UNISINOS

Programa Interdisciplinar de Pós-Graduação em

Computação Aplicada

Mestrado/Doutorado Acadêmico

Gabriel Souto Fischer

HealCity :
Usando Dados de Sinais Vitais dos Cidadãos e a Técnica de Elasticidade para gerência de Ambientes de Saúde no Contexto de Cidades Inteligentes

São Leopoldo
2024

Gabriel Souto Fischer

HEALCITY:
Usando Dados de Sinais Vitais dos Cidadãos e a Técnica de Elasticidade para gerência de Ambientes de Saúde no Contexto de Cidades Inteligentes

Tese apresentada como requisito parcial para a obtenção do título de Doutor pelo Programa de Pós-Graduação em Computação Aplicada da Universidade do Vale do Rio dos Sinos — UNISINOS

Orientador:
Prof. Dr. Rodrigo da Rosa Righi

São Leopoldo
2024

F529h Fischer, Gabriel Souto.
HealCity : usando dados de sinais vitais dos cidadãos e a técnica de elasticidade para gerência de ambientes de saúde no contexto de cidades inteligentes / Gabriel Souto Fischer. – 2024.
155 f. : il. ; 30 cm.

Tese (doutorado) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Computação Aplicada, 2024.
"Orientador: Prof. Dr. Rodrigo da Rosa Righi"

1. Cidades inteligentes. 2. Elasticidade. 3. Hospitais inteligentes. 4. Internet das coisas. 5. Recursos humanos. I. Título.

CDU 004.4

Dados Internacionais de Catalogação na Publicação (CIP)
(Bibliotecária: Silvana Dornelles Studzinski – CRB 10/2524)

AGRADECIMENTOS À CAPES

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

A minha família.

O que é o sucesso?
Rir muito e com frequência; ganhar o respeito de pessoas inteligentes e o afeto das crianças; merecer a consideração de críticos honestos e suportar a traição de falsos amigos; apreciar a beleza, encontrar o melhor nos outros; deixar o mundo um pouco melhor, seja por uma saudável criança, um canteiro de jardim ou uma redimida condição social; saber que ao menos uma vida respirou mais fácil porque você viveu.
Isso é ter tido sucesso
— BESSIE STANLEY

AGRADECIMENTOS

Neste momento estou percorrendo mais uma etapa da minha vida, que é o meu doutorado. Foram muitos anos de luta e dedicação, com diversos desafios vencidos e perdidos para chegar até aqui. Diversos momentos de conciliação entre família, amigos, estudos e trabalho. Então não posso deixar de agradecer a todos aqueles que foram importantes para mim nessa jornada, me ajudando e me compreendendo em todos os momentos.

Primeiramente, agradeço a Deus pelo dom da vida e todas as coisas boas que aconteceram e ainda vão acontecer em minha vida.

Agradeço a meus pais, pela paciência, carinho, compreensão, em todos os momentos de minha formação acadêmica, e principalmente durante a escrita de minha monografia da graduação e de minha dissertação. Mas não apenas por isso, agradeço-os pelo apoio emocional em todos os momentos de minha vida e pelo amor incondicional recebido por eles. Agradeço a eles por sempre apoiarem e incentivarem os meus estudos.

Agradeço a meu irmão, por ser simplesmente ele, por tornar mais engraçados momentos de frustração e por me ajudar a rir em qualquer situação.

Agradeço a minha esposa Michelle, por estar ao meu lado durante toda minha pós graduação, me ajudando, me incentivando e me compreendendo. Mas não apenas por isso, também agradeço a ela pelo apoio emocional, estando ao meu lado nos momentos felizes e me apoiando incondicionalmente nos momentos tristes.

Agradeço ao meu filho Arthur, por ser a grande alegria de minha vida, por tornar qualquer dia cansativo ou frustrante em um dia melhor, unicamente ao sorrir quando me vê chegar em casa. Mas não apenas por isso, também agradeço a ele por me dar a força necessária para superar qualquer desafio por ele.

Agradeço aos meus amigos, que em muitos momentos foram deixados de lado em função das inúmeras atividades acadêmicas aos quais tive que realizar, principalmente a escrita de minha dissertação de mestrado e agora de minha tese de doutorado, e por sempre estarem ao meu lado, rindo comigo nos bons momentos e me apoiando nos momentos ruins.

Agradeço a todos os meus professores e principalmente a meu orientador Rodrigo da Rosa Righi que através de sua dedicação, incentivo e compreensão contribuiu muito para o desenvolvimento deste e de outros importantes trabalhos.

E agradeço a todos que direta ou indiretamente me ajudaram na minha jornada, tanto acadêmica, como pessoal, e também profissional.

"*Jamais considere seus estudos como uma obrigação, mas como uma oportunidade invejável para aprender a conhecer a influência libertadora da beleza do reino do espírito, para seu próprio prazer pessoal e para proveito da comunidade à qual seu futuro trabalho pertencer*".

(Albert Einstein)

