua a compressão de dados através de tabelas de codificação (Golomb, 1966). Contudo, suas tabelas de codificação não necessitam de armazenamento como as tabelas de Huffman. Neste sentido, Sayood em (Sayood, 2017) esclareceu que os códigos de Golomb pertencem a uma família de códigos projetados para codificar números inteiros não negativos, de modo que, quanto maior o número, menor a probabilidade de ocorrência. Logo, para determinar uma tabela de codificação de Golomb, o primeiro passo é computar o valor de q , conforme a Equação 2.2.

$$q = \left\lfloor \frac{n}{m} \right\rfloor \quad (2.2)$$

Em seguida, o código é construído em duas partes, o primeiro é o valor de q , codificado em unário, e o segundo é o valor binário de r dado pela Equação 2.3, o qual é codificado de maneira exclusiva.

$$r = n - qm \quad (2.3)$$

Os primeiros valores de $2^c - m$ de r são codificados como inteiros em $c - 1$ bits cada, e os demais são codificados em c bits, onde c é dado pela equação 2.4.

$$c = \lceil \log_2 m \rceil \quad (2.4)$$

Em consequente, é possível determinar a tabela de codificação de Golomb. Onde por definição, $m = 3$ produz $c = 2$ e $r = \{0, 1, 2\}$ por $(2^2 - 3 = 1)$. Deste modo, o primeiro resíduo (r) é codificado em $c - 1 = 1$ bit para se tornar 0. Ademais, os dois resíduos restantes são codificados em 2 bits, tornando-se 10 e 11. De modo semelhante, onde $m = 5$ e produz $c = 3$ com $r = \{0, 1, 2, 3, 4\}$ pelo meio de $(2^3 - 5 = 3)$. Os três primeiros resíduos (r) são codificados com $c - 1 = 2$ bits, os resíduos restantes são codificados com 3 bits cada. Logo, o prefixo para a codificação é definido por 00, 01, 10, 110, 111.

O método de codificação de Golomb define uma tabela padrão conforme o valor de m . Esta representação pode ser visualizada na Tabela 2.5, onde é apresentado valores de $m = 2$ até $m = 9$. Esta definição resulta em pouca utilização dos recursos computacionais, uma vez que não é preciso estabelecer valores pré-calculados para armazenar tabelas. Além disso, a codificação de Golomb permite a compressão das informações em tempo real, de maneira que, as informações são comprimidas diretamente, sem a necessidade de consultar tabelas, como a codificação de Huffman. No entanto, a codificação de Golomb é restrita apenas a valores positivos, o que inviabiliza seu uso em situações como, sensoriamento de temperatura, eletrocardiograma (ECG), acelerômetro, entre outros, onde existe a presença de valores negativos.

Tabela 2.5: Tabela de Codificação de Golomb com $m = 2$ a 9

m/n	0	1	2	3	4	5	6	7	8	9
2	0 0	0 1	10 0	10 1	110 0	110 1	1110 0	1110 1	11110 0	11110 1
3	0 0	0 10	0 11	10 0	10 10	10 11	110 0	110 10	110 11	1110 0
4	0 00	0 01	0 10	0 11	10 00	10 01	10 10	10 11	110 00	110 01
5	0 00	0 01	0 10	0 110	0 111	10 00	10 01	10 10	10 110	10 111
6	0 00	0 01	0 100	0 101	0 110	0 111	10 00	10 01	10 100	10 101
7	0 00	0 001	0 011	0 100	0 101	0 110	0 111	10 00	10 010	10 011
8	0 000	0 001	0 011	0 011	0 100	0 101	0 110	0 111	10 000	10 001
9	0 000	0 001	0 011	0 011	0 100	0 101	0 110	0 1110	10 1111	10 000
10	0 000	0 001	0 011	0 011	0 100	0 101	0 1110	0 1101	0 1110	0 1111

Teuhola em (Teuhola, 1978) propôs uma extensão a codificação de Golomb. Esta extensão requer pouco poder computacional e é capaz de compactar os dados em tempo real. Denominada Exponencial de Golomb (Exp-Golomb) de ordem k , sua principal vantagem é codificar inteiros não negativos em códigos de tamanho variável. Logo, é possível definir valores mais curtos para símbolos mais frequentes. Neste sentido, a codificação Exp-Golomb consiste de duas partes: prefixo e sufixo. A parte do prefixo, consiste de um código unário definido por uma sequência de zeros seguido por um bit 1, Esta sequência pode ser obtida através da Equação 2.5. Onde y denota a solução da equação, n é o numero inteiro ≥ 0 e k define a ordem do Exp-Golomb.

$$y = \left\lceil \log_2 \left(\frac{n}{2^k} + 1 \right) \right\rceil \quad (2.5)$$

A representação do sufixo é definida pela Equação 2.6. Onde z é o resultado da Equação para a representação binária usando $k + y$ bits significantes, n é um inteiro ≥ 0 , e y é o resultado da Equação 2.5.

$$z = n + 2^k(1 - 2^y) \quad (2.6)$$

Como exemplo, seja $n = 1$ de ordem Exp-Golomb $k = 0$, de acordo com a Equação 2.5 temos $y = 1$. Logo, a representação unária de n é 0 seguida por um *bit* 1. Em seguida, o sufixo é obtido através da Equação 2.6, com resultado $z = 0$. Assim, a codificação binária para $n = 1$ de ordem $k = 0$ é representada por 010. A Tabela 2.6, ilustra a codificação de Exponencial de Golomb para diversos valores de n de $k = 0$ até $k = 3$.

Tabela 2.6: Tabela de codificação de Exponencial de Golomb para diversos valores de n com ordem $k = 0$ até $k = 3$

n	$k = 0$	$k = 1$	$k = 2$	$k = 3$
0	1	10	100	1000
1	010	11	101	1001
2	011	0100	110	1010
3	00100	0101	111	1011
4	00101	0110	01000	1100
5	00110	0111	01001	1101
6	00111	001000	01010	1110
7	0001000	001001	01011	1111
8	0001001	001010	01100	010000
9	0001010	001011	01101	010001
10	0001011	001100	01110	010010

Segundo Marcelloni em (Marcelloni e Vecchio, 2010) métodos de compressão local com perdas devem desconsiderar apenas as informações do ruído. Assim, é possível alcançar altas taxas de compressão sem perder informações relevantes. Para este objetivo, o uso da entropia da informação entre diferentes amostras é considerado. Contudo, Marcelloni reitera que a diferença entre consecutivas amostras deve ser regular, e comumente pequena. De modo que, a imprevisibilidade da informação, é afetada pelo ruído.

Neste contexto, métodos de modulação de pulso propõem quantizar a diferença entre consecutivas amostras, reduzir o ruído e compactar a informação. Deste modo, a quantidade da informação é reduzida em número de *bits*, e o erro de tolerância é reduzido ao máximo (Marcelloni e Vecchio, 2010). DPCM (*Differential Pulse Code Modulation*) é um método de modulação de pulso comumente utilizado para compressão de dados local com perda. Este método explora a entropia da informação entre diferentes amostras de sinais digitalizados, e efetua a compressão conforme a diferença entre estas amostras. Para tal finalidade, o codificador DPCM calcula as diferenças $d_i = s_i - \hat{s}_{i-1}$ entre consecutivas amostras de s_i . A primeira amostra de dados, s_0 , é codificada separadamente ou é comprimida como um dado bruto (*raw data*), em ambos os casos, o decodificador pode reconstruir exatamente o valor de s_0 .

De acordo com Marcelloni (Marcelloni e Vecchio, 2010) embora a quantização escalar resulta na compactação com perdas. A quantidade da informação codificada, não refere-se de fato na diferença, mas sim, a versão quantizada denotada por \hat{d}_i , e a diferença q_i entre d_i e \hat{d}_i é denotada como erro de quantização. Deste modo, Dankovic *et. al* em (Danković et al., 2017) reforçaram que o método de modulação por codificação diferencial de pulso DPCM tem sido uma das técnicas mais efetivas para o processamento e transmissão de sinais com base no filtro de predição. Esta técnica é amplamente utilizada em telecomunicações, fala, codificação de imagens, entre outras. Contudo, aferir diferentes amostras e armazenar consecutivos valores quantizados requer uso excessivo de memória. Portanto, redes como, WSNs e WBANs, onde grande parte dos dispositivos dispõem de poucos recursos computacionais, o DPCM pode se tornar impróprio.

Embora a técnica de compressão de dados local possua diversos métodos, e sua aplicabilidade compreende diversas áreas, seu uso em WSNs e WBANs é amplamente estudado. Segundo Li e Liang em (Li e Liang, 2016), a eficiência energética é um desafio crítico para ambas, visto que os dispositivos são tipicamente operados por baterias, e em muitos cenários reais, a substituição da bateria é virtualmente impossível. Logo, a falha de um subconjunto de dispositivos devido à sua falta de energia pode resultar na falha de toda a rede. Além disso, mesmo em circunstâncias onde a substituição da bateria seja possível, melhorar a eficiência energética pode reduzir significativamente as frequentes substituições das baterias e, portanto, reduzir o custo de manutenção da rede. Deste modo, reduzir a quantidade da informação através de compressão de dados local, pode melhorar a eficiência energética e aumentar o tempo de vida da rede.

2.3 RESUMO

Este capítulo apresentou os conceitos importantes acerca das características das WBANs, ilustrou a sua composição, bem como as formas de comunicação direta e indireta. Além disso, expôs os desafios dos recursos de energia e a evolução dos dispositivos sensores e nanossistemas para o surgimento de dispositivos vestíveis e inteligentes. Também descreveu sobre as WBANs, de modo a esclarecer sua topologia, arquitetura e o padrão 802.15.6. Em seguida, o conceito de entropia da informação apresentou os fundamentos necessários para compreender a redundância da informação e a taxa de bits necessária por símbolo. Logo após, apresentou-se os principais conceitos e fundamentos de fusão de dados, de modo a esclarecer, sua classificação, objetivos e utilização. Da mesma forma, apresentou a abordagem de compressão de dados local, suas duas classes de compressão, com perdas e sem perdas, e seus principais métodos de compressão.

3 MÉTODOS DE COMPRESSÃO DE DADOS LOCAL APLICÁVEIS ÀS WBANS

Este capítulo revisa à literatura acerca dos métodos de compressão de dados local das redes de sensores sem fio (WSN) que moldam-se para as redes corporais sem fio (WBANS), além de métodos de compressão de dados local voltados especificamente para as redes WBANS. A Seção 3.1 oferece uma visão geral do uso de compressão de dados a fim de elencar trabalhos que propuseram explorar suas diferentes classes e métodos para reduzir o consumo de energia. A Seção 3.2 apresenta quais métodos de compressão de dados sem perdas têm sido aplicados nas WSNs e WBANS. A Seção 3.3 apresenta os principais métodos de compressão de dados com perdas para ambas as redes. A Seção 3.4 discute os métodos híbridos, ou seja, métodos que fazem o uso de ambas as classes de compressão, com perdas e sem perdas. Por fim, na Seção 3.5, uma tabela ilustra as principais vantagens e desvantagens dos métodos apresentados, assim como uma profunda análise acerca do que foi apresentado.

3.1 VISÃO GERAL

As redes corporais sem fio (WBANS) oferecem o monitoramento em tempo real das pessoas. Neste ambiente, diversos dispositivos corporais coletam, processam e transmitem informações fisiológicas para aplicações de modo a facilitar no diagnóstico (Asif e Sumra, 2017). Entretanto, oferecer um serviço de monitoramento em tempo real é um desafio de elevada complexidade, visto que, o consumo de energia dos dispositivos aumenta com excessivas transmissões (Khan e Pathan, 2018).

De modo a prover o uso consciente ou aprimorar o consumo de energia dos dispositivos, deu-se origem a uma sucessiva série de pesquisas. Deste modo, diversos trabalhos como (Schellhammer et al., 2004; Marcelloni e Vecchio, 2009, 2008; Kimura e Latifi, 2005; van der Byl et al., 2009) afirmam que o maior consumidor de energia nos dispositivos é a transmissão da informação. Neste sentido, uma abordagem para otimizar a energia dos dispositivos é reduzir a quantidade de transmissão através de compressão de dados local (Pu, 2005). Porém, Barr e Asanovic em (Barr e Asanović, 2006) concluíram que a compressão antes da transmissão em dispositivos alimentados por bateria aumenta o consumo de energia se não for introduzida de forma consciente, visto que, os métodos de compressão visam economizar armazenamento e não energia. Além disso, expuseram que a maioria dos métodos de compressão de dados não são praticáveis em dispositivos com poucos recursos computacionais, dada a complexidade dos métodos de compressão e a utilização de grande parte da memória. Assim, métodos apropriadas devem ser adotadas para suprir estas características.

Haja Vista a necessidade de novos métodos para compressão de dados local que atendam os requisitos para dispositivos com poucos recursos, diversos trabalhos foram publicados explorando diferentes classes e métodos de compressão. Assim, para garantir a confiabilidade da informação e diminuir a quantidade de transmissão, a classe de compressão de dados sem perdas foi utilizada. Diversos autores propuseram o uso de tabelas de codificação de Huffman (Marcelloni e Vecchio, 2008; Maurya et al., 2011; Kolo et al., 2012; Tharini e Ranjan, 2009; Antonopoulos e Voros, 2016; Deepu et al., 2017) para analisar a probabilidade de cada informação, o valor de entropia, e assim diminuir a quantidade de *bits* para transmissão. De outro modo, autores sugeriram métodos como a codificação de Golomb-Rice e Exponencial de Golomb (Marcelloni e Vecchio, 2009; Antonopoulos e Voros, 2016) que, assim como a codificação de Huffman, usa probabilidade e valor de entropia para codificar informações com mais ocorrências em

menos símbolos. Além disso, autores exploraram o uso de transformada wavelet discreta (DWT) para fornecer uma representação de frequência do conjunto de amostras consecutivas, e assim reduzir a quantidade da informação (Azar et al., 2018a; Manna e Misra, 2018). Porém, alguns autores verificaram que nem todo sinal necessita de extrema confiabilidade como, temperatura, umidade relativa do ar entre outros. Logo, admite-se um limiar de perda da informação na compressão (Vecchio et al., 2014).

Neste contexto, alguns trabalhos propuseram estabelecer um determinado limiar entre amostras consecutivas (Schoellhammer et al., 2004; Azar et al., 2018b). Deste modo, cada amostra coletada é comparada com este limiar e transmitida caso ultrapasse este valor. Outros autores sugeriram utilizar a modulação por codificação diferencial de pulso (DPCM) e transmitir a diferença entre amostras consecutivas (Marcelloni e Vecchio, 2010). Assim, um determinado valor da informação é descartada, e conseqüentemente menos *bits* são transmitidos. Apesar de diversos autores optarem por diferentes classes de compressão de dados local, encontra-se na literatura trabalhos que utilizam ambas as classes, denominados híbridos. Desta maneira, autores exibiram um método com codificação de Huffman e um limiar para predição (Deepu et al., 2017), e uma adaptação de DPCM com Exponencial de Golomb (Giorgi, 2017).

Apesar da diferença entre as classes, é possível aferir que a avaliação entre os trabalhos pesquisados possuem características equivalentes. Inicialmente, observa-se que independente do tipo de rede e do tipo de sinal apurado, o objetivo do aumento da taxa de compressão é evidente. Em suma, diversos autores afirmam que a alta taxa de compressão garante a economia de energia. Contudo, é indispensável ponderar custos computacionais para o processamento da compressão. Em seguida, percebe-se que eficiência energética e tempo real possuem a mesma quantidade de trabalhos. A respeito do uso das classes de compressão local, sua disparidade é nítida. Logo, é possível afirmar que a utilização dos métodos de compressão de dados sem perdas se dá majoritariamente. De modo a mensurar os métodos, todos os autores avaliaram a taxa de compressão. Além disso, diversos trabalhos analisaram o consumo energético dos dispositivos. Contudo, apesar de diversos trabalhos afirmarem ter como objetivo tempo real, poucas avaliações consideraram a latência na transmissão da informação. A Tabela 3.1 ilustra o estado-da-arte dos métodos de compressão de dados local para WSNs e WBANs, onde todos os trabalhos estão classificados de acordo com o tipo de rede, tipos de serviços, classes de compressão, heterogeneidade dos dispositivos, sinais e tipo de comunicação. Além disso, é apresentado os resultados alcançados para cada trabalho com relação a taxa de compressão, eficiência energética, bem como se o método proposto foi analisado no contexto de aplicações de tempo real. Essas relações ficarão mais compreensíveis à medida que forem detalhadas nos próximos tópicos.

Tabela 3.1: estado-da-arte dos trabalhos de compressão de dados local para WSNs e WBANs

REFERÊNCIAS	AMBIENTE				REQUISITOS				
	Tipo de Rede	Tipo de Serviços	Classe de Compressão	Dispositivos Heterogêneos	Sinais Heterogêneos	Comunicação	Compressão	Eficiência Energética	Tempo Real
Schoellhammer LTC 2004 (Schoellhammer et al., 2004)	WSN	Ambientais	Com Perdas	ND ¹	Sim Temperatura / Umidade / Vento	ND ¹	Sim ≈ 93,80%	ND ¹	Não
Marcelloni e Vecchio SIMPLE 2008 (Marcelloni e Vecchio, 2008)	WSN	Ambientais	Sem Perdas	Não	Sim Temperatura / Umidade	ND ¹	Sim ≈ 67,33%	ND ¹	Não
Marcelloni e Vecchio DPCM 2010 (Marcelloni e Vecchio, 2010)	WSN	Ambientais	Com Perdas	Não	Não Temperatura	ND ¹	Sim ≈ 94,12%	Sim ≈ 62,23%	Não
Maurya et. al MDPC 2011 (Maurya et al., 2011)	WSN	ND ¹	Sem Perdas	Não	ND ¹	ND ¹	Sim ≈ 67,60%	ND ¹	Não
Kolo et. al ALDC 2012 (Kolo et al., 2012)	WSN	Ambientais	Sem Perdas	Não	Sim Umidade / Temperatura	ND ¹	Sim ≈ 74,02%	ND ¹	Não
Tharini e Ranjan 2009 (Tharini e Ranjan, 2009)	WSN	Ambientais	Sem Perdas	Não	Não Temperatura	ND ¹	Sim ≈ 55,00%	ND ¹	Não
Marcelloni e Vecchio LEC 2009 (Marcelloni e Vecchio, 2009)	WSN/WBAN	Ambientais / Corporais	Sem Perdas	Não	Sim Temperatura / Umidade	ND ¹	Sim ≈ 70,81%	Sim ≈ 32,00%	Não
Giorgi 2017 (Giorgi, 2017)	WBAN	Corporais	Com Perdas / Sem Perdas	ND ¹	Sim Acelerômetro / ECG Freq. Cardíaca Fetal	Direta	Sim ≈ 95,00%	Sim ND ¹	Não
Azar et. al ALTC 2018 (Azar et al., 2018b)	WBAN	Corporais	Com Perdas	Não	Sim Freq. Respiratória Freq. Cardíaca / Temp. Corporal	ND ¹	Sim ≈ 95,10%	Sim ND ¹	Não
Azar et. al DWTLS 2018 (Azar et al., 2018a)	WBAN	Corporais	Sem Perdas	Não	Sim Pressão Sang. Temperatura / PPG / Acel.	ND ¹	Sim ≈ 92,90%	ND ¹	Não
Antonopoulos e Voros RTH 2016 (Antonopoulos e Voros, 2016)	WBAN	Corporais	Sem Perdas	Não	Sim ECG / EEG	ND ¹	Sim ≈ 62,50%	ND ¹	Sim Atraso ND ¹
Antonopoulos e Voros GVM 2016 (Antonopoulos e Voros, 2016)	WBAN	Corporais	Sem Perdas	Não	Sim ECG / EEG	ND ¹	Sim ≈ 68,00%	ND ¹	Sim Atraso ND ¹
Deepu et. al 2017 (Deepu et al., 2017)	WBAN	Corporais	Com Perdas / Sem Perdas	Não	Não ECG	ND ¹	Sim ≈ 9,54%	Sim ≈ 45,00%	Sim Atraso ND ¹
Manna e Misra 2018 (Manna e Misra, 2018)	WBAN	Corporais	Com Perdas	Sim	Sim ECG / Temperatura	Direta	Sim ≈ 67%	ND ¹	Sim
Tsai e Kuo 2018 (Tsai e Kuo, 2018)	WBAN	Corporais	Sem Perdas	Não	Não ECG	ND ¹	Sim ≈ 70,14%	ND ¹	Não
Proposta	WBAN	Corporais	Com Perdas / Sem Perdas	Sim	Sim Temperatura Corporal PPG / ECG	Direta	Sim	Sim	Sim

¹ Não Definido.

3.2 COMPRESSÃO DE DADOS LOCAL SEM PERDAS

Compressão de dados local sem perdas garante a confiabilidade da informação, sua utilização emprega aplicações que exigem alta precisão da informação e não toleram corrompimento por processos de compressão. Nas WSNs por exemplo, monitoramento de gases vulcânicos ou nível de água de um rio exigem enorme precisão. Logo, necessitam da confiabilidade da informação. Contudo, o tamanho e a complexidade dos métodos ocupam grande parte da memória dos dispositivos. Neste sentido, diversos trabalhos (Marcelloni e Vecchio, 2009; Azar et al., 2018a; Marcelloni e Vecchio, 2008; Maurya et al., 2011; Kolo et al., 2012; Tharini e Ranjan, 2009; Antonopoulos e Voros, 2016) propuseram métodos de compressão de dados local sem perdas adequados para WSNs.

Marcelloni e Vecchio em (Marcelloni e Vecchio, 2008) desenvolveram um método simples de compressão de dados que usa poucos recursos computacionais para as WSNs. O método de compressão explora o alto valor de correlação que comumente existe entre amostras consecutivas coletadas por um dispositivo. Deste modo, cada medida do dispositivo m_i é convertida para digital através de um ADC (do inglês, *analog-to-digital converter*) para a representação binária r_i com R bits, onde R é a resolução do ADC. Neste sentido, para cada nova aquisição m_i , o algoritmo calcula a diferença $d_i = r_i - r_{i-1}$, a qual é a entrada para o codificador de entropia. O codificador de entropia executa a compressão sem perdas, e codifica as diferenças de forma compacta, com base em suas características estatísticas. Cada d_i é representado como uma sequência de bits bs_i composta de duas partes $s_i|a_i$, onde s_i codifica o número n_i de bits necessários para representar d_i , e a_i é a representação de d_i . Contudo, a parte a_i da sequência de bits bs_i é um tamanho variável. Logo, seu código inteiro é representado por complemento de 2 e precisa ser gerado da seguinte forma:

1. Se $d_i > 0$, a_i corresponde aos n_i bits menos significativos da representação por complemento de 2 de d_i
2. Se $d_i < 0$, efetua o complemento de 2 de d_i , subtrai 1 e utiliza os n_i bits menos significativos
3. Se $d_i = 0$, s_i é codificado como 00 e a_i não é representado.

Logo, se $d_i = 0$, então $n_i = 0$, senão $n_i = \lceil \log_2(|d_i|) \rceil$. Assim, n_i é no máximo igual ao valor de R . E, s_i é um código de comprimento variável gerado a partir de n_i através da codificação de Huffman. Assim, o procedimento utilizado por Marcelloni e Vecchio para gerar a_i garante que todos os valores possíveis possuam códigos diferentes. De modo que, para valores $d_i = 0$, $d_i = +1$, $d_i = -1$, $d_i = +255$ e $d_i = -255$, seus valores codificados seriam respectivamente 00, 010|1, 010|0, 111110|11111111 e 111110|00000000. Uma vez que o bs_i é gerado, ele é anexado ao fluxo de bits que forma a versão compactada da sequência de medidas m_i . Conforme descrito na Seção 2.2.4, a ideia básica da codificação de Huffman é mapear um alfabeto para uma melhor representação, composto de sequência de bits de tamanhos variáveis, para que os símbolos que ocorrem frequentemente tenham uma representação menor do que aqueles que ocorrem raramente. No algoritmo proposto por Marcelloni e Vecchio, os símbolos são $R + 1$, e as probabilidades diminuem com o aumento dos valores. A Tabela 3.2 ilustra a tabela de codificação utilizada para o algoritmo de compressão de dados sem perdas.

No cenário estabelecido, um dispositivo coletava amostras a cada 2 minutos durante um período de 48 horas, com um total de 1440 amostras. Para mensurar a taxa de compressão do método proposto, dados reais com valores de temperatura e umidade relativa do ar foram

Tabela 3.2: Tabela de Codificação de Huffman para o método de compressão sem perdas de Marcelloni e Vecchio

n_i	s_i	d_i
0	00	0
1	010	-1,+1
2	011	-3,-3,+2,+3
3	100	-7,...,-4,+4,...,+7
4	101	-15,...,-8,+8,...,+15
5	110	-31,...,-16,+16,...,+31
6	1110	-63,...,-32,+32,...,+63
7	11110	-127,...,-64,+64,...,+127
8	111110	-255,...,-128,+128,...,+255
9	1111110	-511,...,-256,+256,...,+511
10	11111110	-1023,...,-512,+512,...,+1023
11	111111110	-2047,...,-1024,+1024,...,+2047
12	1111111110	-4095,...,-2048,+2048,...,+4095
13	11111111110	-8191,...,-4096,+4096,...,+8191
14	111111111110	-16383,...,-8192,+8192,...,+16383

Tabela 3.3: Taxas de compressão com o método simples de Marcelloni e Vecchio

	Temperatura	Umidade Relativa do Ar
Dados Originais	23040 bits	23040 bits
Após Compressão	7605 bits	7527 bits
Taxa de Compressão	66.99%	67.33%

extraídos de um *dataset*. Os resultados alcançados são visualizados na Tabela 3.3, onde, para temperatura o método atingiu 66.99% de taxa de compressão e umidade relativa do ar 67.33%.

Como a transmissão de um *bit* necessita de energia comparável á milhares de instruções (Pottie e Kaiser, 2000). Logo, a redução de *bits* após a compressão produz um consumo de energia para a transmissão. Porém, o consumo real do dispositivo, sem ou após o uso do método não é apresentado, impossibilitando a análise quanto ao tempo de vida do dispositivo. Além disso, armazenar a tabela de codificação exige uma quantidade significativa de memória, o que impede o seu uso em dispositivos com poucos recursos de armazenamento.

No método proposto por Marcelloni e Vecchio em (Marcelloni e Vecchio, 2008) à alta correlação entre amostras consecutivas e os princípios da compressão por entropia propiciaram altas taxas de compressão. Porém, armazenar tabelas de codificação é impraticável em dispositivos com poucos recursos de memória. Nesta perspectiva, Marcelloni e Vecchio (Marcelloni e Vecchio, 2009) propuseram um método simples de compressão por entropia sem perdas LEC (do inglês, *Lossless Entropy Code*). Este método explora uma versão modificada do código Exponencial-Golomb (Exp-Golomb) de ordem 0 (Teuhola, 1978) com o padrão de compressão JPEG (Pennebaker e Mitchell, 1992) em que uma tabela de codificação fixa é criada, cujo tamanho depende apenas da precisão do conversor ADC. Logo, com o tamanho da tabela fixada *a priori*, o método LEC não sofre com o crescimento da tabela de codificação. Além disso, LEC é capaz de calcular uma versão comprimida de cada amostra em tempo real, o que reduz a utilização de memória.

O método de compressão LEC segue a mesma proposta que o método simples de compressão de dados visto anteriormente, onde, cada medida do dispositivo m_i é convertida para digital através de um ADC para a representação binária r_i com R bits, onde R é a resolução do ADC, que é, o número 2^R de valores discretos que o ADC gera na faixa de valores analógicos.

Cada d_i é representado como uma sequência de *bits* bs_i composta de duas partes $s_i|a_i$, onde s_i codifica o número n_i de *bits* necessários para representar d_i , e a_i é a representação de d_i . A Figura 3.1 ilustra o diagrama de blocos do método de compressão e descompressão proposto por (Marcelloni e Vecchio, 2009).

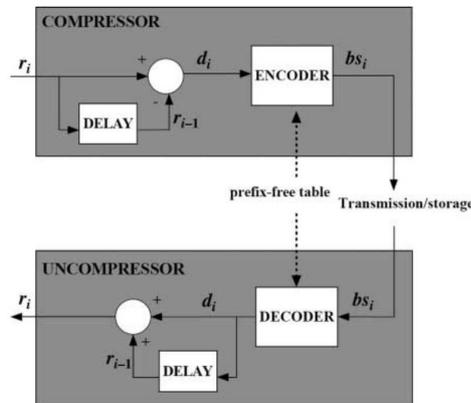


Figura 3.1: Diagrama de bloco do compressor e descompressor do método LEC

Para demonstrar a eficácia do método de compressão, sinais como temperatura, umidade relativa do ar, radiação solar e ECG de várias bases de dados reais foram mensurados. Primeiramente, os sinais de temperatura e umidade relativa do ar que caracterizam-se por apresentar maior correlação entre amostras consecutivas foram analisados. Os resultados de cada base de dados e as taxas de compressão são visualizados na Tabela 3.4. De acordo com a Tabela 3.4, as maiores taxas de compressão alcançadas com o método LEC foram, de 70.81% para temperatura e 62.14% para umidade relativa do ar. Além disso, observa-se que o método LEC alcança as maiores taxas de compressão em bases de dados caracterizadas por uma baixa entropia H_d e, em geral, uma baixa variabilidade entre amostras consecutivas (ou seja, valores baixos da média e desvio padrão das diferenças entre amostras consecutivas).

Tabela 3.4: Taxas de compressão com o método LEC

Base de Dados	Temperatura				Umidade Relativa do Ar			
	Dados Originais	Após Compressão	Taxa de Compressão	H_d	Após Compressão	Taxa de Compressão	H_d	
A	202432	70069	65.39%	5.10	76635	62.14%	5.84	
B	1038608	303194	70.81%	4.05	396442	61.83%	5.85	
C	381008	156099	59.03%	6.15	180192	52.71%	7.14	
D	344268	158994	53.83%	6.82	178725	48.10%	7.67	

Após verificar os sinais que apresentam uma alta correlação entre as amostras, sinais como, radiação solar e ECG que apresentam pouca correlação foram mensurados. Neste contexto, o método LEC atingiu uma taxa de compressão de 70.31% para sinais de radiação solar e ECG 71.53% para sinais ECG. Deste modo, verifica-se que, apesar da pouca correlação entre amostras consecutivas, o método LEC apresenta uma alta taxa de compressão. Este feito só é possível devido ao baixo valor de entropia que cada sinal possui, 4.24% para radiação solar e 3.78% para sinais ECG. Assim, como o LEC é um método de compressão por entropia, independente da suavidade dos sinais, seu desempenho aumenta quando a entropia diminui. Em vista disso, embora a temperatura e a umidade relativa variem de forma lenta e suave, o número de valores possíveis da diferença entre amostras consecutivas é, em média, maior do que o número de valores possíveis da diferença entre amostras consecutivas de radiação solar e sinais de ECG. Logo, os sinais de radiação solar e ECG são caracterizados por conter grandes diferenças entre amostras consecutivas, porém, essas diferenças são repetitivas. Assim, verifica-se que um maior número de símbolos é necessário para codificar sinais de temperatura e umidade relativa do ar.

O método de compressão por entropia sem perdas LEC, apresentou altas taxas de compressão com diferentes tipos de sinais. Além disso, o método demonstrou que o baixo valor de entropia implica na alta taxa de compressão, sem a necessidade da correlação entre amostras consecutivas. Neste sentido, Marcelloni e Vecchio consideraram que o uso do método LEC salva $\approx 32\%$ de energia na transmissão de dados. Contudo, a heterogeneidade dos dispositivos não é levada em consideração, todas as simulações contaram com dispositivos homogêneos. Em uma rede WBAN, cada dispositivos coleta um tipo de sinal diferente e a heterogeneidade é evidente. Além disso, a transmissão e a recepção dos dados só acontece após o uso total de *payload* do pacote, o que acarreta em alta latência.

Após o método de compressão sem perdas LEC, uma série de pesquisas foram desencadeadas (Maurya et al., 2011; Kolo et al., 2012; Tharini e Ranjan, 2009). Neste sentido, cada pesquisador procurou aprimorar o método LEC de forma a mitigar sua principal desvantagem, a adaptabilidade. Maurya *et. al* em (Maurya et al., 2011) propuseram o MPDC (do inglês, *Median Predictor based Data Compression*) que faz alterações no bloco de predição, (*delay*) da Figura 3.1. No método proposto, o bloco de predição explora a variação de três amostras consecutivas para estimar o valor de \hat{r}_i , que permite reduzir o valor de d_i , e não apenas o valor da última amostra como LEC. Com esta variação, o método MPDC alcançou taxas de compressão de 67.60%. Contudo, seus resultados não são conclusivos, de modo que, o tipo de sinal não está explicitado. Além disso, armazenar três amostras no bloco de predição gera uso excessivo de memória, o que é escasso para os dispositivos nas WBANs. Ainda assim, MPDC faz a transmissão e a recepção dos dados de modo semelhante ao LEC, o que impossibilita o seu uso em aplicações nas WBANs de tempo real.

Kolo *et. al* em (Kolo et al., 2012) propuseram o ALDC (do inglês, *Adaptive Lossless Data Compression*). Este método utiliza três tabelas de codificação preestabelecidas para melhorar a estatística de diferentes sinais para compressão. De maneira oposta ao LEC, que efetua a compressão dos dados em tempo real, ALDC coleta uma sequência de amostras, computa todas as diferenças entre elas, e em seguida comprime todas as diferenças através das três tabelas pré-estabelecidas. Assim, o ALDC consegue superar $\approx 3\%$ taxa de compressão do método LEC. Por outro lado, o método ALDC exige armazenar uma sequência de amostras e três tabelas de codificação. Neste sentido, a melhoria na taxa de compressão não é relevante devido ao alto custo dos recursos de memória do dispositivo. Além disso, o método não realiza a compressão de dados em tempo real, o que inviabiliza a sua aplicação.

Outro variante do método LEC foi proposto por Tharini e Ranjan em (Tharini e Ranjan, 2009). De modo contrário ao LEC, que utiliza uma tabela de codificação definida, este método explora o uso de uma tabela de codificação dinâmica para compactar os grupos de n_i (Tabela 3.3). Assim, esta tabela dinâmica analisa e gerencia amostras consecutivas para codificar símbolos mais frequentes com menos *bits*. O uso de uma tabela dinâmica para os valores de n_i oferece um maior grau de adaptabilidade, e particularmente útil ao comprimir sinais com pouca correlação, além de aumentar a taxa de compressão. Contudo, gerenciar e analisar a tabela dinâmica entre amostras consecutivas requer muito processamento, além de aumentar drasticamente o consumo de energia.

Antonopoulos e Voros em (Antonopoulos e Voros, 2016) propuseram dois métodos, o primeiro RTH (do inglês, *Real-Time Huffman*), explora a correlação entre amostras consecutivas a fim de oferecer uma eficiente tabela dinâmica da codificação de Huffman. O segundo, denominado GVM (*GolombVariableM*) faz uso da codificação de Golomb-Rice para comprimir amostras consecutivas. Neste método o uso de predição é inexistente, logo, o sinal é codificado apenas pela diferença entre as amostras. Para mensurar os métodos, base de dados reais com sinais de ECG e EEG foram utilizados. Com relação a taxa de compressão do sinal ECG, o método RTH atingiu $\approx 62.5\%$ e o GVM $\approx 68\%$, estes valores superaram a taxa de compressão do método

LEC, que atingiu $\approx 60.0\%$. Além disso, Antonopoulos e Voros afirmam o emprego dos métodos em aplicações de tempo real, e que a latência entre a compressão e transmissão das amostras são mínimas. Porém, assim como o método proposto por (Tharini e Ranjan, 2009) as constantes atualizações da tabela dinâmica do método RTH apresentam altas taxas de processamento. Logo, provoca o aumento do consumo de energia. Neste sentido, o consumo de energia dos dispositivos não é apresentado. Assim, não é possível identificar se os métodos são aplicáveis nas WBANs de tempo real, uma vez que, o consumo de energia dos dispositivos e seu tempo de vida são cruciais para o funcionamento dos serviços de coleta.

Azar *et. al* em (Azar et al., 2018a) propuseram um método de compressão baseado na transformada wavelet discreta (DWT) com o uso de LS (Sweldens, 1998). A transformada fornece uma representação de frequência do conjunto de amostras consecutivas. Neste sentido, a transformação de redundantes amostras em um determinado período de tempo é comprimida com menos *bits*. O método proposto por Azar *et. al* resulta na diminuição da redundância da informação, e alcança altas de compressão de dados $\approx 92.9\%$ em sinais de temperatura. Contudo, em concordância com os autores, o custo computacional precisa ser levado em consideração, uma vez que, a alta taxa de compressão resultou no consumo excessivo de energia dos dispositivos.

3.3 COMPRESSÃO DE DADOS LOCAL COM PERDAS

Embora os métodos de compressão local com perdas provoquem uma certa distinção quanto as amostras entre origem e destino, é eminente a alta taxa de compressão quando comparado com os métodos de compressão sem perdas. Deste modo, aplicações que não necessitam de informações precisas fazem o uso desta compressão para aumentar o tempo de vida dos dispositivos. Neste sentido, trabalhos como Schoellhammer *et. al* em (Schoellhammer et al., 2004), Marcelloni e Vecchio em (Marcelloni e Vecchio, 2010) e Azar *et. al* em (Azar et al., 2018b) discutem a compactação com perdas e suas abordagens para aprimorar seus métodos de compressão.

Schoellhammer *et. al* em (Schoellhammer et al., 2004) propuseram um método para aplicações em monitoramento climático LTC (do inglês, *Lightweight Temporal Compression*). Este método implementa um algoritmo interpolador linear, de modo que, entre todos os dados adquiridos localmente pelo dispositivo, apenas um número limitado de valores selecionados são transmitidos, onde a informação original é reconstruída dentro de um dado erro de tolerância e . Neste contexto, Schoellhammer *et. al* elucidaram que o método de compressão de dados com perdas LTC é esclarecido da seguinte forma.

Considere um conjunto de dados x e um conjunto vazio s , conforme demonstra a Figura 3.2 (a), o método de compressão é iniciado e obtém o primeiro ponto de dados (t_1, v_1) de x e armazena em z . Então, o segundo ponto (t_2, v_2) é utilizado para definir os limites superior $UL(t_2, v_2 + e)$ e inferior $LL(t_2, v_2 - e)$. Por conseguinte, os valores de *highLine* e *LowLine* são calculados, com *HighLine* conectado a UL e z , e *LowLine* conectado a LL e z . Para cada ponto de dados (t_j, v_j) , onde $j > 2$, o ponto em um segmento vertical usando a margem e é transformado. Se o ponto de dados pertencer ao conjunto de todas as linhas possíveis, os valores de UL e LL são utilizados e *highLine* e *LowLine* são recalculados, conforme ilustrado na Figura 3.2 (b). Caso um ponto (t_j, v_j) não pertencer ao conjunto de todas as linhas possíveis, como visto na Figura 3.2 (c), z é emitido para s , e z é setado para ser o ponto $(t_{j-1}, (UL + LL)/2)$.

O método proposto por Schoellhammer *et. al* (Schoellhammer et al., 2004) alcançou uma taxa de compressão de 93.8% para sinais de temperatura. Além disso, com o descarte de amostras dentro da margem de erro e , uma significativa economia dos dispositivos é atingida. Contudo, este método possui uma alta taxa de latência, uma vez que, o método só permite

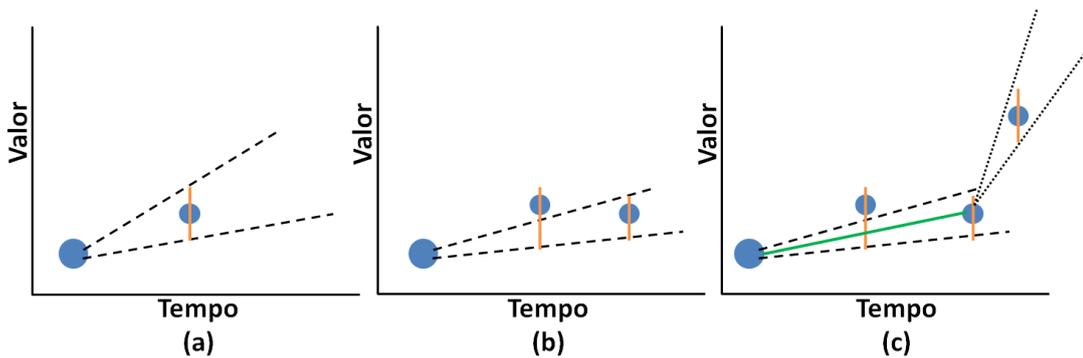


Figura 3.2: Reprodução de compressão de dados local com perda LTC (Schoellhammer et al., 2004)

reconstruir os pontos intermediários da margem de erro e entre duas amostras no dispositivo. Logo, o método não é adequado para aplicações WBAN de tempo real.