RESUMO

Cidades inteligentes podem melhorar a qualidade de vida dos cidadãos otimizando a utilização de recursos. Em um ambiente conectado à IoT, a saúde das pessoas pode ser monitorada constantemente, o que pode ajudar a identificar problemas médicos antes que se tornem sérios. No entanto, ambientes de saúde superlotados podem levar a longos tempos de espera para que os pacientes recebam tratamento. A pandemia global de COVID-19 exacerbou esse problema. Em um ambiente cada vez mais conectado, como as cidades inteligentes, a saúde das pessoas pode ser monitorada a todo momento para que situações que necessitem de suporte médico possam ser identificadas previamente. Assim, as cidades inteligentes adaptariam a alocação hospitalar dos profissionais de saúde às demandas dos pacientes com base em dados previamente coletados. A literatura apresenta alternativas para abordar esse problema tais como o compartilhamento de recursos humanos entre hospitais, ajustes nos turnos de trabalho e de recursos humanos dinamicamente de forma a ajustar a capacidade de atendimento à demanda. No entanto, ainda há necessidade de uma solução que possa ajustar dinamicamente os recursos humanos em múltiplos ambientes de saúde, o que é a realidade das cidades. Nesse contexto, este trabalho apresenta o HealCity, um modelo para otimização do uso de recursos humanos focado em cidades inteligentes que pode monitorar o uso de ambientes de saúde pelos pacientes e adaptar a alocação de profissionais de saúde para atender às suas necessidades. O HealCity usa dados de sinais vitais e técnicas de previsão para antecipar quando a demanda por um determinado ambiente excederá sua capacidade e sugere ações para alocar profissionais de saúde de acordo. Além disso, o HealCity introduz o conceito de *Elasticidade Multinível de Recursos Humanos em Cidades Inteligentes*, uma extensão do conceito de elasticidade de recursos em *Cloud computing* para gerenciar o uso de recursos humanos em diferentes níveis de uma cidade inteligente. Um algoritmo também foi desenvolvido para gerenciar os futuros pacientes na cidade inteligente, identificando automaticamente o ambiente de saúde apropriado para um possível paciente futuro. O HealCity foi avaliado através da simulação de uma cidade inteligente composta por quatro ambientes hospitalares e obteve resultados promissores: comparado aos ambientes médicos com alocações fixas de recursos humanos, o modelo foi capaz de reduzir o tempo de espera para atendimento em até 87,62%, com um aumento no custo de apenas 9,68%. Neste contexto acredita-se que a maleabilidade na gerência de recursos humanos possa ser uma boa alternativa para mitigar problemas de saúde, que aparecem em menor ou maior escala em praticamente todos os países do globo.

Palavras-chave: Internet das Coisas. Saúde. Hospitais Inteligentes. Sensores. Cidades Inteligentes. Recursos Humanos. Elasticidade. Sinais Vitais.

ABSTRACT

Smart cities can improve the quality of life of citizens by optimizing the utilization of resources. In an IoT-connected environment, people's health can be constantly monitored, which can help identify medical problems before they become serious. However, overcrowded healthcare environments can lead to long waiting times for patients to receive treatment. The global COVID-19 pandemic has exacerbated this problem. In an increasingly connected environment, such as smart cities, people's health can be monitored at all times so that situations requiring medical support can be identified in advance. Thus, smart cities would adapt the hospital allocation of healthcare professionals to patient demands based on previously collected data. The literature presents alternatives to address this problem, such as sharing human resources between hospitals, adjustments to work shifts and human resources on-the-fly in order to adjust the capacity to demand. However, there is still a need for a solution that can adjust human resources on-the-fly in multiple healthcare settings, which is the reality of cities. In this context, this work presents HealCity, a smart city-focused model for human resources optimization that can monitor patients' use of healthcare settings and adapt the allocation of health professionals to meet their needs. HealCity uses vital signs data and prediction techniques to anticipate when the demand for a given environment will exceed its capacity and suggests actions to allocate health professionals accordingly. Additionally, HealCity introduce the concept of *Multilevel Human Resources Elasticity in Smart Cities*, an extension of the concept of resource elasticity in *Cloud computing* to manage human resources at different levels of a smart city. An algorithm was also developed to manage future patients in the smart city, automatically identifying the appropriate hospital for a possible future patient. HealCity was evaluated by simulating a smart city composed of four healthcare environments and obtained promising results: compared to medical environment with static professional allocations, the model was able to reduce waiting time for care by up to 87.62%, with an increase in cost of only 9.68%. In this context, it is believed that flexibility in human resources management can be a good alternative to mitigate health problems, which appear to a greater or lesser extent in practically every country in the world.

Keywords: Internet of Things. Health. Smart Hospitals. Sensors. Smart Cities. Human Resources. Elasticity. Vital Signs.

LISTA DE FIGURAS

Figura 1 – Notícias do (i) Jornal *Limerick Post*, onde é possível ver que o problema da superlotação hospitalar não se limita apenas ao Brasil, também atingindo países da União Europeia; (ii) Jornal Zero Hora, onde é possível ver restrições sendo implantadas em alguns hospitais de Porto Alegre em função das superlotações ocorridas recentemente em julho de 2024; e (iii) Jornal g1, onde é possível ver uma superlotação ocorrida recentemente em março de 2024 em hospitais públicos de Campinas. 35

Figura 2 – Exemplo de um cenário onde há mais atendentes examinando do que medicando pacientes em dois ambientes de saúde. Embora o número de pacientes que aguardam exames seja consideravelmente menor do que aqueles que aguardam para receber algum medicamento, gera insatisfação para os pacientes que aguardam medicação. Além disso, existem pessoas com problemas de saúde em suas casas ou locais de trabalho que a qualquer momento podem acabar indo para um desses dois ambientes de saúde. 36

Figura 3 – Componentes do Sistema RFID. 42

Figura 4 – Componentes de uma cidade inteligente. 49

Figura 5 – Topologia da arquitetura hierárquica do VitalSense e seus principais componentes. 51

Figura 6 – O Sistema de Triagem de Manchester. 54

Figura 7 – Elasticidade Vertical e Horizontal. 58

Figura 8 – Configuração da Elasticidade de Recursos. 59

Figura 9 – Escopo da Elasticidade de Recursos. 60

Figura 10 – Objetivos da Elasticidade de Recursos. 61

Figura 11 – Métodos da Elasticidade de Recursos. 61

Figura 12 – Arquitetura da Elasticidade de Recursos. 62

Figura 13 – Provedor da Elasticidade de Recursos. 62

Figura 14 – Modos da Elasticidade de Recursos. 63

Figura 15 – Cenário hipotético em que não existe qualquer estratégia de ajuste de recursos humanos ou indicação de ambiente de saúde adequado. Neste cenário, os pacientes tendem a se dirigir ao ambiente de saúde mais próximo e o número de recursos humanos em cada ambiente de saúde mantém-se constante. . . 73

Figura 16 – Cenário hipotético em que existe uma estratégia de ajuste de recursos humanos e a indicação de ambiente de saúde adequado para um novo paciente. Neste cenário, os pacientes se dirigirem ao ambiente de saúde mais preparado para o seu atendimento e o número de recursos humanos em cada ambiente de saúde pode variar. 74

Figura 17 – Comparação entre os cenários propostos, onde no segundo cenário, existe a impressão de uma *nuvem de recursos humanos*, onde a quantidade de atendentes é provisionada conforme a necessidade. 75

Figura 18 – Estrutura hierárquica baseada em árvore da cidade inteligente com foco no monitoramento dos parâmetros de saúde dos pacientes. As pessoas usam sensores que transmitem sinais vitais para uma infraestrutura *Fog-Cloud* que fornece informações de saúde diretamente para os ambientes de saúde. Nesta estrutura, os cidadãos estão no nível mais baixo, interagindo com dispositivos de ponta, enquanto os ambientes de saúde estão no nível mais alto, interagindo com dados já processados pela infraestrutura em nuvem. 77

Figura 19 – Componentes e visão da rede no modelo HealCity com um (i) serviço web (HealCity App); (ii) serviço de tomada de decisão (Serviços HealCity); (iii) uma rede de sensores para captar os sinais vitais e a localização dos cidadãos; e (iv) gestores dos ambientes de saúde, recursos humanos, pacientes ou pessoas em geral. 79

Figura 20 – Detalhamento da arquitetura e fluxo de dados do modelo HealCity, ilustrando a trajetória dos dados começando nas pessoas que fornecem sinais vitais e localização ao VitalSense e ao módulo *Capture*, respectivamente. Esses dados são posteriormente processados em vários módulos, culminando na exibição das notificações de elasticidade para os recursos humanos, ou dos *dashboards* aos gestores no *HealCity App*. 80

Figura 21 – Entradas e Saídas do modelo HealCity. 82

Figura 22 – Solução hierárquica escalável proposta no modelo HealCity, onde podemos adicionar mais ambientes de saúde em qualquer nó da Fog e quantos nós Fog forem necessários. 83

Figura 23 – Exemplo de Elasticidade Multinível de Recursos Humanos em cidades inteligentes com (i) elasticidade proativa a nível da sala, (ii) elasticidade proativa a nível de ambiente de saúde e (iii) elasticidade proativa a nível regional. . . 85

Figura 24 – Elasticidade preditiva baseada em tempo de espera adotado pelo HealCity. . 85

Figura 25 – Calculando o Número Médio de Atendentes em uma sala do ambiente de saúde entre os instantes de tempo 1 e 5, onde a cada instante de tempo a quantidade de atendentes alocados é alterada, chegando a um número médio de 3 atendentes alocados nessa sala hipotética entre os instantes de tempo. . 88

Figura 26 – Calculando o Tempo Estimado para Atendimento em uma sala do ambiente de saúde com 4 pacientes aguardando e tempo médio de atendimento de 10 minutos, chegando ao Tempo Estimado de Atendimento de 40 minutos. . . 89

Figura 27 – Calculando o tempo estimado de atendimento em uma sala do ambiente de saúde usando alocação paralela de atendentes, com 4 pacientes aguardando, tempo médio de atendimento de 10 minutos e 2 atendentes, chegando ao tempo estimado de atendimento, através do Speedup Elástico de Recursos Humanos, de 20 minutos. 90

Figura 28 – Fluxograma de processos a nível de sala para verificar a possibilidade de alocar ou liberar recursos humanos em uma sala qualquer do ambiente de saúde inteligente. 92

Figura 29 – Exemplo de realocação através do algoritmo de *Round Robin* vislumbrado, com uma fila de 4 salas, e 14 atendentes, onde foi previsto que a Sala 1 precisa alocar 3 atendentes. 94

Figura 30 – Exemplo de realocação através do algoritmo *List Scheduling* adaptado, com uma lista de 4 salas, ordenada pela quantidade de recursos humanos disponíveis para elasticidade, e 12 atendentes, onde foi previsto que a Sala 1 precisa alocar mais 5 profissionais da saúde. 96
Figura 31 – Elasticidade Proativa a Nível Regional. 99
Figura 32 – Classificações do modelo HealCity para os atributos propostos por Al-Dhuraibi et al. (2017). 101
Figura 33 – Exemplo de paciente com sinais vitais alterados, em uma cidade inteligente com três ambientes de saúde disponíveis. Mesmo que haja ambientes de saúde mais próximos, o mais indicado para o paciente é o mais distante, visto que é o ambiente de saúde que possui a melhor capacidade de atendimento no momento. 104
Figura 34 – Fluxograma de interações entre os algoritmos propostos nas diferentes camadas da hierarquia do modelo HealCity. 107
Figura 35 – Fluxo de atendimento do ambiente de saúde. 111
Figura 36 – Imagem da planta teórica desenvolvida. 115
Figura 37 – Representação gráfica dos padrões de carga utilizados na avaliação do HealCity. 118
Figura 38 – Representação gráfica de um exemplo de alocação de recursos humanos ao longo do tempo de atendimento no ambiente de saúde com e sem o uso do modelo HealCity, onde no Ambiente de Saúde Atual têm-se $M2 = 10$ atendentes e no Ambiente de Saúde Inteligente têm-se $M2 = 8,75$ atendentes. . 120
Figura 39 – Leitor RFID Intermec IF2 instalado no laboratório de Internet das Coisas e Aplicações Distribuídas do PPGCA da Universidade do Vale do Rio dos Sinos onde em (A) a antena foi instalada acima da porta e em (B) a antena foi instalada ao lado da porta. 122
Figura 40 – Instante de tempo em que foi identificada a necessidade de (i) realocar atendentes internamente nos ambientes de saúde. Para cada ambiente de saúde, foi proposta a realocação de um enfermeiro da sala de coleta para o eletrocardiograma, e um enfermeiro da sala de medicação para a triagem. 125
Figura 41 – Instante de tempo em que foi identificada a possibilidade de (ii) realocar um atendente entre diferentes ambientes de saúde. Neste caso um doutor foi realocado dos consultórios médicos do Ambiente de Saúde 3 para os consultórios do Ambiente de Saúde 2. 126
Figura 42 – Instante de tempo em que foi identificada a possibilidade de (iii) alocar e (iv) desalocar um recurso humano. Neste caso o Ambiente de Saúde 3 pode liberar um médico alocado fora de seu turno normal de trabalho, enquanto o Ambiente de Saúde 4 alocou um novo recurso humano. 127
Figura 43 – Tempo máximo de espera registrado por ambiente de saúde para cada um dos cenários propostos, C1 (em vermelho), C2 (em laranja) e C3 (em verde), para o (a) Ambiente de Saúde 1, (b) Ambiente de Saúde 2, (c) Ambiente de Saúde 3 e (d) Ambiente de Saúde 4, e a média do tempo máximo de espera encontrado na (e) Cidade inteligente. 128
Figura 44 – Custo dos recursos humanos comparado com a média do tempo máximo de espera na cidade inteligente em (a) C1 e C3 e (b) C2 e C3. 130