Marcelloni e Vecchio em (Marcelloni e Vecchio, 2010) propuseram realizar a compressão de dados local com perdas através do método de modulação por codificação diferencial de pulso (DPCM), que realiza a quantização das diferenças entre amostras consecutivas. Como diferentes combinações dos parâmetros no processo de quantização determinam diferentes limiares entre o desempenho de compressão e a perda de informação, este método infere um conjunto de combinações dos parâmetros a fim de obter melhores resultados. Neste sentido, o usuário define estas combinações conforme a sua aplicação. A Figura 3.3 ilustra o diagrama de blocos do método DPCM proposto por Marcelloni e Vecchio, onde, no codificador, s_i provém da amostra atual, d_i denota a diferença entre $s_i - \hat{s}_{i-1}$, e \hat{d}_i é a versão quantizada de d_i . De modo semelhante, o decodificador recebe a informação e_i , a qual é originada pelo quantizador Q_1 no codificador, o quantizador Q_2 gera o valor de \hat{d}_i , e s_i denota o valor da informação $\hat{d}_i + \hat{s}_{i-1}$. O método proposto por Marcelloni e Vecchio superou o método LTC proposto por Schoellhammer *et al.* (Schoellhammer et al., 2004) e atingiu altas taxas de compressão $\approx 94.12\%$ com pouca complexidade. Contudo, a quantização entre um conjunto de amostras demanda muitos recursos de memória, e o tempo de processamento entre a quantização e a transmissão produz um nível de latência não especificado pelos autores.

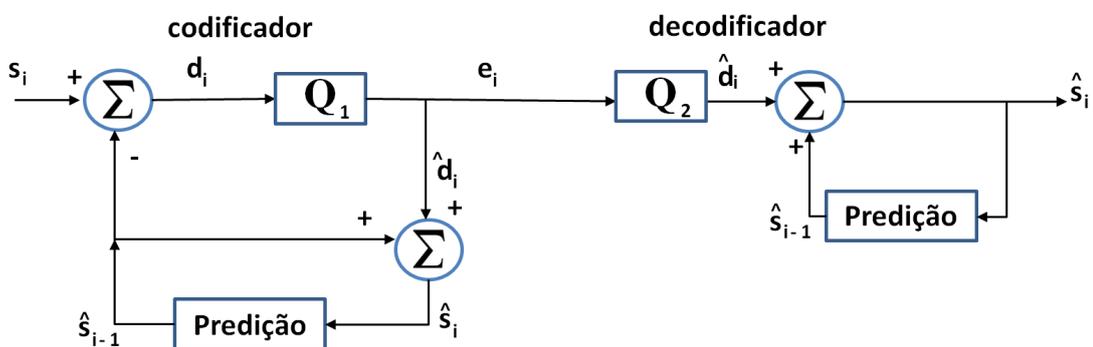


Figura 3.3: Modelo de um quantizador codificador e decodificador DPCM (Gray e Neuhoff, 1998)

Azar *et al.* em (Azar et al., 2018b) propuseram uma alteração ao método LTC para minimizar a taxa de erro de reconstrução e . Assim, quando um ponto (t_j, v_j) não pertencer ao conjunto de todas as linhas possíveis, conforme visto na Figura 3.2, dois novos pontos são adicionados ao conjunto s , $(t_j - 1, v_j - 1)$ e (t_j, v_j) . Contudo, esta alteração sofre com a perda de compressão de dados, uma vez que, os pontos adicionais são acrescentados ao conjunto de

amostras s . Para solucionar esta adversidade, Azar *et. al* propuseram incorporar o método DPCM na fase de compressão. A utilização de ambos os métodos atingiu uma taxa de compressão de $\approx 95.1\%$, além de diminuir a quantidade de transmissões. Porém, o método considera armazenar o conjunto s durante um período de 60 segundos, de modo a aferir as amostras consecutivas. Por consequência, não é possível utilizá-lo para aplicações WBAN de tempo real que consideram uma latência máxima de 250ms para aplicações não médicas (of Electrical e Engineers, 2012)

Manna e Misra em (Manna e Misra, 2018) utilizaram compressão de dados local no dispositivo central *sink*. Assim, todo dado recebido pelos dispositivos sensores eram comprimidos em tempo real e transmitidos para uma base de dados na nuvem que armazenava a informação comprimida. A implementação proposta utiliza a técnica de transformada wavelet discreta (DWT) para reduzir o tamanho da informação recebida por cada dispositivo sensor. No cenário avaliado os sinais de temperatura e ECG foram utilizados para mensurar a taxa de compressão e o consumo energético do *sink*. Os resultados demonstraram que o consumo de energia diminuiu na medida que menos dados são transmitidos. Neste sentido, realizar transmissões apenas com o *payload* carregado favorece a eficiência energética. Contudo, é necessário um *trade-off* entre minimizar a quantidade de transmissões e aumentar a latência, haja visto os padrões estabelecidos para as WBANs (of Electrical e Engineers, 2012).

3.4 COMPRESSÃO DE DADOS LOCAL HÍBRIDA

Os métodos de compressão de dados local com perdas e sem perdas dispõem de características distintas. Entre elas, destacam-se a confiabilidade da informação e a taxa de compressão. Compressão sem perdas por exemplo, garante a confiabilidade da informação através de tabelas previamente determinadas. Por outro lado, compressão de dados com perdas alcança maiores taxas de compressão com o uso de uma margem de erro. Contudo, em determinados ambientes o uso de ambos os métodos é necessário, como nas redes WBANs, onde diversos dispositivos capturam sinais distintos. Sinais como ECG e pressão arterial por exemplo, demandam da confiabilidade da informação. Logo, uma pequena alteração compromete no diagnóstico médico. Em contrapartida, considera-se uma breve taxa de erro em sinais como acelerômetro, temperatura e umidade. Neste sentido, trabalhos como (Giorgi, 2017; Deepu et al., 2017; Tsai e Kuo, 2018) propuseram métodos híbridos a fim de mensurar diversos sinais em uma mesma aplicação.

Giorgi em (Giorgi, 2017) propôs o uso da modulação por codificação diferencial de pulso (DPCM) para compressão com perdas, e uma alteração nas tabelas de codificação de Exponencial de Golomb para compressão sem perdas. Deste modo, o método busca diminuir a quantidade de *bits*, e assim reduzir o consumo de energia dos dispositivos. O método obteve taxas de compressão de $\approx 95\%$, e não é específico para apenas um sinal, suas simulações contaram com dados de bases reais com sinais de ECG, acelerômetro e frequência cardíaca fetal. Contudo, ele não detalha o consumo de energia, o que impossibilita afirmar que seu método aumenta o tempo de vida dos dispositivos.

Deepu *et. al* em (Deepu et al., 2017) propuseram um método híbrido para diminuir o consumo de energia dos dispositivos e melhorar a qualidade de sinal ECG. Este método dispõe de um bloco de compressão com perdas, uma unidade de descompressão e um codificador por entropia, como ilustra a Figura 3.4. Assim, primeiramente os dados são comprimidos com o método de compressão com perdas, cuja qualidade é suficiente para avaliação preliminar dos sinais ECG. Nos casos que exige-se análise mais minuciosa, o sinal ECG é comprimido com compressão de dados sem perdas.

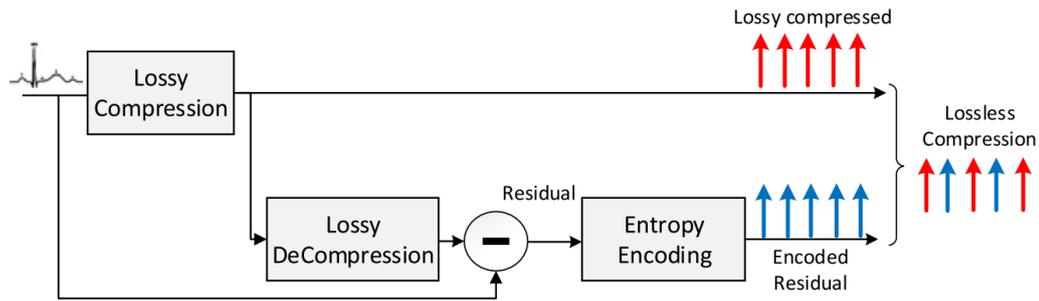


Figura 3.4: Diagrama de bloco híbrido proposto por Deepu

O método proposto por Deepu *et. al* apresenta uma considerável redução de energia dos dispositivos, conforme ilustrado na Tabela 3.5. Além disso, segundo Deepu *et. al* seu método aplica-se em situações de tempo real para outros sinais fisiológicos e sensores sem fio. Apesar do método apresentar $\approx 45\%$ de economia de energia dos dispositivos com a compressão sem perdas em tempo real, e os autores informarem que seu método pode ser aplicado em tempo real, informações como latência para compressão e transmissão não foram apresentadas. Logo, não é possível assegurar que o método pode ser aplicado para diferentes tipos de sinais para transmissão em tempo real.

Tabela 3.5: Consumo de energia dos dispositivos com diferentes duty cycles apresentados por Deepu *et. al* (Deepu *et al.*, 2017)

Duty Cycle	Método de Compressão	Consumo para transmissão sem fio
Tempo Real	Sem compressão	1.883mA
Tempo Real	Sem perdas	0.988mA
Tempo Real	Com perdas	0.335mA
5s	Sem compressão	1.799mA
5s	Sem perdas	0.921mA
5s	Com perdas	0.209mA
10s	Sem compressão	1.753mA
10s	Sem perdas	0.879mA
10s	Com perdas	0.167mA

Tsai e Kuo em (Tsai e Kuo, 2018) propuseram um método para compressão de sinais ECG sem fio sem perdas com a combinação de três diferentes métodos para aumentar o espaço de armazenamento do dispositivo e reduzir a latência na transmissão. A primeira técnica utilizada é uma predição linear adaptativa, a segunda técnica é a codificação de Golomb-Rice, usada com um tamanho dinâmico para codificar o erro de predição, e a terceira é uma adaptação para permitir o processo de descompressão em tempo real. A Figura 3.5 ilustra o diagrama de bloco de compressão e descompressão proposto por (Tsai e Kuo, 2018). Assim, percebe-se que um valor de predição $\hat{y}[n]$ é usado para derivar o valor presente de amostras anteriores. Logo, o valor de predição $e(n)$ é produzido pela diferença $x[n] - \hat{y}[n]$. Para melhorar o desempenho de compressão para o sinal ECG, utiliza-se um bloco de erro de predição linear adaptativo e um bloco de codificação de Golomb-Rice. Assim, aplica-se o valor do erro de predição $e(n)$ na codificação de Golomb-Rice e para determinar o parâmetro k .

Com a predição linear e a constante atualização do valor de k para a codificação de Golomb-Rice, o método proposto por (Tsai e Kuo, 2018) apresenta altas taxas de compressão. Contudo, não é possível determinar que o método proposto é aplicável em tempo real, uma vez que, o valor de latência entre amostras consecutivas não é apresentado. Além disso, a atualização

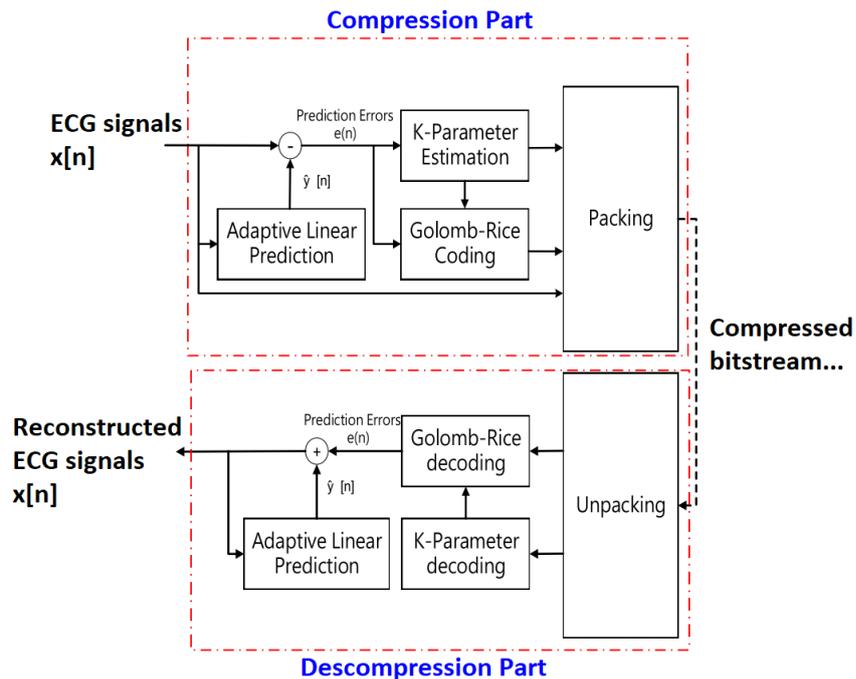


Figura 3.5: Diagrama de bloco híbrido do compressor e descompressor

dinâmica da tabela de codificação de Golomb-Rice através do bloco de parâmetro k produz grande consumo de energia, como descrito por Marcelloni e Vecchio em (Vecchio et al., 2014).

3.5 DISCUSSÃO

Esta seção apresenta uma síntese do estado-da-arte sobre o uso dos métodos de compressão de dados local com eficiência energética de WSNs aplicáveis no contexto das WBANs. A Tabela 3.1 assinala os pontos revelantes dos trabalhos pesquisados e analisados, classificando-os conforme seu ambiente e requisitos observados. Essa classificação do ambiente envolve duas propriedades, os tipos de rede WSNs e WBANs onde comumente encontra-se dispositivos com poucos recursos computacionais e são a essência desta dissertação, e os tipos de serviços, compostos pelos serviços ambientais e corporais. Assim, analisou-se diversos documentos, inclusive um que não definiu qual tipo de serviço avaliou (Maurya et al., 2011), e outro que analisou ambos os ambientes. Em seguida, tendo o cômputo da eficiência energética e tempo real como alvos, verificou-se os requisitos de cada trabalho, a fim de constatar quais trabalhos satisfazem premência. Neste sentido, deu-se a análise das classes de compressão local utilizadas, a heterogeneidade dos dispositivos e sinais, a comunicação entre os dispositivos e o *sink* e, por fim, a taxa de compressão, que elucida o valor aproximado da compressão após a transformação da informação original.

Com referência as classes de compressão de dados local, o emprego da classe sem perdas deu-se majoritariamente pelos autores. Este feito é de fácil compreensão, uma vez que, dada a multiplicidade dos trabalhos, grande parte é ocupada pelas WBANs. Neste sentido, diversos autores preocuparam-se com sinais contínuos como ECG e EEG, e a confiabilidade destes sinais é imprescindível. No âmbito da classe de compressão com perdas, os trabalhos pesquisados atentaram-se ao requisito compressão, em princípio, os métodos desta classe têm em vista reduzir a quantidade da informação com um aceitável nível de erro. Assim, um aumento na taxa de compressão é alcançado. Encontrou-se ainda, trabalhos que fizeram uso de ambas as classes, a fim de mensurar diferentes tipos de sinais.

No que tange a imposição de dispositivos heterogêneos, dois trabalhos não apresentaram informações com relação aos dispositivos. Ademais, todos consideraram apenas um tipo. Haja visto, que uma rede WBAN compreende de um conjunto de dispositivos que coletam diferentes tipos de sinais, a heterogeneidade dos dispositivos é perceptível. Nesse momento, ficou evidente que a maioria dos trabalhos almejavam garantir confiabilidade ou taxa de compressão e eficiência em energia. Logo, não é possível corroborar que os trabalhos de fato aumentam o tempo de vida dos dispositivos. Ainda que, dispositivos homogêneos compuseram cada cenário, em sua maioria, os trabalhos avaliados usam sinais heterogêneos para mensurar a taxa de compressão, com exceção de um trabalho que não apresentou o tipo de sinal, e outros dois trabalhos específicos para sinais ECG. A heterogeneidade dos sinais é necessária à avaliação do método proposto, visto que, é possível aferir a disparidade da taxa de compressão entre sinais distintos. Além disso, cada sinal detém de características específicas, e essa dissemelhança caracteriza as redes WBANs.

A análise do estado-da-arte permitiu constatar que o cenário da WBAN ainda carece de soluções de compressão de dados local para favorecer a eficiência energética dos dispositivos e a transmissão em tempo real das informações. Pela Tabela 3.1 observou-se que os trabalhos desconsideraram um cenário com a heterogeneidade dos dispositivos. Além disso, apenas um trabalho expressa a comunicação entre os dispositivos e o *sink*. Contudo, este trabalho não avalia o consumo de energia dos dispositivos. Consequentemente, é cabível afirmar que poucos trabalhos de fato apresentam eficiência energética após a compressão e transmissão das informações, sendo estes, não adequados para aplicações de tempo real nas WBANs. Neste contexto, um trabalho apresentou dois métodos para aplicações em tempo real. No entanto, a eficiência energética não é esclarecida, tampouco o tempo de vida dos dispositivos. Com exceção do trabalho apresentado por (Deepu et al., 2017), altas taxas de compressão foram apresentadas. Porém, não é possível assegurar que a alta taxa de compressão aprimora a eficiência energética, uma vez que, conforme exposto por (Azar et al., 2018a) o custo de processamento para altas taxas de compressão aflige os recursos de energia. No cenário das WBANs, aplicações em tempo real dispõem de dispositivos e sinais heterogêneos. Estas aplicações são comumente destacadas pela quantidade de transmissão de informações para as centrais de saúde, e pela carência de recursos de energia de seus dispositivos. Logo, necessita-se de soluções em WBANs de tempo real para reduzir o consumo de energia dos dispositivos corporais e assim aumentar seu tempo de vida, haja vista a sua importância para o monitoramento das pessoas. Além disso, torna-se crucial que estas soluções atendam os critérios de latência máxima estabelecida.

3.6 RESUMO

Ao longo deste capítulo observou-se que a literatura oferece diversos métodos de compressão de dados local para WSNs e WBANs. Primeiramente, foi elucidada uma visão geral da importância do monitoramento em tempo real das redes WBANs na vida das pessoas, o uso de compressão de dados local nas redes WSNs e WBANs, as classes de compressão e o estado-da-arte de maneira abstrata. Em seguida, apresentou-se os trabalhos de compressão de dados local sem perdas, suas principais características e os resultados alcançados por cada trabalho. De mesmo modo, foi discorrido sobre os trabalhos de compressão de dados local com perdas, e sua eficiência nas altas taxas de compressão. Logo após, fez-se a explanação de trabalhos híbridos, com o uso de compressão de dados local com e sem perdas no mesmo método. Por fim, uma análise sobre os trabalhos asseverou a carência de métodos de compressão de dados local com eficiência em energia para aplicações em tempo real nas WBANs.

4 O MECANISMO GROWN PARA COMPRESSÃO DE DADOS LOCAL EM TEMPO REAL EM REDES WBANS

Este capítulo apresenta um mecanismo de compressão de dados local em tempo real energeticamente eficiente para redes WBANs. O mecanismo chamado GROWN (*EnerGy-Efficient Local Data Compression for TRansmission Over WBAN*) gerencia a redundância da informação nos dispositivos corporais e emprega métodos de compressão de dados local encontrados nas WSNs e limiares predefinidos, reduzindo a transmissão das informações e o consumo de energia sem impactar de forma intensa no atraso total da transmissão da informação para o *sink* em WBANs de tempo real. A Seção 4.1 apresenta uma visão geral do mecanismo proposto, suas características, modelo de rede e de comunicação, bem como sua forma de atuação. A Seção 4.2 descreve a arquitetura do mecanismo, detalhando o processo de compressão de dados e descompressão. Ao término da seção é apresentado um exemplo de funcionamento do método.