Figura 45 – Quantidade elástica de recursos humanos utilizados para cada ambiente de saúde, e para o total da cidade inteligente, no Cenário C3, com o uso da Elasticidade de Recursos Humanos para Saúde em Cidades Inteligentes. Neste gráfico, realocação refere-se as movimentações de profissionais que ocorrem nos ambientes de saúde, em que não há acréscimo ou redução de RH no sistema. 131
Figura 46 – Área de leitura de *tags*-RFID em torno da antena do leitor Intermec IF2. . . 133
Figura 47 – Área de leitura frontal de *tags*-RFID da antena do leitor Intermec IF2. . . . 133
Figura 48 – Instalação proposta das antenas do leitor Intermec IF2 em dois cenários: (A) com uma porta simples e (B) com uma porta dupla, onde em ambos os exemplos as portas possuem 2,1 metros de altura. 134
Figura 49 – Resultados dos testes de carga no *middleware Fosstrak*. 135

LISTA DE TABELAS

Tabela 1 – Lista de dispositivos com sensores capazes de capturar os sinais vitais apresentados. 48
Tabela 2 – Trabalhos selecionados. 66
Tabela 3 – Comparação entre os trabalhos relacionados. 71
Tabela 4 – Nomenclaturas para os formalismos matemáticos do modelo HealCity. . . . 86
Tabela 5 – Distribuições triangulares de probabilidade para tempos de atendimento. . . 112
Tabela 6 – *Thresholds* de tempo para Gerenciamento de Elasticidade do modelo HealCity. 114
Tabela 7 – Funções para expressar os diferentes comportamentos de carga. Em $carga(x)$, x é o instante de tempo que será realizada a entrada de pacientes. 118
Tabela 8 – Métricas para avaliação e resultados esperados em cada um dos cenários propostos. 121
Tabela 9 – Maior tempo de espera e média dos maiores tempos de espera registrados nos cenários propostos para cada ambiente de saúde e para a média da cidade inteligente. 129
Tabela 10 – Quantidade média elástica de Recursos Humanos utilizados em cada ambiente de saúde, juntamente com a média da cidade inteligente para cada cenário proposto. 132
Tabela 11 – Métricas de avaliação e resultados encontrados, onde os melhores resultados no M1 são destacados em verde e os piores em vermelho. Para as métricas Custo de RH e Quantidade Elástica de RH, foi utilizada a média dos resultados hora-a-hora do período simulado. 136
Tabela 12 – Avaliação qualitativa comparativa do modelo HealCity versus Trabalhos Relacionados. 137

LISTA DE ALGORITMOS

Algoritmo 1 – Elasticidade Proativa a Nível de Sala 91
Algoritmo 2 – Elasticidade Proativa a Nível de Hospital 93
Algoritmo 3 – Realocação através do algoritmo de *Human Resource List Scheduling* . . 95
Algoritmo 4 – Desalocação de Recursos Humanos 97
Algoritmo 5 – Seleciona o ambiente de saúde apropriado para o paciente 103

LISTA DE ABREVIATURAS

Aprox.	Aproximadamente
Art.	Artigo
cm	centímetro, -s
Dr.	Doutor, -a
m	metro, -s
ms	milíssegundo, -s
Prof.	Professor, -a
s	segundo, -s
/s	por segundo
Unisinos	Universidade do Vale do Rio dos Sinos
Wi-Fi	Wireless Fidelity

LISTA DE SIGLAS

ACM	*Association for Computing Machinery*
ADT	*Admission, Discharge and Transfer*
ALE	*Application Level Events*
API	*Application Programming Interface*
AR	Auto-Regressivo puro
ARIMA	*Auto-Regressive Integrated Moving Average*
ARMA	*Auto-Regressive and Moving Average*
BLE	*Bluetooth Low Energy*
COVID	*Corona Virus Disease*
CPS	*Cyber-Physical Systems*
CPU	*Central Process Unit*
DMP	*Distributed Movement Prediction*
EPC	*Electronic Product Code*
EPCIS	*Electronic Product Code Information Service*
ERAD-RS	Escola Regional de Alto Desempenho da Região Sul
FAPERGS	Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul
GARCH	*Generalized Auto Regressive Conditional Heteroskedasticity*
GBM	*Gradient Boosted Machines*
GDPR	*General Data Protection Regulation*
GNSS	*Global Navigation Satellite System*
GPS	*Global Positioning System*
GPU	*Graphic Process Unit*
HF	*High Frequency*
HIS	*Hospital Information Systems*
HRLS	*Human Resource List Scheduling*
IA	Inteligência Artificial
IaaS	*Infrastructure as a Service*
ID	Identificador
IEC	*International Electrotechnical Commission*
IEEE	*Institute of Electrical and Electronics Engineers*
IET	*Institution of Engineering and Technology*
IJTM	*International Journal of Technology Management*
IoT	*Internet of Things*