4.1 VISÃO GERAL

O GROWN atua nos dispositivos corporais e emprega a compressão com e sem perdas por meio de métodos de compressão de dados local encontrados nas WSNs. Logo, os dados transmitidos são comprimidos sem perdas por meio de tabelas de codificação e diante de limiares predefinidos para a compressão com perdas. Assim, apenas determinados valores serão comprimidos para a transmissão. O objetivo do mecanismo é reduzir o consumo de energia dos dispositivos sem afetar drasticamente a latência para a transmissão das informações para o *sink*.

4.1.1 Modelo de rede

O GROWN executa sobre um conjunto de dispositivos corporais portáteis (nós) de sensoriamento de sinais fisiológicos interligados numa rede WBAN composta por um conjunto de dispositivos denotados por $D = \{d_1, d_2, d_3\}$, onde $d_j \in D$ conforme Figura 4.1. Esses nós possuem capacidade de processamento e de comunicação para coletar e disseminar sinais fisiológicos. Cada nó possui um identificador único (Id), imutável no tempo, que o identifica na rede. Um dispositivo d_j coleta um tipo de sinal fisiológico, processa e transmite a informação, através de um meio sem fio, a um dispositivo *sink*. Assume-se que há um controle de perdas de mensagens entre os dispositivos corporais e o *sink* realizado pela tecnologia de comunicação sem fio disponível. Os nós operam de maneira estática ao longo do tempo e formam uma rede com topologia em estrela.

Conforme descrito por Movassaghi *et. al* em (Movassaghi et al., 2014) a arquitetura geral das WBANs possui três níveis de comunicação: *intra-body*, *inter-body* e *beyond-body*. Na comunicação *intra-body* manifesta-se a interação entre os dispositivos e o *sink*. Em vista disso, os dispositivos monitoram, coletam e transmitem os sinais fisiológicos para o *sink* através de tecnologias de comunicação sem fio de curto alcance, como Bluetooth e ZigBee. Uma vez no *sink*, os dados são então processados e encaminhados para o ponto de acesso através da comunicação *inter-body*. O terceiro nível, *beyond-body* possui como principal foco áreas metropolitanas e aplicações específicas (Movassaghi et al., 2014). Neste sentido, o GROWN preocupa-se em melhorias no nível de comunicação *intra-body*, e assim diminuir a quantidade da informação para transmissão ao *sink*.

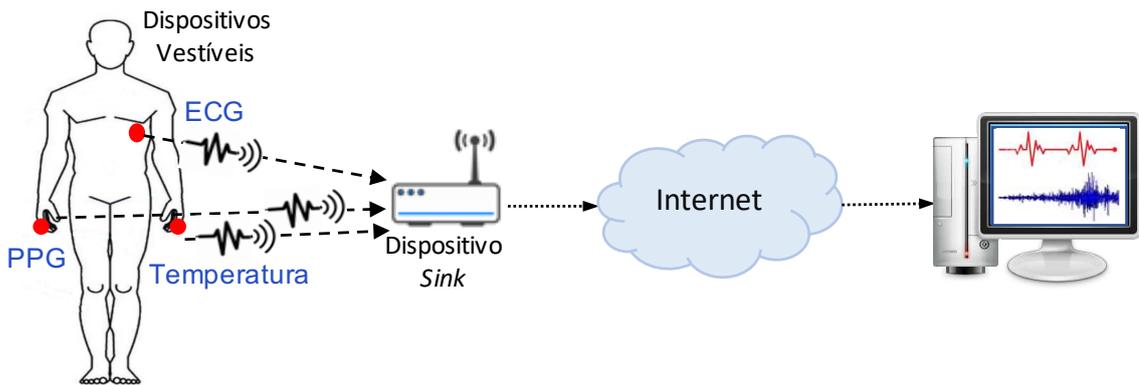


Figura 4.1: Modelo do ambiente de coleta e disseminação de sinais fisiológicos

4.2 ARQUITETURA GROWN

Esta seção apresenta a arquitetura do mecanismo GROWN (*EnerGy-Efficient Local Data Compression for TRansmission Over WBAN*), bem como o funcionamento do mecanismo. Assim, são descritos seus componentes, suas características e a operação de cada módulo. A arquitetura do GROWN é composta pelos módulos (i) **Gerência de Compressão** e (ii) **Gerência de Descompressão**, conforme ilustra a Figura 4.2. O módulo *Gerência de Compressão* localizado nos dispositivos corporais é responsável por gerenciar a aquisição dos sinais fisiológicos coletados dos sensores corporais, analisar os limiares predefinidos estabelecidos para compressão, realizar a compressão e a transmissão dos dados; e o *Gerência de Descompressão* localizado no *sink*, é responsável por receber o dado comprimido pelo módulo anterior, realizar a descompressão, atualizar o último dado recebido de cada dispositivo e apresentar a informação após a descompressão.

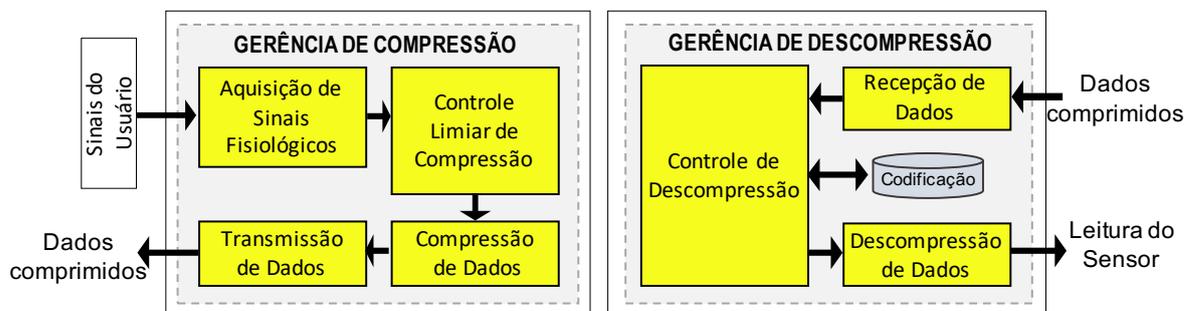


Figura 4.2: Arquitetura da abordagem proposta

4.2.1 Módulo gerência de compressão

Este módulo obtém sinais fisiológicos do usuário, por meio de sensores vestíveis, e converte esses sinais analógicos para digital. Após análise, comprime esses dados e envia para um dispositivo *sink*. Ele é composto por componentes que convertem o sinal fisiológico analógico para digital; analisam o dado digital e definem a forma de compressão adequada; realizam a compressão dos dados; e transmitem esses dados ao dispositivo *sink*, por meio de uma conexão sem fio, para posterior descompressão.

Os sensores vestíveis iniciam sua operação coletando sinais fisiológicos do usuário e, na medida que as leituras são realizadas, são recebidos pelo componente *Aquisição de Sinais*

Algoritmo 1: Controle de Compressão

```

1 procedure CONTROLECOMPRESSAO (leitura_atual)
2   valor_compressao  $\leftarrow$  0
3   if (primeira_leitura == True)
4     leitura_anterior  $\leftarrow$  leitura_atual
5     primeira_leitura  $\leftarrow$  False
6     valor_compressao  $\leftarrow$  leitura_atual
7   else
8     aux_leitura  $\leftarrow$  |leitura_anterior - leitura_atual|
9     if ((t_limiar > 0 AND aux_leitura > t_limiar) OR (t_limiar == 0))
10      valor_compressao  $\leftarrow$  (leitura_atual - leitura_anterior)
11      leitura_anterior  $\leftarrow$  leitura_atual
12    endif
13  endif
14  return (valor_compressao)
15 end procedure

```

Fisiológicos. Esse componente converte o sinal fisiológico analógico para digital e o envia ao componente *Controle Limiar de Compressão*. Conforme descreve o Algoritmo 1, esse componente recebe a leitura atual do sensor (l.1). Se for a primeira leitura do sensor (l.3), armazena essa leitura para comparação com as próximas (l.4), registrando (l.6) e enviando essa leitura para compressão (l.14). Nas próximas leituras, ele verifica a diferença entre a leitura atual e a anterior (l.8), a fim de auxiliar a tomada de decisão sobre a compressão dessa informação. Para minimizar a redundância nas transmissões de informações, esse componente emprega um limiar predefinido estabelecido. Ele verifica se a diferença absoluta entre as leituras atual e anterior é superior ao limiar preestabelecido ou se não foi estabelecido um limiar (l.9). Nessas condições, ele verifica a diferença entre a leitura atual e anterior (l.10), armazena a leitura atual para comparação com as próximas (l.11) e a envia para compressão (l.14). Conforme ilustra a Figura 4.2, o componente *Compressão de Dados* recebe a leitura processada, comprime e encaminha os dados ao componente *Transmissão*. Este componente envia os dados comprimidos ao dispositivo *sink*, por meio da tecnologia sem fio disponível, à medida que eles chegam.

O componente *Compressão de Dados* codifica a informação recebida do componente *Controle Limiar de Compressão* (e_{ts}) modificando a tabela de codificação Exponencial de Golomb (Exp-Golomb) (Teuhola, 1978) de ordem $K = 3$. As amostras recebidas ($e_{ts} \neq 0$) são representadas por uma sequência de bits bs_i , que se divide em duas partes $s_i|a_i$. A primeira parte (s_i) identifica o grupo ao qual e_{ts} pertence e ilustra o valor de n_i , que equivale à quantidade de bits necessária para representar e_{ts} . Assim, os primeiros $2^k - 1$ valores de n_i possuem grupos s_i representados por $\lfloor \log_2(|e_{ts}|) \rfloor + 1$, onde s_i é representado por k bits, conforme ilustra a Tabela 4.1

Tabela 4.1: Tabela com valores dos índices n_i e grupos s_i

n_i	0	1	2	3	4	5	6
s_i	000	001	010	011	100	101	110

Após este nível ($n_i > 6$), o valor do grupo s_i é determinado por $\lfloor \log_2(|e_{ts}|) - 1 \rfloor$, onde os primeiros $s_i - 1$ valores de s_i são representados por 1 seguidos do valor 0. A parte a_i corresponde à representação binária de e_{ts} , que é um código de tamanho inteiro variável. Assim, o valor de e_{ts} é gerado conforme definido em (Marcelloni e Vecchio, 2008), assegurando valores diferentes para as entradas de a_i em cada grupo s_i , conforme as regras disponíveis na Tabela 4.2.

Tabela 4.2: Regras de compressão dos dados

Condição	Regra de compressão
1 - $e_{ts} < 0$	Efetuar complemento de 2 de e_{ts} , subtrair 1 e utilizar os n_i bits menos significativos
2 - $e_{ts} = 0$	Codificar s_i como 000 e não representar a_i
3 - $e_{ts} > 0$	a_i corresponde aos n_i bits menos significativos do complemento de 2 de e_{ts}

Como exemplo, seja C um conjunto de valores $e_{(ts)}$ recebidos pelo componente *Controle Limiar de Compressão*, e C_i o índice de cada valor $e_{(ts)}$ contido no conjunto C . Deste modo, suponha o conjunto $C = \{-7, -9, 5, 10\}$, e um dispositivo com resolução ADC de 8 bits. Cada valor C_i do conjunto C requer $\lceil \log_2(|e_{(ts)}|) \rceil + 1$ bits para sua representação. Logo, cada valor de C_i necessita de (3, 4, 3, 4) bits. Para os valores positivos, aplica-se a primeira regra, obtendo-se a representação de complemento de 2 (5 = 0000101) e (10 = 00001010). Em seguida a_i obtém os n_i bits menos significativos resultantes do complemento de 2. Logo, as representações de a_i para 5 e 10 são (101, 1010). De modo oposto, para os valores negativos, utiliza-se a segunda regra. Assim, primeiramente obtém-se o complementos de 2 de ($-7 = 11111001$) e ($-9 = 11110111$). Após, subtrai-se 1 de cada representação ($-7 = 11111000$) e ($-9 = 11110110$). Por fim, a_i obtém os n_i bits menos significativos resultantes do complemento de 2 das representações de -7 e -9 , 000 e 0110. Esta representação garante que todo valor de a_i possua um valor diferente em seu grupo s_i . A Tabela 4.3 ilustra os valores de a_i para diferentes grupos de s_i .

Tabela 4.3: Representação da tabela com valores de a_i para os primeiros índices de n_i

$n_i = 1$	s_i	índ.	-1	1						
	001	a_i	0	1						
$n_i = 2$	s_i	índ.	-3	-2	2	3				
	010	a_i	00	01	10	11				
$n_i = 3$	s_i	índ.	-7	-6	-5	-4	4	5	6	7
	011	a_i	000	001	010	011	100	101	110	111

A representação da tabela de codificação sem perdas com os índices n_i , grupos s_i e os valores de $e_{(ts)}$ visualiza-se na Tabela 4.4. Em princípio, a tabela de codificação depende da correlação entre consecutivas amostras de $e_{(ts)}$. Assim, diferenças mais frequentes devem ser associadas a códigos mais curtos. Neste sentido, Kolo *et. al* em (Kolo et al., 2012), elucida que dados típicos coletados por WSNs, as diferenças mais frequentes são aquelas próximas a 0. No contexto das redes WBANs, este cenário corresponde a sinais de temperatura, onde, em sua naturalidade, possui maior correlação entra consecutivas amostras. Logo espera-se transmitir menos dados nestes tipos de sinais, e assim, aumentar a vida útil do dispositivo.

De modo geral, o componente *Compressão de Dados* estabelece uma tabela fixa para todos os dispositivos. Assim, a utilização de recursos como armazenamento e processamento é reduzida. Visto que, o método proposto armazena apenas a última amostra como referência e não utiliza tabelas dinâmicas para compressão. Além disso, a principal vantagem da tabela fixa de codificação com Exp-Golomb, é a capacidade de calcular uma versão comprimida de cada amostra em tempo real, que é um dos principais objetivos deste trabalho.

4.2.2 Módulo gerência de descompressão

Este módulo opera no dispositivo *sink*, que detém recursos de energia, armazenamento e processamento para executar a descompressão de dados enviados por dispositivos corporais

Tabela 4.4: Tabela de codificação proposta

n_i	s_i	e_{ts}	bit	byte
0	000	0	3	5
1	001	-1,+1	4	5
2	010	-3,-2,+2,+3	5	5
3	011	-7,-6,-5,-4,+4,+5,+6,+7	6	5
4	100	-15,-14,...,-8,+8,...,+14,+15	7	5
5	101	-31,-30,...,-16,+16,...,+30,+31	8	5
6	110	-63,-62,...,-32,+32,...,+62,+63	9	5
7	11110	-127,-126,...,-64,+64,...,+126,+127	12	5
8	111110	-255,-254,...,-128,+128,...,+254,+255	14	5
9	1111110	-511,-510,...,-256,+256,...,+510,+511	16	5
10	11111110	-1023,-1022,...,-512,+512,...,+1022,+1023	18	5
11	111111110	-2047,-2046,...,-1024,+1024,...,+2046,+2047	20	5

pertencentes a sua WBAN. A recuperação do valor correspondente ao sinal fisiológico coletado e enviado por esses dispositivos inicia no componente *Recepção de Dados*. Após receber a mensagem com os dados comprimidos (bs_i) e a identificação do dispositivo de origem (D_{Id}), ele separa essas informações para uso posterior. Essas informações são manuseadas pelo componente *Controle de Descompressão* (CD), conforme descreve ao Algoritmo 2. Esse componente mantém uma *Lista de Referências* (LR), que armazena a identificação (D_{Id}) de todos os dispositivos corporais (i.e., sensores) conectados à rede WBAN e o valor da sua última leitura (\hat{x}_{ts}). A Tabela 4.5 ilustra a formação da LR. Após consulta à LR com o D_{Id} , obtém-se a \hat{x}_{ts} desse dispositivo (l.2).

Algoritmo 2: Controle de Descompressão

```

1 procedure CONTROLEDESCOMPRESSAO (dados_comprimidos, id)
2   ultimaleitura  $\leftarrow$  obtemultimaleitura(id)
3   prefixo  $\leftarrow$  obtemprefixo(dados_comprimidos)
4   sufixo  $\leftarrow$  obtemsufixo(dados_comprimidos)
5   valor_dado  $\leftarrow$  bdcodificacao(prefixo, sufixo)
6   return (valor_dado, id)
7 end procedure

```

O próximo passo do componente CD é obter o prefixo (s_i) (l.3) e o sufixo (l.4) para o bs_i recebido. Em seguida, ele consulta a base de dados *Codificação* (l.5), que armazena os índices n_i , os grupos s_i e seus valores a_i . Cada valor de a_i é exclusivo em todo grupo s_i , permitindo a recuperação da informação por meio de poucas instruções, adequando-se a uma operação em tempo real. Após consultar a base de dados *Codificação* o componente CD obtém o valor correspondente da informação bs_i , adquire à última leitura do dispositivo D_{Id} salva na *Lista de Referências* (\hat{x}_{ts}), atualiza à LR e encaminha os valores para o componente *Descompressão de Dados*. Este componente calcula a diferença entre \hat{x}_{ts} antes da atualização da LR e o valor obtido de bs_i . Finalmente, o valor lido pelo sensor é disponibilizado para uso.

4.2.3 Funcionamento do GROWN

Considere um ambiente de WBAN, onde uma pessoa possua um dispositivo vestível de sensoriamento de temperatura e que este dispositivo interage com um *smartphone* atuando

Tabela 4.5: Referências dos dispositivos da rede

	D_{Id}				
	1	2	3	4	5
\hat{x}_{ts}	8	2	-4	9	7

como o *sink* através de uma tecnologia de comunicação sem fio de curto alcance, como ilustra a Figura 4.3. O sensor inicia sua operação obtendo o sinal fisiológico de temperatura da pessoa e o converte de analógico para digital em uma representação binária (x_{ts}) com R bits através do componente *Aquisição de Sinais Fisiológicos*. Tendo em conta que o valor lido foi 38, então x_{ts} em 8 bits torna-se 00100110. A tomada de decisão sobre a compressão da informação depende do componente *Controle Limiar de Compressão* que possui um filtro de predição (T) pré-definido. Assumindo $T = 1$, esse filtro verifica o valor de x_{ts} , e como trata-se da primeira leitura do sensor, a leitura anterior (\hat{x}_{ts}) recebe o valor 0 (zero). Dessa forma, o erro de predição (e_{ts}), que corresponde à diferença entre leituras sucessivas, é $e_{ts} = x_{ts} - \hat{x}_{ts} = 38$.

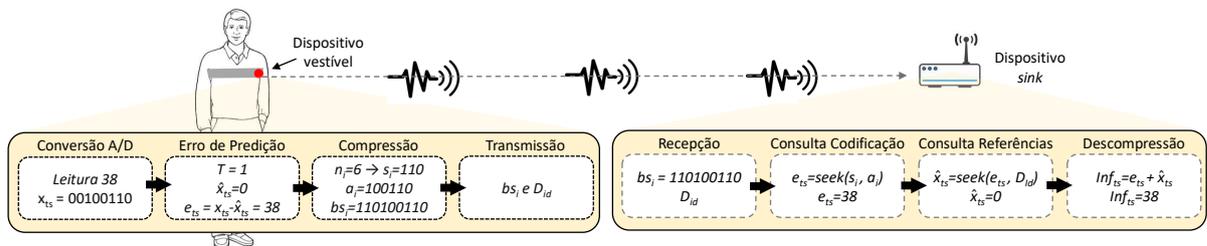


Figura 4.3: Exemplo de funcionamento

O valor de e_{ts} é encaminhado para o componente *Compressão de Dados*, a fim de ser comprimido em uma sequência de bits (bs_i) composta de duas partes $s_i|a_i$. s_i representa o grupo (n_i) ao qual e_{ts} pertence. Assim, $n_i = \lceil \log_2(|38|) \rceil + 1 = 6$. Consultando-se a tabela dos valores de codificação (Tabela 4.4), verifica-se que s_i pertence ao grupo 110. Em seguida, obtém-se a_i conforme a regras de compressão da Tabela 4.2. Neste caso, o valor de e_{ts} é superior a zero. Assim, para $n_i = 6$ utiliza-se os 6 bits menos significativos do complemento de 2 da representação binária de e_{ts} , resultando em $a_i = 100110$ e $bs_i = 110100110$. Em seguida, a informação bs_i é transmitida ao *sink* com a identificação do dispositivo vestível (D_{Id}). Já o *sink* ao receber a mensagem do dispositivo vestível verifica sua identificação (D_{Id}) e os dados recebidos (bs_i). Os dados são decompostos, obtendo-se $s_i = 110$ e $a_i = 100110$. Consultando-se a *Tabela de Codificação* com s_i e a_i , obtém-se o erro de predição (e_{ts}). Em seguida, uma consulta à *Tabela de Referências* usando e_{ts} e D_{Id} retorna o último valor recebido desse dispositivo, $\hat{x}_{ts} = 0$. Finalmente, a informação (Inf_{ts}) correspondente ao sinal do dispositivo vestível é recuperada, $Inf_{ts} = e_{ts} + \hat{x}_{ts} = 38$.

Considere uma segunda coleta do sinal de temperatura com conversão de analógico para digital no valor 37. Logo, temos a representação binária de $x_{(ts)} = 00100101$. Este valor passa pelo filtro de predição, que verifica a diferença entre à leitura atual x_{ts} e à última leitura transmitida \hat{x}_{ts} . Assim, temos $e_{ts} = x_{ts} - \hat{x}_{ts} = -1$. Neste caso, conforme descrito no Algoritmo 1 obtém-se o valor absoluto 1. Este valor não é superior ao limiar predefinido $T = 1$. Assim, a segunda leitura será representada como 0|000 e não será transmitida,

Uma terceira leitura adquirida pelo dispositivo obteve valor 36. Da mesma forma, temos a representação binária da informação $x_{(ts)} = 00100100$. Esta informação passa pelo filtro de predição, e obtém-se o valor do erro de predição (diferença entre as leituras) $e_{ts} = x_{ts} - \hat{x}_{ts} = -2$. Assim, o valor absoluto do erro de predição é superior ao limiar predefinido $T = 1$. Logo,

o *Controle Limiar de Compressão* atualiza à última informação transmitida \hat{x}_{ts} e encaminha para o componente *Compressão de Dados* o erro de predição e_{ts} . No componente *Compressão de Dados*, $e_{(ts)}$ obtém o grupo $s_i = 010$ através de $\lfloor \log_2(|-2|) \rfloor + 1 = 2$. Para comprimir valores negativos, utiliza-se a 1ª regra da Tabela 4.2. Logo, temos o complemento de 2 de $e_{(ts)}$ com resolução de 8 bits 11111110, subtrai-se 1 do resultado 11111101 e utiliza-se os ni bits menos significativos. Tendo como resultado da expressão o valor 01. Assim, é obtida a representação $bs_i = 01001$ para transmissão. Após o *sink* receber a transmissão do dispositivo vestível ele verifica a identificação (D_{Id}) e os dados recebidos (bs_i). Os dados são decompostos, obtendo-se $s_i = 010$ e $a_i = 01$. Consultando-se a *Tabela de Codificação* com s_i e a_i , obtém-se o erro de predição (e_{ts}). Em seguida, uma consulta à *Tabela de Referências* usando e_{ts} e D_{Id} retorna o último valor recebido desse dispositivo, $\hat{x}_{ts} = 28$. Finalmente, a informação (Inf_{ts}) correspondente ao sinal do dispositivo vestível é recuperada, $Inf_{ts} = e_{ts}(-2) + \hat{x}_{ts}(38) = 36$. A Tabela 4.6 ilustra a representação de cada valor ao longo do tempo na transmissão das informações.

Tabela 4.6: Representação dos exemplos ao longo do tempo com limiar $T = 1$

Leitura	Dispositivo Vestível					Dispositivo sink				
	x_{ts}	\hat{x}_{ts}	e_{ts}	si	ai	bs_i	D_{Id}	e_{ts}	\hat{x}_{ts}	Inf_{ts}
1	28	0	28	110	100110	110100110	1	28	0	28
2	27	28	-1	0	000	0	1	0	28	28
3	26	28	-2	010	01	01001	1	-2	28	26

4.3 RESUMO

Este capítulo apresentou o mecanismo GROWN para gerência de compressão de dados local em tempo real energeticamente eficiente para redes WBANs. Inicialmente, apresentou-se uma visão geral do mecanismo e onde o GROWN atua numa WBAN. Em seguida, detalhou-se cada componente da arquitetura GROWN para compressão e descompressão. Por fim, o capítulo detalhou o funcionamento do GROWN e como ele atua na redução da redundância da informação e transmissão.

5 AVALIAÇÃO

Este capítulo detalha a avaliação de desempenho do GROWN e objetiva analisar sua eficiência energética numa WBAN. A avaliação considera os resultados obtidos a partir do funcionamento do mecanismo no ambiente de experimentação onde ele foi implementado. As métricas buscam comprovar a eficácia e eficiência do GROWN para reduzir o consumo energético dos dispositivos e seu uso em aplicações WBAN de tempo real. A Seção 5.1 descreve a implementação do GROWN. A Seção 5.2 apresenta o cenário experimentado, suas configurações, e as métricas usadas na avaliação. A Seção 5.3 descreve os resultados alcançados acompanhados das respectivas análises. Por fim, a Seção 5.4 faz uma discussão dos resultados alcançados.

5.1 IMPLEMENTAÇÃO

O GROWN foi implementado a partir de diferentes plataformas de desenvolvimento para atender às necessidades dos dispositivos corporais, do *sink* e aproximar ao máximo a experimentação de um ambiente real. Seu código-fonte está disponível no GitHub¹ para futuros downloads, comparações e experimentações. O módulo de descompressão, localizado no *sink*, foi implementado no Android Studio², versão 3.6.1. O Android Studio é um ambiente de desenvolvimento integrado (IDE) oficial para desenvolvimento de *apps* Android, onde desenvolveu-se um aplicativo para realizar a conexão com os dispositivos e a descompressão. O aplicativo desenvolvido possui três etapas de operação conforme ilustra a Figura 5.1.

A primeira etapa refere-se no *scan* de dispositivos que suportam a tecnologia *Bluetooth Low Energy* (BLE). O padrão BLE (SIG, 2014) é uma versão otimizada do tradicional *Bluetooth* para dispositivos com recursos limitados. Ele possibilita reduzir o consumo de energia na transmissão em dispositivos que não necessitam grandes volumes de dados, o que o torna adequado para monitoramento nas WBANs (Ghamari et al., 2016). Além disso, a escolha por esta tecnologia deu-se pela facilidade de encontrar módulos e componentes BLE para os dispositivos corporais, baixo custo e a presença desta tecnologia em diversos *smartphones* do mercado. A segunda etapa considera escolher o dispositivo para o pareamento e início do recebimento das informações. Por fim, a terceira etapa ilustra o dispositivo corporal conectado e a aplicação demonstrando o valor recebido e a informação correta após consulta à *Tabela de Referências*. Este aplicativo foi instalado em um celular Motorola, modelo Moto G⁴ Plus, que opera com o sistema operacional Android, versão 7.0.

A implementação dos módulos de compressão com e sem perdas ocorreu através da linguagem C++ e compilada para placas Arduino³ Uno e Mega ilustradas na Figura 5.2. Cada placa Arduino possui um microcontrolador ATmega de 16 MHz, mesmo microcontrolador utilizado em diversos dispositivos vestíveis (Nakayama et al., 2019). Elas operam com uma tensão de 5 V e possuem 10 bits de resolução em cada porta analógica para leitura dos sensores fisiológicos. Todos os componentes de hardware seguem as versões padrão do fabricante, i.e., não sofreram qualquer alteração física. Além disso, para aprimorar a leitura dos sinais, os dispositivos foram compilados com as bibliotecas oficiais de cada sensor. As demais características de cada placa são apresentadas na Tabela 5.1.

¹<https://github.com/cainapassos/GROWN>

²<https://developer.android.com/studio>

³<https://www.arduino.cc/>

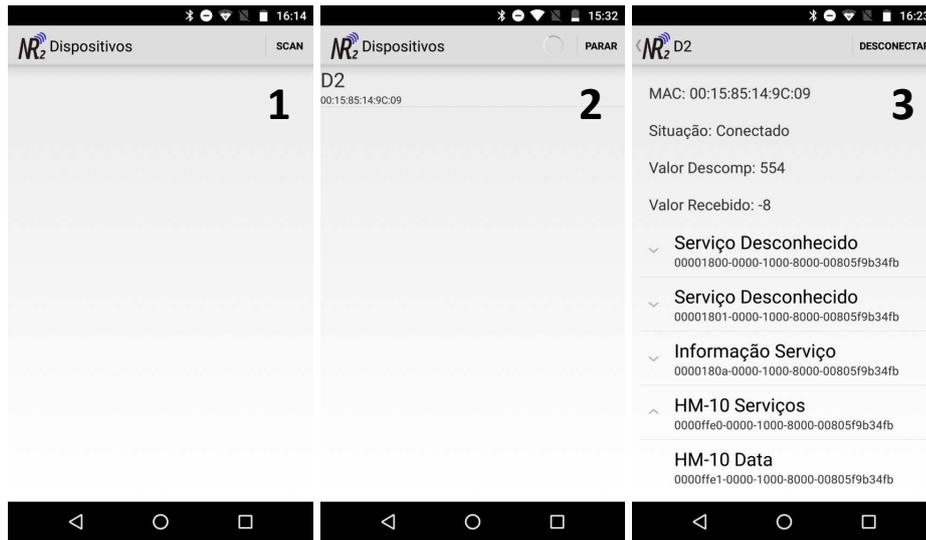


Figura 5.1: Etapas do aplicativo no dispositivo *sink*

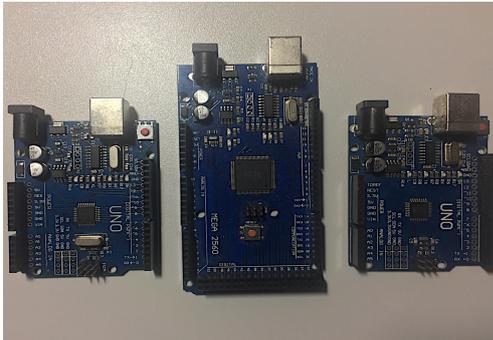


Figura 5.2: Placas Arduino utilizadas no *testbed*

Tabela 5.1: Especificações Técnicas

Característica	UNO	MEGA
Microcontrolador	ATmega328P	ATmega2560
Memória Flash	32 KB	256 KB
SRAM	2 KB	8 KB
EEPROM	1 KB	4 KB
Corrente por pino E/S	20 mA	50 mA

Para implementação dos dispositivos corporais cada placa Arduino obteve um sensor fisiológico que corresponde ao sensoriamento de Temperatura, ECG (eletrocardiograma) e PPG (fotopletismograma). Além disso, cada dispositivo detém de componentes essenciais para funcionamento conforme ilustra a Figura 5.3. Todos os dispositivos foram produzidos em *Protoboards* Mb102 de 830 pontos. As *protoboards* fornecem energia para todos os sensores e placas Arduino, responsáveis pela gerência dos sensores e do GROWN na compressão. Elas são alimentadas por fontes de tensão ajustável reguladas para operar em 5 V. Cada fonte é alimentada por uma bateria de 9 V com capacidade energética de 400 mAh. Para comunicação entre o dispositivo e o *sink* usou-se um módulo BLE HM-10. Este módulo utiliza o chip CC2540⁴ BLE (Texas Instrument) para transmissão e recepção das informações. Por fim, cada dispositivo dispõe de um sensor de corrente modelo INA219 (Instruments, 2020) para aferir seu consumo energético.

O dispositivo Temperatura utilizou uma placa Arduino Uno para gerência do GROWN. Além dos componentes essenciais, a coleta do sinal fisiológico ocorreu através do sensor Melexis MLX90614⁵. Este sensor é um termômetro infravermelho para medições de temperatura sem contato de alta precisão com 17 bits de resolução. A comunicação entre o sensor e a placa Arduino ocorreu através dos pinos SDA e SCL com o protocolo I²C. Este protocolo permite a comunicação entre circuitos integrados em uma mesma placa com somente duas vias (Kumari e

⁴<http://www.ti.com/lit/ds/symlink/cc2540.pdf?ts=1588109113783>

⁵<https://www.melexis.com/en/product/mlx90614/digital-plug-play-infrared-thermometer-to-can>

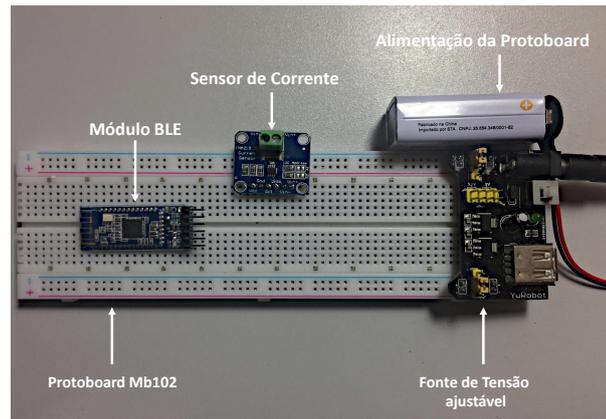


Figura 5.3: Componentes essenciais para os dispositivos corporais

Gayathri, 2017). A Figura 5.4 ilustra o dispositivo Temperatura e todos os seus componentes para funcionamento.

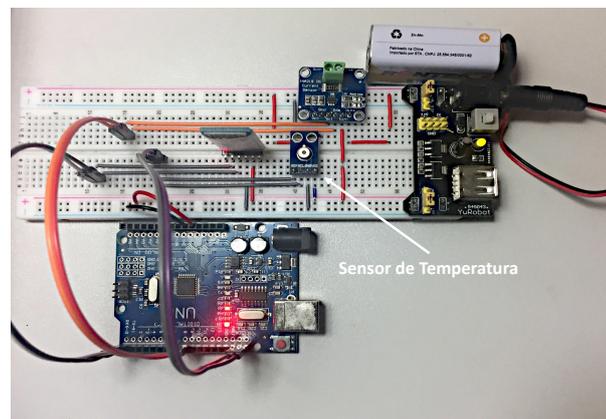


Figura 5.4: Dispositivo Temperatura

Para gerência do GROWN no dispositivo ECG utilizou-se a placa Arduino Mega. Além disso, todo sinal fisiológico foi coletado por um Módulo Sensor Cardíaco ECG AD8232⁶ utilizado em diversas pesquisas científicas (Jayarathna et al., 2020; Kalaivani et al., 2020; Lakshmanan et al., 2020). Este módulo foi projetado para extrair, amplificar e filtrar pequenos sinais através de seus eletrodos remotos (Lakshmanan et al., 2020). Estes eletrodos são pré-fixados no usuário e definidos pelos pontos braço direito (do inglês, *Right Arm* - RA), braço esquerdo (do inglês, *Left Arm* - LA) e perna esquerda (do inglês, *Left Leg* - LL). A Figura 5.5 demonstra o dispositivo ECG com seus eletrodos e componentes para funcionamento.

O dispositivo PPG tomou como base o sensor da empresa PulseSensor⁷. Este sensor coleta os sinais através de um processo em que um transmissor LED emite luminosidade em alguma parte do corpo humano como pulso ou dedo, e paralelamente um diodo fotossensível mede a intensidade da luz absorvida pelos tecidos apiteliais (Nakayama et al., 2019; Tavares et al., 2020). Assim, se a quantidade de luz incidente no sensor permanecer constante, o valor do sinal permanece próximo de 512, ponto médio da resolução ADC de 10 bits do dispositivo Arduino. Logo, as medições apresentam as mudanças no volume sanguíneo (Liang et al., 2018). A Figura 5.6 ilustra o dispositivo PPG, o sensor captando a luminosidade através do LED e os demais componentes.

⁶<https://cdn.sparkfun.com/datasheets/Sensors/Biometric/AD8232.pdf>

⁷<https://pulsesensor.com/>

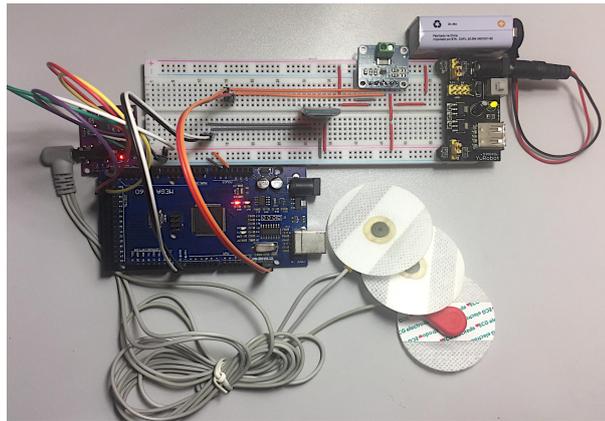


Figura 5.5: Dispositivo ECG

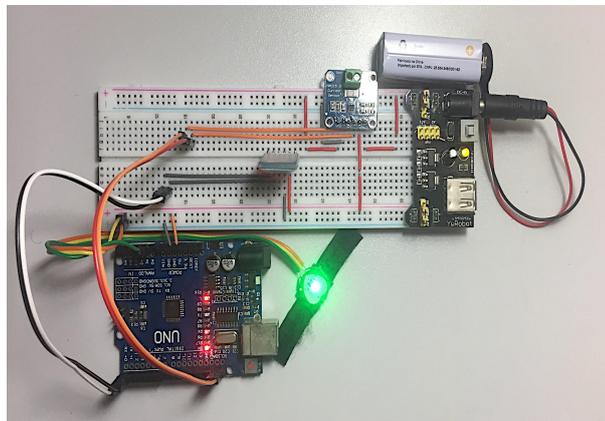


Figura 5.6: Dispositivo PPG

Para aprimorar a análise do GROWN, criou-se um quarto dispositivo (**ECGBD**) com um módulo SD e uma placa Arduino Mega. O módulo SD permite a leitura de dados originados de um cartão de memória com capacidade de 16GB. Este cartão possui uma base dados ECG (MIT-BIH Arrhythmia⁸) em modo texto. A base de dados é composta por 48 trechos de meia hora de sinais ECG obtidos de 47 indivíduos com resolução de 11 bits. Assim, a placa Arduino Mega responsável pela gerência do dispositivo realiza leituras no cartão SD e obtém os valores de cada linha da base de dados. Por fim, estes valores são comprimidos pelo GROWN e transmitidos para o dispositivo *sink* para descompressão. A Figura 5.7 demonstra o dispositivo (**ECGBD**).

5.2 AMBIENTE DE EXPERIMENTAÇÃO E MÉTRICAS

O ambiente de experimentação (*Testbed*) avaliado consistiu de quatro dispositivos corporais fixos e um dispositivo celular atuando como *sink* da rede WBAN, conforme ilustrado na Figura 4.1. Esses dispositivos corporais correspondem ao sensoriamento de Temperatura, ECG, e PPG. A comunicação entre os dispositivos corporais e o *sink* utilizou a tecnologia *Bluetooth Low Energy* (BLE), estabelecendo uma rede *ad-hoc* no padrão IEEE 802.15.4, por um período de operação de 600s. Este período de tempo é suficiente para a obtenção de 600 amostras de corrente instantânea de cada dispositivo. Assim, os dispositivos possuem um módulo de comunicação sem fio *bluetooth* modelo HM-10 BLE V4. Durante todo o experimento, o celular permaneceu em modo avião para evitar interrupções com outros processos, permanecendo habilitada apenas

⁸<https://alpha.physionet.org/content/mitdb/1.0.0/>

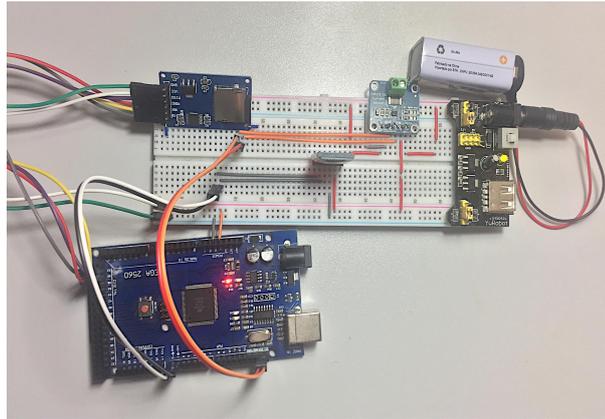


Figura 5.7: Dispositivo ECGBD

a tecnologia de comunicação sem fio *Bluetooth* BLE. Por fim, diante dos resultados obtidos, os dados foram analisados através da ferramenta GNU PLOT (Williams et al., 2020), versão 5.2. Os resultados exibidos são média dos resultados obtidos a partir de 10 repetições de experimentação e um intervalo de confiança de 95%.