ISO	*International Organization for Standardization*
ITU	*International Telecommunication Union*
JVM	*Java Virtual Machine*
LED	*Light Emissor Diode*
LF	*Low Frequency*
LGPD	Lei Geral de Proteção de Dados Pessoais
MA	Médias Móveis puro
MIT	*Massachusetts Institute of Technology*
NFC	*Near Field Communication*
P2P	*Peer-to-Peer*
PaaS	*Platform as a Service*
PPGCA	Programa de Pós-Graduação em Computação Aplicada
QoS	*Quality of Service*
QoE	*Quality of Experience*
PPSUS	Programa Pesquisa para o SUS
RAM	*Random Access Memory*
RAS	*Ressource Allocation Strategy*
RF	*Radio Frequency*
RH	Recursos Humanos
RFID	*Radio Frequency Identification*
RIS	*Radiology Information System*
RNA	Redes Neurais Artificiais
RSS	*Received Signal Strength*
RSSI	*Received Signal Strength Indicator*
RTLS	*Real-Time Location System*
SBCAS	Simpósio Brasileiro de Computação Aplicada à Saúde
SQL	*Structured Query Language*
SVM	*Support Vector Machines*
SVR	*Support Vector Regression*
STM	Sistema de Triagem de Manchester
SUS	Sistema Único de Saúde
TI	Threshold Inferior
TS	Threshold Superior
UWB	*Ultra Wide Band*

VM	*Virtual Machine*
WPA	*Wi-Fi Protected Access*
WPAN	*Wireless Personal Area Networks*
WSN	*Wireless Sensor Network*

LISTA DE SÍMBOLOS

ºC	Graus Celsius
GHz	GigaHertz
GB	Gigabytes
MHz	MegaHertz
Mbps	Megabits por segundos
Kbits	Kilobits
Kbits/s	Kilobits por segundo
KHz	KiloHertz

SUMÁRIO

1 INTRODUÇÃO **33**
1.1 Motivação 34
1.2 Questão de Pesquisa 36
1.3 Objetivos 38
1.4 Organização do Texto 39

2 REFERENCIAL TEÓRICO **41**
2.1 Internet das Coisas de Saúde 41
2.1.1 Tecnologia RFID 42
2.1.2 *Real-Time Location Systems* 43
2.1.3 Sensores para captura de sinais vitais 44
2.2 Cidades Inteligentes 48
2.2.1 Modelo VitalSense 50
2.2.2 Legislação Aplicável 51
2.3 Processos Hospitalares 52
2.3.1 Ambulatório e Emergência 53
2.3.2 Exames de Diagnóstico e Tratamento 54
2.3.3 Internação Clínica ou Cirúrgica 55
2.4 Técnicas de Predição 55
2.4.1 Modelo Auto-Regressive Integrated Moving Average (ARIMA) 56
2.5 Elasticidade de Recursos 57
2.5.1 Configuração 59
2.5.2 Escopo 59
2.5.3 Objetivo 60
2.5.4 Método 61
2.5.5 Arquitetura 62
2.5.6 Provedor 62
2.5.7 Modo de Operação 62
2.6 Considerações Parciais 64

3 TRABALHOS RELACIONADOS **65**
3.1 Metodologia de Pesquisa e Seleção dos Trabalhos 65
3.2 Using Data Mining to Predict Hospital Admissions From the Emergency Department (GRAHAM et al., 2018) 66
3.3 Sustainable Approach for a Smart Human Resource Management in Healthcare (FRANKE et al., 2018) 66
3.4 A Smart healthcare reward model for resource allocation in smart city (OUEIDA; ALOQAILY; IONESCU, 2019) 67
3.5 ElHealth: Using Internet of Things and Data Prediction for Elastic Management of Human Resources in Smart Hospitals (FISCHER et al., 2020) 67
3.6 Optimizing human resource cost of an emergency hospital using multi-objective Bat algorithm (APORNAK et al., 2021a) 67
3.7 Human resources optimization in hospital emergency using the genetic algorithm approach (APORNAK et al., 2021b) 68
3.8 A cooperative robust human resource allocation problem for healthcare systems for disaster management (HAFEZALKOTOB et al., 2022) 68

3.9 An Analysis of Waiting Time for Emergency Treatment and Optimal Allocation of Nursing Manpower (LIAO; CHU; HO, 2022) . . . 69
3.10 Real-time resource allocation in the emergency department: A case study (DUMA; ARINGHIERI, 2023) . . . 69
3.11 Data-driven hospitals staff and resources allocation using agent-based simulation and deep reinforcement learning (LAZEBNIK, 2023) . . . 69
3.12 Estimation of New Resource Allocation in Hospital's (or Medical Care) Inpatient Department using Discrete Event Simulation (RASIDI et al., 2024) . . . 70
3.13 Análise . . . 70
3.14 Considerações Parciais . . . 72

4 MODELO HEALCITY . . . **73**
4.1 Motivação e Caso de Uso . . . 73
4.2 Decisões de Projeto . . . 76
4.3 Arquitetura . . . 78
4.4 Elasticidade Multinível de Recursos Humanos em Cidades Inteligentes . . . 84
4.4.1 Elasticidade Proativa a Nível de Sala . . . 86
4.4.2 Elasticidade Proativa a Nível de Ambiente de Saúde . . . 93
4.4.3 Elasticidade Proativa a Nível Regional . . . 98
4.5 Gerenciamento de futuros pacientes na cidade inteligente . . . 102
4.6 Considerações Parciais . . . 107

5 METODOLOGIA DE AVALIAÇÃO . . . **109**
5.1 Implementação e Ambiente de Testes . . . 109
5.1.1 Protótipo desenvolvido . . . 109
5.1.2 Simulador do ambiente de saúde . . . 110
5.1.3 Ambiente de Simulação . . . 111
5.2 Parâmetros . . . 111
5.3 Infraestrutura . . . 115
5.4 Cenários . . . 116
5.5 *Workload* de Entrada . . . 117
5.6 Métricas de Avaliação . . . 119
5.7 Metodologia de Testes para o Sistema RTLS . . . 121
5.8 Considerações Parciais . . . 123

6 RESULTADOS E DISCUSSÃO . . . **125**
6.1 Funcionamento do Modelo . . . 125
6.2 Métrica M1: Tempo Máximo de Espera . . . 126
6.3 Métrica M2: Custo de Recursos Humanos . . . 129
6.4 Métrica M3: Quantidade Elástica de Recursos Humanos utilizados . . . 130
6.5 Testes do Sistema RTLS . . . 132
6.6 Discussão . . . 135

7 CONCLUSÃO . . . **139**
7.1 Lições Aprendidas e Transferência de Tecnologia . . . 140
7.2 Contribuições Esperadas . . . 142
7.3 Limitações . . . 143
7.4 Trabalhos Futuros . . . 143
7.5 Publicações . . . 144

REFERÊNCIAS . . . **147**

1 INTRODUÇÃO

A Internet das Coisas (*Internet of Things*, ou IoT) é um conceito onde uma vasta gama de objetos e "coisas" são conectadas através de uma estrutura de rede e coletam, analisam e compartilham dados com outros objetos, programas e plataformas (EGHMAZI et al., 2024). O desenvolvimento desse paradigma está em crescimento constante (OLSZEWSKI; IWANOWSKI; GRANISZEWSKI, 2024). Isso ocorre por causa dos esforços contínuos da comunidade de pesquisadores, em função do IoT permitir aplicações ilimitadas para resolver necessidades também ilimitadas tornando a vida mais confortável e agradável (KUMAR, 2024). Além disso, a Internet das Coisas torna possível para nossos dispositivos interagirem não apenas uns com os outros, mas com serviços e outras pessoas em uma escala global (QIN; SHENG; CURRY, 2020). Já os Sistemas Ciber-Físicos (*Cyber-Physical Systems*, ou CPS) são sistemas que integram sensores, computação, controle e redes de comunicação em objetos físicos e infraestrutura, onde os mesmos são conectados através da Internet e acabam interagindo entre si para executar uma tarefa comum (SINGH et al., 2022). O conceito de cidade inteligente refere-se ao uso de várias tecnologias de informação, sensoriamento e computação como a Internet das Coisas, os Sistemas Ciber-Físicos e as Redes de Sensores Sem-fio (*Wireless Sensor Networks*, ou WSN) (ALI et al., 2023; SIDOROV, 2023; JANANI et al., 2023).

Cidades inteligentes geralmente são consideradas uma integração de tecnologia de informação e comunicação, dispositivos físicos conectados a rede, infraestrutura social e de negócios para melhorar o bem-estar dos cidadãos e a capacidade dos serviços governamentais na tomada de decisões nas operações da cidade (ALSHAMAILA et al., 2023; LUCIC et al., 2023). As cidades inteligentes podem aumentar a eficiência da utilização de recursos, otimizar a gestão e os serviços urbanos, e melhorar a qualidade de vida dos cidadãos (XIUFAN; DECHENG, 2023; WANG; ZHOU, 2023). Inúmeros sistemas podem ser imaginados para cidades inteligentes, incluindo edifícios inteligentes (SEN et al., 2023), gerenciamento de resíduos (RAI et al., 2024), sensores ambientais (ZHANG, 2023), gerenciamento inteligente de energia (AKRAM et al., 2024) e melhorias na saúde (FISCHER et al., 2024).

De acordo com Rejeb et al. (2023), as tecnologias baseadas em IoT estão transformando as aplicações de saúde, permitindo abordagens mais personalizadas, preventivas e colaborativas para o atendimento ao paciente e promovendo ambientes de saúde aprimorados. Em cidades cada vez mais inteligentes, os dados dos pacientes podem ser capturados da forma mais transparente possível, com os sinais vitais sendo coletados em tempo real por sensores presos às roupas das pessoas (KALASIN; SURAREUNGCHAI, 2023) ou nos ambientes por onde elas passam (TRIPATHY et al., 2023). Além disso, a IoT usada em cidades inteligentes fornece meios para que os sistemas de saúde extraiam e analisem dados, combinados com técnicas de aprendizado de máquina, para detectar precocemente distúrbios de saúde (BAUCAS; SPACHOS; PLATANIOTIS, 2023). Com os sinais vitais e os dados de outros pacientes sendo capturados ininterruptamente, não é apenas possível identificar antecipadamente os distúrbios de saúde que

podem ocorrer, mas também planejar e estabelecer contramedidas para mitigar os danos que podem ser causados.