Para avaliação do GROWN na redução do consumo energético e sua eficácia para aplicações WBAN de tempo real, os dispositivos corporais foram pré-configurados de três maneiras distintas. Assim, as configurações foram:

- *Configuração 1*: O funcionamento dos dispositivos ocorre sem a atuação do GROWN. Assim, todos os dados coletados pelos sensores corporais são transmitidos diretamente para o *sink*. Esta configuração representa uma situação real onde não existe a execução de rotinas de compressão nos dispositivos e descompressão pelo *sink*. Logo, encontra-se o consumo real de energia de todos os dispositivos, bem como a latência total para a transmissão das informações.
- *Configuração 2*: O funcionamento dos dispositivos ocorre com o GROWN e o filtro de predição T do método de compressão com perdas pré-estabelecido em 1 para todos os dispositivos. Logo, toda diferença entre amostras consecutivas superior a 1 será encaminhada para compressão e transmissão. Assim, espera-se obter redução no consumo de energia dos dispositivos, visto que, os sinais corporais apresentam certa correlação entre amostras consecutivas. Além disso, a latência deve sofrer alteração, dado que o processo de compressão e descompressão demanda um certo período de tempo.
- *Configuração 3*: Operação semelhante a Configuração 2. Porém, o filtro de predição T do método de compressão com perdas é pré-estabelecido com 0 (zero) para todos os dispositivos. Logo, todos os dados serão comprimidos sem perdas, e a informação transmitida será a compressão da diferença entre amostras consecutivas dos sensores. Assim como a Configuração 2, espera-se uma redução no consumo de energia dos dispositivos e um aumento na latência visto o uso dos processos de compressão e descompressão do GROWN.

A Figura 5.8 ilustra a operação de coleta e transmissão do dispositivo Temperatura no ambiente de experimentação com o GROWN. Neste ambiente, o usuário mantém o dedo indicador próximo ao sensor MLX90614 pré-fixado na *proto-board* Mb102, enquanto o sensor realiza a leitura do sinal por infravermelho, sem a necessidade do contato direto. O dispositivo

Temperatura recebe a leitura, realiza a compressão e encaminha ao módulo *Bluetooth* para transmissão. Este módulo transmite o dado comprimido ao *sink*, que realiza a descompressão.

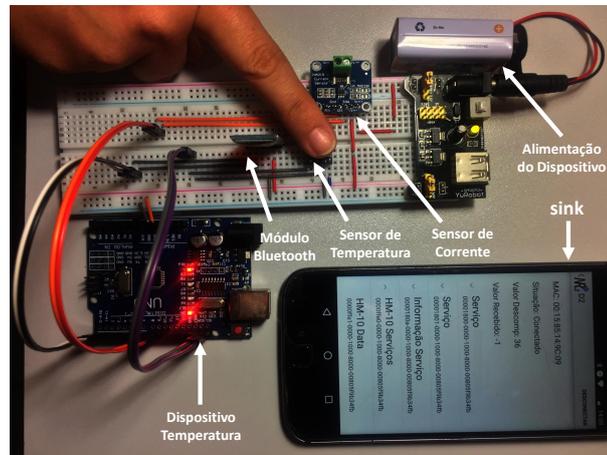


Figura 5.8: Testbed do dispositivo Temperatura

A avaliação do GROWN ocorreu a partir da análise dos resultados obtidos em experimentações por métricas definidas previamente. Estas métricas visam refletir o comportamento do mecanismo quanto ao tempo de compressão e descompressão dos sinais fisiológicos, o tempo médio para transmissão de uma informação, o consumo energético dos dispositivos e a taxa de compressão para cada sinal fisiológico. Assim, as métricas escolhidas para avaliar o mecanismo foram definidas com base nos trabalhos de (Patel et al., 2013) para aferir o **Atraso de Compressão** e o **Atraso de Descompressão** dos dados, de (Vergutz et al., 2017) para aferir o **Atraso Médio** de cada rodada de experimentação, e (Marcelloni e Vecchio, 2009) para mensurar a taxa de compressão dos pacotes transmitidos. O **Consumo Energético do Dispositivo** foi aferido por um sensor de corrente e tensão desenvolvido pela Texas Instruments, modelo INA219 (Instruments, 2020). Este sensor monitora a queda de tensão ocorrida sobre um resistor *shunt* de 1 m Ω , que é proporcional à corrente que circula através do resistor, conforme a lei de Ohm (Parks, 2007). Para maior precisão, a calibração do sensor de corrente seguiu as recomendações do seu *datasheet*.⁹ As métricas para eficácia do GROWN são descritas abaixo:

- **Atraso de Compressão (AC)** indica o tempo de processamento de compressão da informação. Ela equivale à razão entre o somatório das diferenças entre o instante de término da compressão (T_{fcom}^i) e o instante de início da compressão (T_{icom}^i), e o total de amostras transmitidas (Inf_{ts}).

$$AC = \sum_{i=1}^n \frac{T_{fcom}^i - T_{icom}^i}{Inf_{ts}^i} \quad (5.1)$$

- **Atraso de Descompressão (AD)** indica o tempo de processamento de descompressão da informação. Ela corresponde à razão entre o somatório das diferenças entre o instante de término de descompressão (T_{fdes}^i) e o instante de início da descompressão (T_{ides}^i), e Inf_{ts} .

$$AD = \sum_{i=1}^n \frac{T_{fdes}^i - T_{ides}^i}{Inf_{ts}^i} \quad (5.2)$$

⁹<http://www.ti.com/lit/ds/symlink/ina219.pdf>

- **Atraso Médio (AM)** computa a média do somatório dos atrasos de compressão (**AC**) Eq. (5.1), transmissão (**ATR**) e descompressão (**AD**) Eq. (5.2) em cada experimentação y , conforme o total de transmissões de dados ocorridas (T_{tr}), por um dispositivo x durante todas as experimentações (N_e).

$$AM = \sum_{x=1}^{N_e} \sum_{y=1}^{T_{tr}} \frac{AC_{xy} + AD_{xy} + ATR_{xy}}{T_{tr} \times N_e} \quad (5.3)$$

- **Consumo Energético do Dispositivo (CED)** indica o consumo de corrente elétrica por um dispositivo durante o período de coleta, processamento e transmissão dos dados. Ela equivale à multiplicação da corrente consumida (I) pelo tempo (t).

$$CED = I \times t \quad (5.4)$$

- **Taxa de Compressão (TC)** indica a redução da quantidade de pacotes transmitidos com o uso da compressão. Ela equivale à multiplicação por 100 do resultado da subtração de 1 da razão entre a quantidade de pacotes comprimidos transmitidos ($comp_pkt$) e a quantidade de pacotes transmitidos sem a compressão ($orig_pkt$).

$$TC = \left(1 - \frac{comp_pkt}{orig_pkt} \right) \times 100 \quad (5.5)$$

5.3 RESULTADOS E ANÁLISES

Esta seção apresenta a avaliação de desempenho do mecanismo GROWN quanto à sua eficiência energética para reduzir o consumo de energia dos dispositivos, sua eficácia para aplicações WBAN de tempo real e as taxas de compressão alcançadas na redução dos pacotes transmitidos. Os resultados foram obtidos a partir de 10 repetições de experimentação de cada configuração apresentada na Seção 5.2 nos dispositivos corporais. Assim, as subseções demonstram os resultados para cada métrica avaliada através de gráficos e tabelas com os intervalos de confiança de 95%.

5.3.1 Atraso de compressão e descompressão

A compressão e a descompressão dos dados impactam diretamente a latência na disseminação e apresentação das informações, visto que impõem um processamento adicional nos dispositivos corporais e no dispositivo *sink*. As variações nos atrasos durante a compressão (**AC**) e descompressão dos dados (**AD**) ao longo do tempo são apresentadas na Figura 5.9. Estas variações apresentadas pelo AC são devidas à correlação entre amostras consecutivas. Deste modo, quanto maior a correlação menor é a diferença entre as amostras. Assim, o GROWN atribui códigos mais curtos para compressão e diminui seu tempo de atraso. Observa-se mais intensidade desta correlação no sinal de temperatura, que apresentou um atraso médio de compressão de ≈ 1 ms, conforme visto no Gráfico 5.9(g), visto que este sinal apresenta poucas alterações ao longo do tempo. Em contrapartida, o sinal ECG apresentou o maior atraso médio de compressão com ≈ 3 ms, conforme visto no Gráfico 5.9(c). Assim, percebe-se a diferença da correlação entre amostras consecutivas após os 300 segundos de experimentação. Em sua totalidade, o sinal ECG apresentou queda na correlação e aumentou ≈ 2 ms para comprimir uma informação quando comparado com o sinal de temperatura.

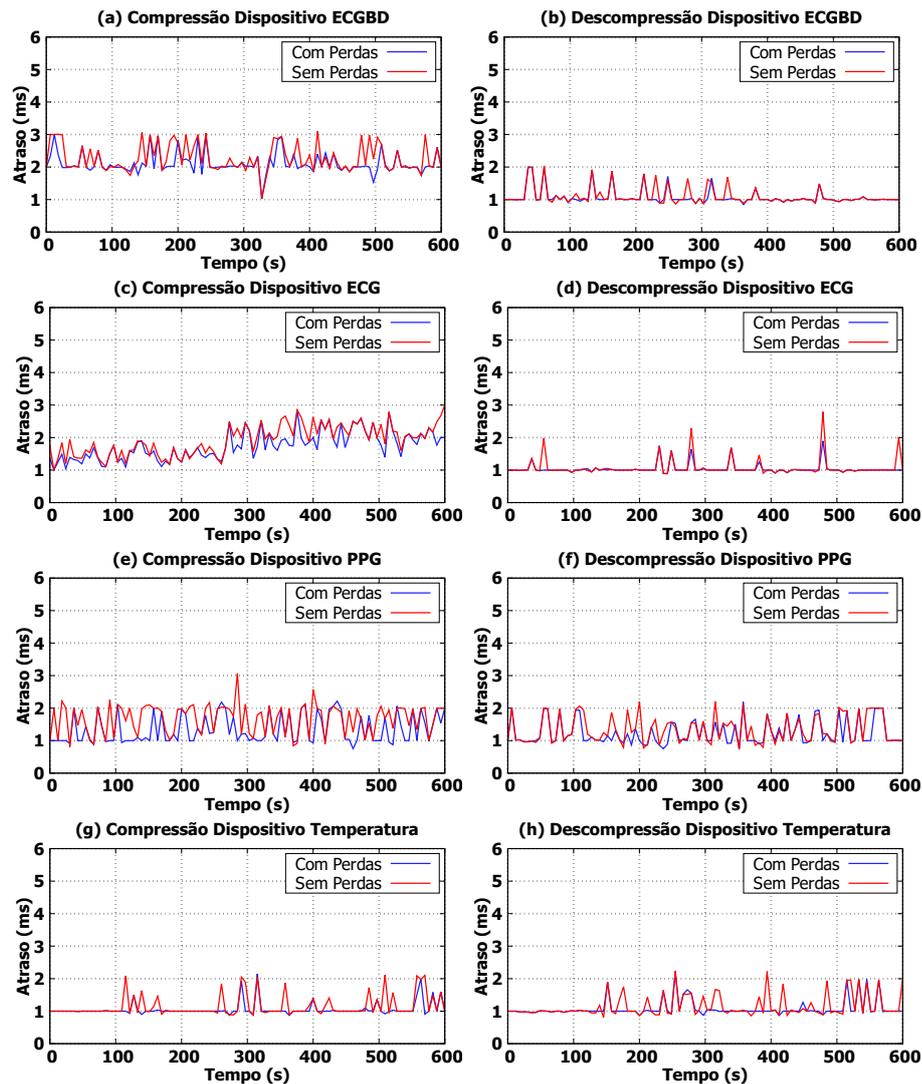


Figura 5.9: (AC) e (AD) alcançado pelo GROWN nos sinais sensorizados

No mesmo sentido, o sinal PPG apresentou um atraso médio de compressão de $\approx 2\text{ms}$, como visto no Gráfico 5.9(e). Verifica-se que o sinal PPG apresenta variações ao longo de toda à experimentação. Logo, a correlação entre amostras consecutivas é afetada e o tempo médio de compressão aumenta. De modo similar, o GROWN obteve o mesmo comportamento no dispositivo ECGBD, que possui a base de dados real, com média de $\approx 3\text{ms}$ na compressão. Este comportamento repete-se no dispositivo *sink*. Contudo, devido ao seu maior poder computacional, e a consulta direta na *Tabela de Codificação* para obter os valores de s_i e a_i a descompressão impõe um atraso menor, com $\approx 1\text{ms}$ para todos os dispositivos, como exibidos nos Gráficos 5.9(b), 5.9(d) e 5.9(h).

No ambiente de experimentação as configurações 2 e 3 representam o uso do GROWN com perdas e sem perdas respectivamente. Verifica-se que em todos os Gráficos o método com perdas (configuração 2) apresenta uma queda no atraso para compressão e descompressão em relação ao método sem perdas (configuração 3). Constata-se que os Gráficos 5.9(a), 5.9(c) e 5.9(e) possuem as maiores diferenças entre os dois métodos. Logo, percebe-se que o GROWN identifica a correlação entre amostras consecutivas, e através do filtro de predição T , ignora informações inferiores a 1 para transmissão. Assim, menos informações são encaminhadas para compressão.

5.3.2 Atraso médio

O Atraso médio (**AM**) deve-se a compressão, transmissão e descompressão com o uso do GROWN sendo exibido no gráfico da Figura 5.2. O impacto da baixa correlação entre amostras consecutivas é determinante para o aumento do tempo de compressão nos dispositivos com poucos recursos computacionais. Observa-se este impacto principalmente nos dispositivos ECG, ECGBD e PPG com perdas (**CP**) e sem perdas (**SP**), que apresentaram o maior atraso no tempo de compressão e descompressão ≈ 3 ms, contra ≈ 1 ms do dispositivo Temperatura (**CP**) e (**SP**). Os resultados mostram que a transmissão dos dados impôs um atraso médio de ≈ 49 ms a todos os dispositivos. Durante o período das experimentações, verificou-se que este atraso ocorre até a disponibilidade da informação no *sink*, com e sem o emprego do GROWN. Esse comportamento é amplamente explicitado no trabalho de (Gatouillat et al., 2018), que verificou um intervalo entre amostras de $\approx 48,75$ ms em conexões *Bluetooth* BLE no Android 7.0. Em geral, o emprego do GROWN aumentou em 7,84% o tempo total de transmissão para os dispositivos ECG e ECGBD, 6% para o PPG, e 2,04% para o de Temperatura. Assim, constata-se que o tempo de compressão e descompressão não afeta drasticamente o atraso entre amostras consecutivas. Este resultado demonstra a aplicabilidade do GROWN nas redes WBANs de tempo real, que podem apresentar uma latência máxima de 125ms para aplicações médicas e 250ms para aplicações não médicas (Movassaghi et al., 2014).

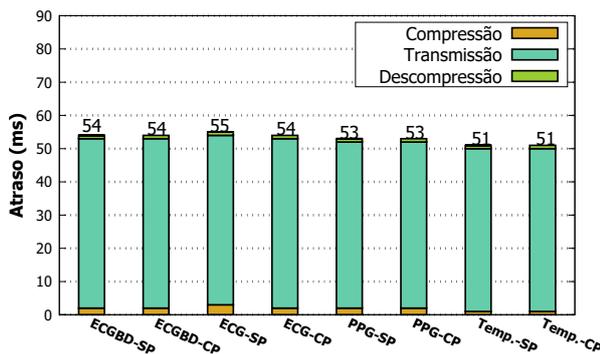


Tabela 5.2: Atraso médio dos dispositivos sensorizados (**AM**)

Tabela 5.3: Intervalo de Confiança do (**AM**)

Dispositivo	Valores Compressão			Valores Descompressão		
	IC	Val. Mín	Val. Máx	IC	Val. Mín	Val. Máx
ECGBD-SP	0,053446	2,9886	3,0954	0,037399	1,2796	1,3544
ECGBD-CP	0,027551	2,9675	3,0565	0,017187	1,2433	1,2987
ECG-SP	0,052504	2,9955	3,1005	0,020742	1,2793	1,3207
ECG-CP	0,033612	2,9594	3,0266	0,024196	1,2318	1,2802
PPG-SP	0,020973	2,0140	2,0560	0,019442	1,1666	1,2054
PPG-CP	0,010753	1,9822	2,0038	0,006930	1,1581	1,1719
Temp-SP	0,040091	1,0059	1,0861	0,007937	1,0761	1,0919
Temp-CP	0,018017	0,9970	1,0330	0,009682	1,0563	1,0757

5.3.3 Consumo energético do dispositivo

O emprego do GROWN com perdas (**CGCP**) e sem perdas (**CGSP**) reduziu o consumo de energia dos dispositivos (**CED**) através da compressão e correlação entre as amostras consecutivas, conforme visto no gráfico da Figura 5.4. Observa-se pelo gráfico que o dispositivo Temperatura sem o GROWN (**CSG**) consumiu, em média, 38,28mAh, e com os métodos **CGCP** e **CGSP**, ele consumiu 36,65mAh e 36,81mAh, respectivamente. Logo, constata-se uma diferença de 1,63mAh **CGCP** e de 1,47mAh **CGSP** sobre o consumo do **CSG**, que indicam a eficiência energética que o dispositivo obteve com o emprego do GROWN. A eficiência no dispositivo ECG alcançou 1,02mAh para ambos os métodos. Isto representa uma queda na correlação entre amostras consecutivas para sinais ECG. Já no dispositivo PPG, que possui o sinal com menor correlação entre amostras consecutivas, o GROWN gerou uma eficiência energética de 0,89mAh no **CGCP** e de 0,31mAh no **CGSP**. No ambiente de experimentação, a modificação do dispositivo ECGBD para realizar consecutivas leituras de amostras ECG no módulo SD aumentou o consumo energético. Esta diferença corresponde ao consumo de ≈ 16 mA do módulo SD que

permanece ativo em todo o período de experimentação, visto que o módulo ECG AD8232 segundo seu *datasheet*¹⁰ consome apenas $\approx 170 \mu\text{A}$. Assim, o dispositivo sem o GROWN **CSG**, consumiu, em média, 73,29mAh ao longo das experimentações e obteve uma eficiência energética de 1,14mAh **CGCP** e de 0,90mAh **CGSP**. Ao obter uma eficiência energética propicia-se o aumento do tempo de vida dos dispositivos, como apresentado na Tabela 5.6. Deste modo, o GROWN aumentou o tempo de vida de todos os dispositivos, sobretudo do dispositivo Temperatura, cujo sinal apresenta a maior correlação entre amostras consecutivas.