1.1 Motivação

Dentre as diversas situações-problemas do nosso dia-a-dia que podem ser minimizadas utilizando-se do paradigma das cidades inteligentes, existe uma que todas as pessoas se deparam ou vão se deparar em algum momento da vida: a necessidade de atendimento médico. Assim os hospitais estão entre os mais importantes pontos de atendimento capazes de garantir um tratamento adequado à população. Diversos são os ambientes, recursos e processos dentro de um ambiente de saúde para realizar essa tarefa. Considerando a importância de tais ambientes, aumentar a eficiência com que os recursos e processos de um ambiente de saúde são controlados torna-se uma preocupação central. Tal preocupação é particularmente relevante no contexto de países subdesenvolvidos, onde o elevado número de pacientes associado à falta de recursos leva a tempos de espera excessivamente elevados (KOMALASARI, 2023; DENG et al., 2023; BIYA et al., 2022). O problema da superlotação de hospitais é de certa forma antigo (MCCARTHY, 2011) e extremamente recente (KOMALASARI, 2023; DENG et al., 2023). Conforme pode ser visto na Figura 1 o problema da superlotação de ambientes médicos na literatura está em consonância com o que ocorre de fato nas cidades e é presenciado pela população em geral, o que mostra que é um problema que ainda persiste e que deve ser tratado. A pandemia global da doença do Coronavírus (COVID-19) exacerbou esse problema e trouxe um enorme impacto no desenvolvimento e operação das cidades ao redor do mundo e no cotidiano das pessoas (XIANG et al., 2023).

De acordo com Butean et al. (2015), por mais fácil ou complicada que seja uma situação, se a equipe médica não reagir a tempo, tudo torna-se duvidoso e inseguro com relação a saúde dos pacientes. Assim, os profissionais de saúde desempenham um papel fundamental no bem-estar dos pacientes (TOOMAN; PIIRMAN; KÄRP, 2022). Esses profissionais vão desde a equipes de enfermeiros que realizam procedimentos de triagem e pequenos tratamentos até aos doutores que atendem as mais diversas especialidades médicas. Nesse tipo de cenário, uma alocação estática de profissionais de saúde para setores de saúde pode ser ineficiente, pois alguns profissionais podem ser mal alocados em setores de baixa demanda, levando à falta de profissionais em setores de alta demanda, conforme exemplificado na Figura 2. De fato, a alocação de atendentes deve sempre se **adaptar** às condições atuais dos setores de saúde. Além disso, em cidades inteligentes, sinais vitais capturados das pessoas ininterruptamente podem ser utilizados para identificar a demanda futura de atendimento. Neste contexto, surge a necessidade dos sistemas dos ambientes de saúde interagirem entre si, para identificar a disponibilidade e uso dos profissionais de saúde alocados na rede de atendimento, de forma a propor ajustes baseados tanto na demanda já existente nos hospitais, como na demanda futura baseada nos sinais vitais em declínio sendo capturados da população. Atualmente, sistemas utilizados pelos órgãos go-

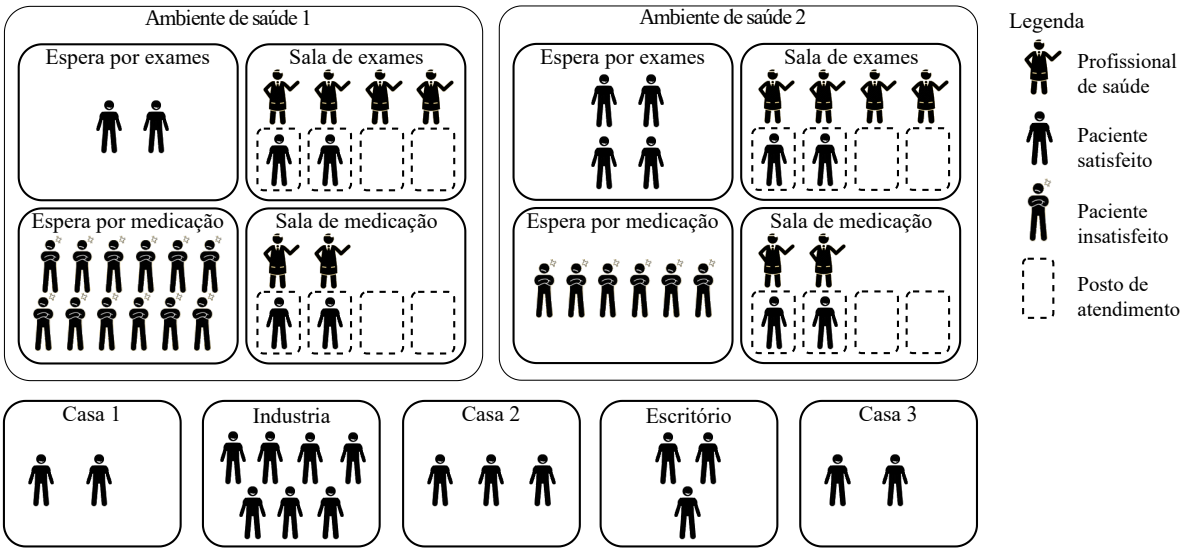
Figura 1 – Notícias do (i) Jornal *Limerick Post*, onde é possível ver que o problema da superlotação hospitalar não se limita apenas ao Brasil, também atingindo países da União Europeia; (ii) Jornal Zero Hora, onde é possível ver restrições sendo implantadas em alguns hospitais de Porto Alegre em função das superlotações ocorridas recentemente em julho de 2024; e (iii) Jornal g1, onde é possível ver uma superlotação ocorrida recentemente em março de 2024 em hospitais públicos de Campinas.

Fonte: English (2024), Zanchettin (2024) e Moreira (2024).

vernamentais de saúde já são capazes de identificar os leitos dos ambientes de saúde vagos e ocupados em sua área de governança e ainda disponibilizar *dashboards* públicos para consulta pela população (RIO GRANDE DO SUL, 2024). Porém não existem sistemas para identificar o uso dos recursos humanos (RH), nem para disponibilizar *dashboards* para consulta dos gestores dos ambientes de saúde ou órgãos governamentais.