Tabela 5.5: Intervalo de Confiança do (CED)

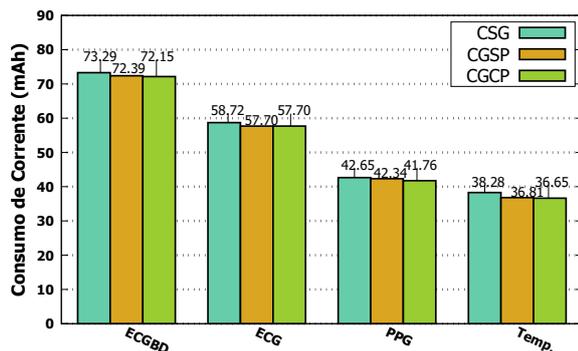


Tabela 5.4: Consumo energético dos dispositivos (CED)

Dispositivo	Valores dos Dispositivos		
	IC	Val. Mín	Val. Máx
ECGBD-CSG	0,008223	73,2838	73,3002
ECGBD-CGSP	0,010895	72,3781	72,3999
ECGBD-CGCP	0,008787	72,1388	72,1672
ECG-CSG	0,003602	58,7181	58,7253
ECG-CGSP	0,007438	57,6966	57,7114
ECG-CGCP	0,003036	57,7010	57,7070
PPG-CSG	0,006321	42,6477	42,6603
PPG-CGSP	0,007040	42,3320	42,3460
PPG-CGCP	0,008315	41,7517	41,7683
Temp-CSG	0,012458	38,2715	38,2965
Temp-CGSP	0,007840	36,8022	36,8178
Temp-CGCP	0,007438	36,6466	36,6614

Apesar do GROWN propiciar eficiência energética nos dispositivos, conforme mostra a Figura 5.4, o módulo de comunicação sem fio *Bluetooth Low Energy* (BLE) HM-10 permaneceu ativo e emparelhado com o *sink* em todo o período de experimentação. Logo, ele contribuiu no aumento do consumo de energia dos dispositivos. Este módulo utiliza o chip CC2540 BLE (Texas Instruments) para transmissão e recepção dos dados que, segundo seu *datasheet*¹¹, consome $\approx 24\text{mA}$ quando ativo TX e $\approx 8.5\text{mA}$ ¹² apenas ativo e emparelhado. Assim, para mitigar o impacto desse consumo de energia, ele foi configurado no modo *sleep*. Esse modo permite manter o dispositivo *dormindo* por um determinado período, podendo ser acordado com um comando *wakeup*. O modo *sleep* foi configurado no dispositivo Temperatura, visto que possui a maior correlação entre amostras consecutivas. Logo, o GROWN detecta a correlação e ativa o modo *sleep* sempre que o dispositivo não transmitir duas amostras consecutivas. A Tabela 5.6 sintetiza a autonomia da bateria em horas obtida pelos dispositivos com o GROWN, e demonstra a autonomia do dispositivo Temperatura com e sem o emprego do modo *sleep*. Essa configuração levou-o a consumir 24,92 mAh com perdas e 25,81 mAh sem perdas, aumentando a vida útil do dispositivo em 53,73% com **CGCP** e 48,37% com **CGSP**.

¹⁰<https://cdn.sparkfun.com/datasheets/Sensors/Biometric/AD8232.pdf>

¹¹<http://www.ti.com/lit/ds/symlink/cc2540.pdf>

¹²http://fab.cba.mit.edu/classes/863.15/doc/tutorials/programming/bluetooth/bluetooth40_en.pdf

Tabela 5.6: Tempo de vida dos dispositivos

Dispositivo	Autonomia da bateria (h)			Consumo (mAh)		
	CSG	CGSP	CGCP	CSG	CGSP	CGCP
ECGBD	5,46	5,53	5,54	73,29	72,39	72,15
ECG	6,81	6,93	6,93	58,72	57,70	57,70
PPG	9,38	9,45	9,58	42,65	42,34	41,76
Temperatura	10,45	10,87	10,91	38,28	36,81	36,65
Temperatura <i>sleep</i>	10,45	15,50	16,05	38,28	25,81	24,92

5.3.4 Taxa de compressão

O GROWN atinge taxas de compressão mais altas com amostras que possuem maior correlação entre elas. As Taxas de Compressão (TC) para os dispositivos ECGBD, ECG, PPG e Temperatura obtidas com o uso do mecanismo encontram-se na Tabela 5.7. Observa-se que o sinal de Temperatura obteve as maiores taxas, 99,16% com o método com perdas e 97,50% com o método sem perdas, e o sinal ECG obteve a mesma taxa de compressão para ambos os métodos 35,68%. Este fato ocorre porque apesar de existir uma pequena correlação entre amostras consecutivas, a diferença entre os sinais excedeu o limiar previamente estabelecido de 1. Assim, a informação foi transmitida. No entanto, isto ocorre com menor frequência no sinal PPG, que apresentou a menor taxa de compressão entre os demais sinais.

Tabela 5.7: Taxa de Compressão (TC)

Sinal	No. de pacotes transmitidos			Taxa de Compressão (%)	
	Original	C/Perda	S/Perda	C/perdas	S/perdas
ECGBD	8.863	5.589	5.757	36,94	35,04
ECG	7.586	4.879	4.879	36,68	36,68
PPG	6.483	5.011	6.016	22,70	07,20
Temp.	120	1	3	99,16	97,50

5.4 DISCUSSÃO

Os resultados obtidos nas experimentações realizadas demonstraram a eficiência do GROWN na redução do consumo energético dos dispositivos e sua aplicação em WBANs de tempo real. Assim, são utilizadas configurações e métricas que foram definidos e descritos na Seção 5.2, sendo que, em todas as configurações buscou-se o mesmo objetivo de avaliar a eficiência do mecanismo. Em todo o processo de experimentação o mecanismo atentou-se para à redundância da informação. Esta redundância determina o grau de aleatoriedade que uma determinada amostra pode assumir. Isto assemelha-se com a entropia da informação (Shannon, 1948), onde, amostras consecutivas que dispõem de uma grande diferença possuem maior grau de entropia. Por outro lado, amostras consecutivas que possuem grande correlação apresentam um menor grau de entropia. Diferentemente do trabalho de (Antonopoulos e Voros, 2016) que explora a correlação e constrói tabelas dinâmicas de tempos em tempos, o GROWN possui uma tabela fixa e explora a correlação com base na diferença entre amostras consecutivas.

Neste sentido, foi constatado que o tempo de compressão e descompressão dos sinais varia conforme à correlação entre as amostras. Visto que, o GROWN atribui códigos mais curtos para compressão em números menores e diminui seu tempo de atraso. Assim,

sempre que a correlação for alta a diferença entre as amostras é pequena, e diminui o tempo para compressão. Este mesmo cenário repete-se na descompressão, porém, o dispositivo *sink* detém mais recursos computacionais. Logo, o tempo de descompressão diminui. Além disso, verificou-se que o emprego do GROWN não afeta drasticamente no atraso médio. Isso ocorre pelo fato do mecanismo apresentar uma tabela fixa para compressão. Assim, não é necessário processamento adicional e alterações nas tabelas de codificação e decodificação. Isto propicia sua aplicabilidade nas WBANs de tempo real que necessitam de um atraso máximo de 250ms para aplicações não médicas (Movassaghi et al., 2014).

Assim como o tempo de compressão e descompressão, a correlação entre amostras consecutivas dos sinais foi determinante para avaliação do GROWN na eficiência energética dos dispositivos. Além disso, as diferentes configurações impostas na Seção 5.2 modificaram o consumo energético. Como esperado, verificou-se que a adição de um limiar no filtro de predição (Configuração 2) favoreceu para a eficiência energética dos dispositivos. Com exceção do dispositivo ECG que apresentou a mesma eficiência para ambas as configurações (2 e 3), os demais dispositivos obtiveram um ganho energético, conforme constatados na Tabela 5.6. Neste contexto, demonstrou-se que a compressão da diferença entre amostras consecutivas sem perdas (Configuração 3) representou um ganho energético para todos os dispositivos. Além disso, a base de dados utilizada para leitura em um módulo SD apresentou comportamentos similares ao dispositivo ECG. Isto solidifica os resultados apresentados e o ambiente de experimentação.

A eficiência energética obtida por cada dispositivo favorece na autonomia da bateria. Assim, o dispositivo permanece ligado por mais tempo. Logo, pode-se transmitir mais informações. Neste contexto, procurou-se mitigar o consumo energético e reduzir o tempo de atividade do módulo HM-10 BLE (*Bluetooth Low Energy*). Para isso, configurações do modo *sleep* foram pré-determinadas no dispositivo Temperatura que apresentou maior correlação entre amostras consecutivas. Estas configurações com o emprego do GROWN propiciaram um grande avanço na eficiência energética do dispositivo conforme apresentado na Tabela 5.6. Contudo, para os demais dispositivos esta solução não é factível, visto que, mesmo com um limiar pré-estabelecido os valores excedem frequentemente, e a informação é comprimida para transmissão. A mensuração da taxa de compressão é realizada na medida em que a quantidade de pacotes transmitidos diminui com o emprego do GROWN. A correlação entre amostras consecutivas proporcionou o aumento da taxa de compressão conforme constatado na Tabela 5.7. Assim, este fato corrobora para o aumento da eficiência energética dos dispositivos. Logo, os dispositivos que obtiveram o menor número de pacotes transmitidos alcançaram as maiores taxas de compressão, e como resultado alcançaram maior eficiência energética.

5.5 RESUMO

Este capítulo apresentou a metodologia de avaliação e a validade dos resultados obtidos pelo mecanismo GROWN, a fim de mensurar a eficiência energética e sua aplicação numa WBAN de tempo real. Os resultados obtidos diante das métricas definidas demonstram que o GROWN propicia o aumento da vida dos dispositivos através da correlação entre amostras consecutivas e seu método de compressão de dados local. As distintas configurações do ambiente de experimentação buscaram exibir a eficiência energética e a taxa de compressão na redução dos pacotes transmitidos. Além disso, os resultados foram discutidos e analisados para melhor entender o mecanismo. O GROWN mostrou-se uma alternativa viável para aplicações WBAN de tempo real, sua gerência para analisar amostras consecutivas e sua tabela fixa de compressão apresentam baixo impacto na latência total.

6 CONCLUSÕES

A evolução das redes de sensores sem fio (WSNs) e nanossistemas têm possibilitado o desenvolvimento de novas redes e serviços para o cotidiano das pessoas. Entre os desenvolvimentos destacam-se as redes corporais sem fio (*Wireless Body Area Networks* - WBANs) que possibilitam o monitoramento contínuo e em tempo real dos sinais fisiológicos das pessoas através de sensores corporais implantáveis ou sobre o corpo humano. Estes sensores são comumente acoplados a placas (dispositivo corporal) que realizam o processamento e a transmissão da informação para um dispositivo central (*sink*). Assim, o *sink* encaminha estas informações para profissionais da saúde ou centrais para possíveis diagnósticos. Contudo, este cenário implica em diversos desafios como o consumo de energia dos dispositivos, vista às sucessivas transmissões entre o dispositivo corporal e o *sink*. Além disso, a falta de um ou mais dispositivos pelo uso excessivo de energia pode inviabilizar o diagnóstico, e colocar em risco à vida das pessoas. Assim, diversos autores propuseram métodos com técnicas de compressão de dados local para mitigar o consumo energético.

Um levantamento sobre o estado da arte demonstrou que embora alguns métodos reduzam o consumo de energia poucos desconsideram a transmissão de informações redundantes. Além disso, outros trabalhos não consideram a latência máxima prevista pela IEEE no padrão das WBANs de 250ms (of Electrical e Engineers, 2012). Logo, detectou-se a necessidade de novas soluções com eficácia na redução do consumo energético dos dispositivos corporais. Essas soluções necessitam gerenciar a redundância da informação transmitida e adaptar-se ao padrão máximo de latência estabelecido para uma WBAN. A fim de lidar com esses problemas, este trabalho propôs o mecanismo GROWN (*EnerGy-Efficient Local Data Compression for TRansmission Over WBAN*) para gerenciar a redundância da informação transmitida e reduzir o consumo de energia dos dispositivos corporais sem impactar de forma intensa no atraso total da transmissão da informação para o *sink*. Ele baseia-se em técnicas de compressão de dados local encontradas nas redes de sensores sem fio (WSNs) e o uso de limiares definidos. O emprego destas técnicas possibilitou o desenvolvimento de uma tabela fixa de compressão de dados que depende da correlação entre amostras consecutivas dos sinais fisiológicos. Assim, o GROWN atribui códigos mais curtos para baixas diferenças entre as amostras.

A avaliação do GROWN foi conduzida por experimentações que buscaram aproximar o *testbed* de um ambiente real. As experimentações buscaram avaliar a eficiência e a eficácia do mecanismo na redução do consumo de energia dos dispositivos corporais e seu emprego em aplicações de tempo real. Os resultados alcançados demonstram que o GROWN propiciou eficiência energética em todos os dispositivos corporais avaliados. Especialmente, nos sinais que apresentam grande correlação entre amostras consecutivas, conforme apresentado no sinal de temperatura. Além disso, as experimentações demonstraram que a correlação entre amostras consecutivas impacta no tempo de compressão e descompressão das informações e na taxa de compressão dos pacotes transmitidos. Particularmente, o mecanismo mostrou-se eficiente mesmo em sinais que apresentam pouca correlação. Logo, apesar do tempo gasto para compressão e descompressão das informações, seu impacto no atraso total é mínimo, e corrobora sua aplicação em WBANs de tempo real.

Conclui-se que ser mais eficiente na economia de energia de modo a oferecer uma transmissão eficaz, garantir o funcionamento dos serviços de coleta e ampliar o tempo de vida dos dispositivos, é possível. Esse objetivo foi alcançado pelo GROWN, visto que ele aumentou o tempo de vida dos dispositivos e consequentemente prolongou os serviços de coleta. O

emprego de métodos de compressão de dados local contribuíram para redução da informação na transmissão. Além disso, foi possível reduzir a redundância da informação coletada pelos sensores. Por fim, entende-se que uso de compressão de dados local interfere no pós-processamento do sinal fisiológico coletado pelo sensor. Tendo em vista que toda informação leva um tempo para ser comprimida. Logo, esta medida altera-se conforme o método empregado. Com isto, este trabalho contribui com um detalhamento do uso de compressão de dados local nas WBANs e como seu emprego pode reduzir o consumo de energia dos dispositivos.

6.1 TRABALHOS FUTUROS

Os resultados alcançados e as conclusões apresentadas acima demonstram que esta pesquisa atingiu os objetivos propostos. Conforme os resultados, o mecanismo mostrou-se eficiente na redução do consumo de energia dos dispositivos. Além disso, o aumento no atraso da informação ocorreu de forma suave, não impactando drasticamente o atraso total. Contudo, novas experimentações precisam avaliar um cenário denso, visto que uma WBAN deve suportar até 256 dispositivos simultâneos. Além disso, o mecanismo demanda de otimizações. Abaixo, detalha-se de modo mais claro as questões a serem investigadas:

1. **Falha na entrega dos pacotes:** No ambiente de experimentação, assume-se que há um controle de perdas de pacotes entre os dispositivos corporais e o *sink* realizado pela tecnologia de comunicação sem fio. Logo, o mecanismo não identifica falhas no controle destes pacotes. Isto pode acarretar em equívocos na apresentação da informação após a descompressão, visto que, o GROWN gerencia a correlação entre amostras consecutivas.
2. **Controle da vida útil dos dispositivos:** O mecanismo proposto opera de modo eficiente no aumento do tempo de vida dos dispositivos. Contudo, não existe um controle da situação atual, apenas uma estimativa do aumento do tempo de vida. Assim, faz-se necessário controlar o resíduo de energia dos dispositivos em tempo real. Este controle propicia a atualização do *sink* quanto a vida de cada dispositivo.
3. **Tecnologia de comunicação sem fio:** O emprego da tecnologia de comunicação sem fio *Bluetooth Low Energy* (BLE) gerou um consumo contínuo no dispositivo, em virtude do pareamento com o *sink*. Esta conexão aumenta os custos de energia. Assim, pretende-se avaliar o mecanismo com outras tecnologias e seu resultado na eficiência energética.
4. **Privacidade dos dados:** Manter os dados seguros torna-se uma tarefa de constante evolução. As WBANs transmitem informações sensíveis das pessoas, e a privacidade sobre estes dados torna-se imprescindível. O GROWN não considera questões de segurança e privacidade dos dados. Logo, precisa-se de uma análise minuciosa de métodos e técnicas aplicáveis nas WBANs.
5. **Mobilidade corporal:** Nas experimentações realizadas os dispositivos operavam de maneira estática, sem a existência de mobilidade corporal. Logo, planeja-se estudar o comportamento dos dispositivos e do *sink* com emprego do mecanismo em um cenário com mobilidade, que implica em falhas de transmissão conforme a postura corporal.

REFERÊNCIAS

- Adhikary, S., Choudhury, S. e Chattopadhyay, S. (2016). A new routing protocol for wban to enhance energy consumption and network lifetime. Em *Proceedings of the 17th international conference on distributed computing and networking*, páginas 1–6.
- Akyildiz, I. F., Su, W., Sankarasubramaniam, Y. e Cayirci, E. (2002). Wireless sensor networks: a survey. *Computer networks*, 38(4):393–422.
- Ali, A., Ming, Y., Chakraborty, S. e Iram, S. (2017). A comprehensive survey on real-time applications of wsn. *Future Internet*, 9(4):77.
- Almeida-Santos, P. S., Rocha, I. e Hein, N. (2014). Utilização da entropia informacional na seleção de indicadores financeiros mais relevantes para tomada de decisão no setor público: o caso dos estados brasileiros. *Perspectivas em Ciência da Informação*, 19(2):83–105.
- Anderson, G. e Horvath, J. (2004). The growing burden of chronic disease in america. *Public health reports*, 119(3):263–270.
- Antonescu, B. e Basagni, S. (2013). Wireless body area networks: challenges, trends and emerging technologies. Em *Proceedings of the 8th international conference on body area networks*, páginas 1–7.
- Antonopoulos, C. P. e Voros, N. S. (2016). Resource efficient data compression algorithms for demanding, wsn based biomedical applications. *Journal of biomedical informatics*, 59:1–14.
- Araújo, D. S. A. d. (2013). *Análise de Agrupamentos Com Base na Teoria da Informação: Uma Abordagem Representativa*. Tese de doutorado, Universidade Federal do Rio Grande do Norte.
- Arrabi, S. e Lach, J. (2009). Adaptive lossless compression in wireless body sensor networks. Em *Proceedings of the Fourth International Conference on Body Area Networks*, páginas 1–8.
- Asif, A. e Sumra, I. A. (2017). Applications of wireless body area network (wban): A survey. *Engineering Science and Technology. International. Research Journal*, páginas 64–71.
- Azar, J., Darazi, R., Habib, C., Makhoul, A. e Demerjian, J. (2018a). Using dwt lifting scheme for lossless data compression in wireless body sensor networks. Em *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)*, páginas 1465–1470. IEEE.
- Azar, J., Makhoul, A., Darazi, R., Demerjian, J. e Couturier, R. (2018b). On the performance of resource-aware compression techniques for vital signs data in wireless body sensor networks. Em *Communications Conference (MENACOMM), IEEE Middle East and North Africa*, páginas 1–6. IEEE.
- Bachir, A., Dohler, M., Watteyne, T. e Leung, K. K. (2010). Mac essentials for wireless sensor networks. *IEEE Communications Surveys & Tutorials*, 12(2):222–248.
- Balaji, S., Nathani, K. e Santhakumar, R. (2019). Iot technology, applications and challenges: a contemporary survey. *Wireless personal communications*, 108(1):363–388.

- Barr, K. C. e Asanović, K. (2006). Energy-aware lossless data compression. *ACM Transactions on Computer Systems (TOCS)*, 24(3):250–291.
- Bleiholder, J. e Naumann, F. (2009). Data fusion. *ACM Computing Surveys (CSUR)*, 41(1):1.
- Bontorin, M., Nogueira, M. e Santos, A. (2015). A bio-inspired technique for energy-aware fpgas on body area networks. *IEEE Latin America Transactions*, 13(12):3707–3713.
- Campani, C. A. P. (2005). *Avaliação da Compressão de Dados e da Qualidade de Imagem em Modelos de Animação Gráfica para Web: uma nova abordagem baseada em Complexidade de Kolmogorov*. Tese de doutorado, UFRGS - Universidade Federal do Rio Grande do Sul.
- Castanedo, F. (2013). A review of data fusion techniques. *The Scientific World Journal*, 2013.
- Cavallari, R., Martelli, F., Rosini, R., Buratti, C. e Verdone, R. (2014). A survey on wireless body area networks: Technologies and design challenges. *IEEE Communications Surveys & Tutorials*, 16(3):1635–1657.
- Chen, M., Gonzalez, S., Vasilakos, A., Cao, H. e Leung, V. C. (2011). Body area networks: A survey. *Mobile networks and applications*, 16(2):171–193.
- Chen, S., Luo, Z. e Gan, H. (2018). An entropy fusion method for feature extraction of eeg. *Neural Computing and Applications*, 29(10):857–863.
- Crosby, G. V., Ghosh, T., Murimi, R. e Chin, C. A. (2012). Wireless body area networks for healthcare: A survey. *International Journal of Ad Hoc, Sensor & Ubiquitous Computing*, 3(3):1.
- Danković, N., Antić, D., Nikolić, S. S., Perić, S. L., Perić, Z. H. e Jocić, A. V. (2017). The probability of stability estimation of an arbitrary order dpcm prediction filter: Comparison between the classical approach and the monte carlo method. *Information Technology And Control*, 46(2):194–204.
- Dasarathy, B. V. (1997). Sensor fusion potential exploitation-innovative architectures and illustrative applications. *Proceedings of the IEEE*, 85(1):24–38.
- Dasarathy, B. V. (2001). Information fusion-what, where, why, when, and how? *Information Fusion*, 2(2):75–76.
- de Geografia e Estatística, I. B. (2020). Projeção da população do brasil e das unidades da federação. <https://www.ibge.gov.br/apps/populacao/projecao/>. Acessado em 20/04/2020.
- Deepu, C. J., Heng, C.-H. e Lian, Y. (2017). A hybrid data compression scheme for power reduction in wireless sensors for iot. *IEEE transactions on biomedical circuits and systems*, 11(2):245–254.
- Dhanvijay, M. M. e Patil, S. C. (2019). Internet of things: A survey of enabling technologies in healthcare and its applications. *Computer Networks*.
- Ding, W., Jing, X., Yan, Z. e Yang, L. T. (2019). A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion. *Information Fusion*, 51:129–144.

- Durrant-Whyte, H., Stevens, M. e Nettleton, E. (2001). Data fusion in decentralised sensing networks. Em *Proceedings of the 4th International Conference on Information Fusion*, páginas 302–307.
- Durrant-Whyte, H. F. (1988). Sensor models and multisensor integration. *The international journal of robotics research*, 7(6):97–113.
- Farach, M., Noordewier, M., Savari, S., Shepp, L., Wyner, A. e Ziv, J. (1995). On the entropy of dna: Algorithms and measurements based on memory and rapid convergence. Em *SODA*, volume 95, páginas 48–57.
- Fouad, H. (2014). Continuous health-monitoring for early detection of patient by web telemedicine system. Em *International Conference on Circuits, Systems and Signal Processing, Saint Petersburg State Politechnical University, Russia*.
- Gatouillat, A., Massot, B., Badr, Y., Sejdić, E. e Gehin, C. (2018). Building iot-enabled wearable medical devices: an application to a wearable, multiparametric, cardiorespiratory sensor. Em *BIODEVICES*, páginas 109–118.
- Ghamari, M., Janko, B., Sherratt, R., Harwin, W., Piechockic, R. e Soltanpur, C. (2016). A survey on wireless body area networks for ehealthcare systems in residential environments. *Sensors*, 16(6):831.
- Giorgi, G. (2017). A combined approach for real-time data compression in wireless body sensor networks. *IEEE Sensors Journal*, 17(18):6129–6135.
- Golomb, S. (1966). Run-length encodings (corresp.). *IEEE transactions on information theory*, 12(3):399–401.
- Gray, R. M. e Neuhoff, D. L. (1998). Quantization. *IEEE transactions on information theory*, 44(6):2325–2383.
- Haghighat, M., Abdel-Mottaleb, M. e Alhalabi, W. (2016). Discriminant correlation analysis: Real-time feature level fusion for multimodal biometric recognition. *IEEE Transactions on Information Forensics and Security*, 11(9):1984–1996.
- Hall, D. L. e Llinas, J. (1997). An introduction to multisensor data fusion. *Proceedings of the IEEE*, 85(1):6–23.
- Hanson, M. A., Powell Jr, H. C., Barth, A. T., Ringgenberg, K., Calhoun, B. H., Aylor, J. H. e Lach, J. (2009). Body area sensor networks: Challenges and opportunities. *Computer*, 42(1).
- Huffman, D. A. (1952). A method for the construction of minimum-redundancy codes. *Proceedings of the IRE*, 40(9):1098–1101.
- Instruments, T. (2020). INA219 Zero-Drift, Bidirectional Current/Power Monitor With I2C Interface. <http://www.ti.com/lit/ds/symlink/ina219.pdf>. Acessado em 20/03/2020.
- Islam, S. R., Kwak, D., Kabir, M. H., Hossain, M. e Kwak, K.-S. (2015). The internet of things for health care: a comprehensive survey. *IEEE Access*, 3:678–708.

- Javaid, N., Ahmad, A., Nadeem, Q., Imran, M. e Haider, N. (2015). im-simple: improved stable increased-throughput multi-hop link efficient routing protocol for wireless body area networks. *Computers in Human Behavior*, 51:1003–1011.
- Javaid, N., Ahmad, A., Tauqir, A., Imran, M., Guizani, M., Khan, Z. A. e Qasim, U. (2016). Modeling induction and routing to monitor hospitalized patients in multi-hop mobility-aware body area sensor networks. *EURASIP Journal on Wireless Communications and Networking*, 2016(1):147.
- Jayarathna, T., Gargiulo, G. D. e Breen, P. P. (2020). Continuous vital monitoring during sleep and light activity using carbon-black elastomer sensors. *Sensors*, 20(6):1583.
- Kalaivani, S., Tharini, C., Saranya, K. e Priyanka, K. (2020). Design and implementation of hybrid compression algorithm for personal health care big data applications. *Wireless Personal Communications*, páginas 1–17.
- Khan, R. A. e Pathan, A.-S. K. (2018). The state-of-the-art wireless body area sensor networks: A survey. *International Journal of Distributed Sensor Networks*, 14(4):1550147718768994.
- Kimura, N. e Latifi, S. (2005). A survey on data compression in wireless sensor networks. Em *Information Technology: Coding and Computing, 2005. ITCC 2005. International Conference on*, volume 2, páginas 8–13. IEEE.
- Klein, L. A. (1999). *Sensor and data fusion concepts and applications*. Society of Photo-Optical Instrumentation Engineers (SPIE).
- Kolo, J. G., Shanmugam, S. A., Lim, D. W. G., Ang, L.-M. e Seng, K. P. (2012). An adaptive lossless data compression scheme for wireless sensor networks. *Journal of Sensors*, 2012.
- Kumari, R. S. S. e Gayathri, C. (2017). Interfacing of mems motion sensor with fpga using i2c protocol. Em *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, páginas 1–5. IEEE.
- Kwak, K. S., Ullah, S. e Ullah, N. (2010). An overview of ieee 802.15. 6 standard. Em *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on*, páginas 1–6. IEEE.
- Lakshmanan, M., Karnan, H. e Natarajan, S. (2020). Smart diagnosis of cardiac arrhythmias using optimal feature rank score algorithm for solar based energy storage ecg acquisition system. Em *Smart Healthcare for Disease Diagnosis and Prevention*, páginas 125–139. Elsevier.
- Latré, B., Braem, B., Moerman, I., Blondia, C. e Demeester, P. (2011). A survey on wireless body area networks. *Wireless Networks*, 17(1):1–18.
- Lay-Ekuakille, A., Griffo, G., Vergallo, P., Massaro, A., Spano, F. e Gigli, G. (2015). Implantable neurorecording sensing system: Wireless transmission of measurements. *IEEE Sens. J*, 15:2603–2613.
- Levin, A. e Stevens, P. E. (2011). Early detection of ckd: the benefits, limitations and effects on prognosis. *Nature Reviews Nephrology*, 7(8):446.
- Lewis, D. (2008). 802.15. 6 call for applications in body area networksresponse summary,”. *15-08-0407-05-0006*.

- Li, J., Zhang, X., Gong, J., Tang, J., Ren, Z., Li, G., Deng, Y. e Cai, J. (2018). Signal-noise identification of magnetotelluric signals using fractal-entropy and clustering algorithm for targeted de-noising. *Fractals*, 26(02):1840011.
- Li, Y. e Liang, Y. (2016). Temporal lossless and lossy compression in wireless sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 12(4):37.
- Liang, Y., Elgendi, M., Chen, Z. e Ward, R. (2018). An optimal filter for short photoplethysmogram signals. *Scientific data*, 5:180076.
- Lima, M. N. (2010). *Saúde Móvel-Conceitos, Iniciativas e Aplicações*. Curitiba: Clube de Autores.
- Lipprandt, M., Eichelberg, M., Thronicke, W., Kruger, J., Druke, I., Willemsen, D., Busch, C., Fiehe, C., Zeeb, E. e Hein, A. (2009). Osami-d: An open service platform for healthcare monitoring applications. Em *Human System Interactions, 2009. HSI'09. 2nd Conference on*, páginas 139–145. IEEE.
- Liu, Z., Banakhr, F., Monte, G. e Huang, V. (2015). Using algorithms on smart transducer: An iee standard perspective. *IEEE Sensors Journal*, 15(5):2523–2530.
- Mack, H. (2017). Remote patient monitoring market grew by 44 percent in 2016, report says. <https://www.mobihealthnews.com/content/remote-patient-monitoring-market-grew-44-percent-2016-report-says>. Acessado em 27/12/2018.
- Mahapatro, A. et al. (2011). Online fault detection and recovery in body sensor networks. Em *Information and Communication Technologies (WICT), 2011 World Congress on*, páginas 407–412. IEEE.
- Manna, T. e Misra, I. S. (2018). Implementation of energy efficient wban using ieee 802.15. 6 scheduled access mac with fast dwt based backhaul data compression for e-healthcare. Em *International Conference on Communication Systems and Networks*, páginas 26–51. Springer.
- Marcelloni, F. e Vecchio, M. (2008). A simple algorithm for data compression in wireless sensor networks. *IEEE communications letters*, 12(6).
- Marcelloni, F. e Vecchio, M. (2009). An efficient lossless compression algorithm for tiny nodes of monitoring wireless sensor networks. *The Computer Journal*, 52(8):969–987.
- Marcelloni, F. e Vecchio, M. (2010). Enabling energy-efficient and lossy-aware data compression in wireless sensor networks by multi-objective evolutionary optimization. *Information Sciences*, 180(10):1924–1941.
- Maurya, A. K., Singh, D. e Sarje, A. K. (2011). Median predictor based data compression algorithm for wireless sensor network. *International Journal of Smart Sensors and Ad Hoc Networks*, 1(1):62–65.
- Movassaghi, S., Abolhasan, M., Lipman, J., Smith, D. e Jamalipour, A. (2014). Wireless body area networks: A survey. *IEEE Communications Surveys & Tutorials*, 16(3):1658–1686.
- Nadeem, Q., Javaid, N., Mohammad, S. N., Khan, M., Sarfraz, S. e Gull, M. (2013). Simple: Stable increased-throughput multi-hop protocol for link efficiency in wireless body area networks. *arXiv preprint arXiv:1307.7102*.

- Naganawa, J.-i., Wangchuk, K., Kim, M., Aoyagi, T. e Takada, J.-i. (2015). Simulation-based scenario-specific channel modeling for wban cooperative transmission schemes. *IEEE J. Biomedical and Health Informatics*, 19(2):559–570.
- Nakayama, F., Lenz, P., Banou, S., Nogueira, M., Santos, A. e Chowdhury, K. R. (2019). A continuous user authentication system based on galvanic coupling communication for s-health. *Wireless Communications and Mobile Computing*, 2019.
- Nali, I. C. (2012). *Estudo avaliativo da informação mútua generalizada e de métricas clássicas como medidas de similaridade para correção em imagens fractais e cerebrais*. Tese de doutorado, Universidade de São Paulo.
- Nascimento, W. S. e Prudente, F. V. (2016). Sobre um estudo da entropia de shannon no contexto da mecânica quântica: uma aplicação ao oscilador harmônico livre e confinado. *Quim. Nova*, 39(6):757–764.
- of Electrical, I. e Engineers, E. (2012). Ieee standard for local and metropolitan area networks - part 15.6: Wireless body area networks. *IEEE Std 802.15.6-2012*, páginas 1–271.
- Parks, J. E. (2007). Ohms Law III Resistors in Series and Parallel. *Department of Physics and Anatomy, University of Tennessee*.
- Patel, D., Bhogan, V. e Janson, A. (2013). Simulation and comparison of various lossless data compression techniques based on compression ratio and processing delay. *International Journal of Computer Applications*, 81(14).
- Pennebaker, W. B. e Mitchell, J. L. (1992). *JPEG: Still image data compression standard*. Springer Science & Business Media.
- Pottie, G. J. e Kaiser, W. J. (2000). Wireless integrated network sensors. *Communications of the ACM*, 43(5):51–58.
- Pu, I. M. (2005). *Fundamental data compression*. Butterworth-Heinemann.
- Qu, Y., Zheng, G., Ma, H., Wang, X., Ji, B. e Wu, H. (2019). A survey of routing protocols in wban for healthcare applications. *Sensors*, 19(7):1638.
- Raja, N. S. M., Arunmozhi, S., Lin, H., Dey, N. e Rajinikanth, V. (2019). A study on segmentation of leukocyte image with shannon's entropy. Em *Histopathological Image Analysis in Medical Decision Making*, páginas 1–27. IGI Global.
- Rashid, B. e Rehmani, M. H. (2016). Applications of wireless sensor networks for urban areas: A survey. *Journal of network and computer applications*, 60:192–219.
- Resque, P., Pinheiro, S., Rosário, D., Cerqueira, E., Vergutz, A., Nogueira, M. e Santos, A. (2019). Assessing data traffic classification to priority access for wireless healthcare application. Em *2019 IEEE Latin-American Conference on Communications (LATINCOM)*, páginas 1–6. IEEE.
- Roy, M., Chowdhury, C. e Aslam, N. (2017). Designing an energy efficient wban routing protocol. Em *2017 9th International Conference on Communication Systems and Networks (COMSNETS)*, páginas 298–305. IEEE.
- Salomon, D. (2004). *Data compression: the complete reference*. Springer Science & Business Media.

- Sayood, K. (2017). *Introduction to data compression*. Morgan Kaufmann.
- Schoellhammer, T., Greenstein, B., Osterweil, E., Wimbrow, M. e Estrin, D. (2004). Lightweight temporal compression of microclimate datasets [wireless sensor networks]. *29th Annual IEEE International Conference on Local Computer Networks*, páginas 516–524.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423.
- SIG, B. (2014). Bluetooth low energy. <https://www.bluetooth.com/specifications/archived-specifications/>. Acessado em 10/04/2020.
- Srisooksai, T., Keamarungsi, K., Lamsrichan, P. e Araki, K. (2012). Practical data compression in wireless sensor networks: A survey. *Journal of Network and Computer Applications*, 35(1):37–59.
- Sukor, M., Ariffin, S., Faisal, N., Yusof, S. S. e Abdallah, A. (2008). Performance study of wireless body area network in medical environment. Em | *Second Asia International Conference on Modelling & Simulation*, páginas 202–206. IEEE.
- Sweldens, W. (1998). The lifting scheme: A construction of second generation wavelets. *SIAM journal on mathematical analysis*, 29(2):511–546.
- Tavares, T., Nogueira, M., Rosário, D., Santos, A. e Cerqueira, E. (2020). Traffic model based on autoregression for ppg signals in wearable networks. *IEEE Networking Letters*, 2(2):49–53.
- Teuhola, J. (1978). A compression method for clustered bit-vectors. *Information processing letters*, 7(6):308–311.
- Tharini, C. e Ranjan, P. V. (2009). Design of modified adaptive huffman data compression algorithm for wireless sensor network. *Journal of Computer Science*, 5(6):466.
- Tsai, T.-H. e Kuo, W.-T. (2018). An efficient ecg lossless compression system for embedded platforms with telemedicine applications. *IEEE Access*, 6:42207–42215.
- Ullah, F., Abdullah, A. H., Kaiwartya, O. e Cao, Y. (2017). Trapy-mac: Traffic priority aware medium access control protocol for wireless body area network. *Journal of medical systems*, 41(6):93.
- Ullah, S., Higgins, H., Braem, B., Latre, B., Blondia, C., Moerman, I., Saleem, S., Rahman, Z. e Kwak, K. S. (2012). A comprehensive survey of wireless body area networks. *Journal of medical systems*, 36(3):1065–1094.
- Uthayakumar, J., Vengattaraman, T. e Dhavachelvan, P. (2018). A survey on data compression techniques: From the perspective of data quality, coding schemes, data type and applications. *Journal of King Saud University-Computer and Information Sciences*.
- van der Byl, A., Neilson, R. e Wilkinson, R. H. (2009). An evaluation of compression techniques for wireless sensor networks. Em *AFRICON 2009*, páginas 1–6.
- Vecchio, M., Giaffreda, R. e Marcelloni, F. (2014). Adaptive lossless entropy compressors for tiny iot devices. *IEEE Transactions on Wireless Communications*, 13(2):1088–1100.

- Vergutz, A., da Silva, R., Vieira, A. B., Nogueira, M. et al. (2017). Um Sistema de Identificação Antecipada e Transmissão Prioritária de Alertas Médicos sobre WBAN e WLAN. Em *Anais do XXXV Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*. SBC.
- Vergütz, A., G Prates, N., Henrique Schwengber, B., Santos, A. e Nogueira, M. (2020). An architecture for the performance management of smart healthcare applications. *Sensors*, 20(19):5566.
- Villanueva, J. M. M. (2009). Fusão de dados das técnicas de tempo de trânsito utilizando transdutores ultra-sônicos para medição da velocidade do vento. *Rio de Janeiro*.
- Wald, L. (1999). Some terms of reference in data fusion. *IEEE Transactions on geoscience and remote sensing*, 37(3):1190–1193.
- WEAVER, W. e SHANNON, C. (1971). A teoria matemática da comunicação. *Comunicação e indústria cultural*, 5:25–37.
- White, F. E. (1991). Data fusion lexicon. Relatório técnico, JOINT DIRECTORS OF LABS WASHINGTON DC.
- Williams, T., Kelley, C. e many others (2020). Gnuplot 5.2: an interactive plotting program. <http://www.gnuplot.info/>. Acessado em 10/04/2020.
- Wu, G., Ren, J., Xia, F. e Xu, Z. (2010). An adaptive fault-tolerant communication scheme for body sensor networks. *Sensors*, 10(11):9590–9608.
- Yang, J., Zhang, X. e Liang, W. (2020). A retrospective analysis of factors affecting surgical site infection in orthopaedic patients. *Journal of International Medical Research*, 48(4):0300060520907776.
- Yick, J., Mukherjee, B. e Ghosal, D. (2008). Wireless sensor network survey. *Computer networks*, 52(12):2292–2330.
- Yuce, M. R. (2010). Implementation of wireless body area networks for healthcare systems. *Sensors and Actuators A: Physical*, 162(1):116–129.
- Zhou, Y. (2017). *Energy efficient wireless body area network design in health monitoring scenarios*. Tese de doutorado, University of British Columbia.
- Zhou, Y., Fang, Y. e Zhang, Y. (2008). Securing wireless sensor networks: a survey. *IEEE Communications Surveys & Tutorials*, 10(3):6–28.

APÊNDICE A – PUBLICAÇÕES

Ao longo do mestrado realizou-se diversas submissões de modo a contribuir com a ciência, disseminar o conhecimento e colaborar com a comunidade acadêmica. Este apêndice descreve as publicações alcançadas, o título e o local de cada publicação e o link de acesso. Por fim, apresenta-se o prêmio conquistado de melhor artigo.

A.1 ARTIGOS PUBLICADOS

- **Compressão de Dados Local em Tempo Real Energeticamente Eficiente para Redes WBAN.** C Passos, C Pedroso, A Batista, M Nogueira, A Santos. *Workshop de Gerência e Operação de Redes e Serviços (SBRC 2020 - XXV WGRS)*.
<https://sol.sbc.org.br/index.php/wgrs/article/view/12453>
- **GROWN: Local Data Compression in Real-Time To Support Energy Efficiency in WBAN.** C Passos, C Pedroso, A Batista, M Nogueira, A Santos. *IEEE Latin-American Conference on Communications (LATINCOM 2020)*.
<https://ieeexplore.ieee.org/document/9282319>

A.2 PREMIAÇÕES

- **Melhor trabalho do Workshop de Gerência e Operação de Redes e Serviços (SBRC 2020 - XXV WGRS)**