Neste contexto, algumas abordagens na literatura científica tentam atuar no gerenciamento de recursos humanos. Este é o caso do trabalho proposto por Graham et al. (2018), que teve como objetivo prever a chegada de pacientes ao pronto-socorro de um hospital para preparar adequadamente a alocação de equipe médica. Porém, o objetivo era exclusivamente identificar padrões específicos nos dados para apoiar os gestores na tomada de decisão, sem de fato propor contramedidas para melhorar a alocação de recursos de saúde. Apornak et al. (2021a) investigaram maneiras de otimizar o custo dos recursos humanos necessários em uma emergência, identificando uma possível alocação ótima de recursos humanos no ambiente de saúde estudado. Entretanto, os autores não propuseram formas para os recursos se adaptarem dinamicamente à demanda dos setores de saúde. Hafezalkotob et al. (2022) investigaram o problema da realocação de recursos de saúde em situações de desastres. Os autores propuseram um formulação linear robusta para permitir a observação de comportamentos colaborativos entre centros mé-

Figura 2 – Exemplo de um cenário onde há mais atendentes examinando do que medicando pacientes em dois ambientes de saúde. Embora o número de pacientes que aguardam exames seja consideravelmente menor do que aqueles que aguardam para receber algum medicamento, gera insatisfação para os pacientes que aguardam medicação. Além disso, existem pessoas com problemas de saúde em suas casas ou locais de trabalho que a qualquer momento podem acabar indo para um desses dois ambientes de saúde.



Fonte: Elaborada pelo autor.

dicos, incluindo transferência de funcionários, leitos e pacientes entre hospitais. Entretanto, assim como nos trabalhos anteriormente citados, os autores não propuseram ajuste dos recursos humanos à demanda de pacientes *on-the-fly*.

Uma visão detalhada sobre a literatura na área e as lacunas encontradas podem ser vistas no Capítulo 3. Com base nesta análise, pode-se dizer que apesar de existirem vários trabalhos atuando na análise e gerenciamento de recursos humanos em ambientes de saúde, não foram encontradas abordagens capazes de ajustar a capacidade de atendimento dos ambientes de saúde dinamicamente em ambientes multi-hospitalares.

Dessa forma, a motivação científica para o presente trabalho está na necessidade de se propor alternativas e soluções a esse problema, uma vez que o mesmo não é tratado de forma eficiente em trabalhos científicos considerados estado da arte, tais como os propostos por Graham et al. (2018), Apornak et al. (2021a) e Hafezalkotob et al. (2022). Dada a situação, isso justifica a criação de um novo modelo para gerenciamento dinâmico da capacidade de atendimento de múltiplos ambientes de saúde.

1.2 Questão de Pesquisa

Manter uma qualidade de serviço desejável para aplicações em ambientes de saúde é um grande desafio, pois o número de pacientes pode mudar dinamicamente. O aumento do número de pessoas necessitando de atendimento médico, em função da sazonalidade ou até de surtos de doenças, aumenta a complexidade do monitoramento de informações como sinais vitais em

tempo real, pois quanto maior o número de pessoas, maior a quantidade de informações que os sistemas precisam manipular. Além de lidar com o volume de dados, os sistemas também devem gerenciar os ambientes para que estes pacientes sejam atendidos da melhor forma possível, com o menor tempo de espera, e com a menor exposição possível a outras doenças que possam estar circulando em ambientes lotados de pacientes. Portanto, é necessário que os sistemas de saúde em cidades inteligentes sejam capazes de identificar e lidar com tais situações para manter um determinado nível de qualidade de serviço desejável, atenuando o tempo de espera por atendimento de saúde das pessoas. Assim, para nortear esta pesquisa, a seguinte afirmação define a atual hipótese de tese:

Hipótese. *Um modelo para controle de recursos humanos em ambientes de saúde de cidades inteligentes é capaz de diminuir o* ***tempo de espera*** *por atendimento nos ambientes de saúde, sem aumentar proporcionalmente o* ***custo****, mantendo este* ***tempo adequado à legislação pertinente****, considerando o* ***número*** *de pacientes nas cidades.*

Primeiramente, **tempo de espera** refere-se ao tempo que os pacientes com algum problema de saúde precisam aguardar dentro dos ambientes de saúde para receber o tratamento adequado. Nesse contexto, diminuir o tempo de espera por atendimento nos ambientes de saúde das cidades, tem com objetivo melhorar o atendimento, fazendo com que os pacientes recebam o tratamento adequado o mais rápido possível, diminuindo, inclusive, a exposição deles à outros pacientes, algo tão importante em tempos de surtos de doenças e pandemias. O **custo** refere-se aos custos que os ambientes de saúde possuem relacionados aos recursos humanos. Com a diminuição do tempo de espera, pressupõe-se que haverá um aumento nos custos. Neste contexto, a afirmação de não aumentar proporcionalmente o custo, tem como objetivo garantir que o ganho em tempo de espera seja superior ao aumento nos custos dos ambientes de saúde. O **tempo adequado à legislação pertinente** refere-se aos possíveis limites de tempo de espera estabelecidos em leis na área de saúde para triagem de pacientes. Nesse contexto, com base na legislação existente, é estabelecido um tempo máximo de espera, ao qual deve ser cumprido para evitar a deterioração do estado de saúde dos pacientes. Por fim, o **número** de pacientes é a quantidade de pessoas precisando de atendimento médico nas cidades, independentemente de estarem em suas casas ou aguardando nos ambientes de saúde. O número de pacientes muda ao longo do tempo, aumentando ou diminuindo a carga nos sistemas de saúde das cidades inteligentes. Com base na hipótese, foi definida a seguinte questão de pesquisa para ser trabalhada nesta proposta de tese:

Questão de Pesquisa. *Quais estratégias um modelo de* ***gerenciamento de recursos humanos*** *para saúde precisa empregar para* ***identificar a demanda por serviços*** *de saúde e propor uma* ***melhor alocação*** *tanto de* ***recursos humanos*** *quanto de* ***atendimento****, considerando a* ***quantidade de pacientes****?*

Dessa forma, este trabalho tem como foco de pesquisa identificar como seria um modelo para, através do uso de sinais vitais capturados por serviços de saúde, e disponibilizados através

de uma infraestrutura que combina *Edge, Fog e Cloud computing*, informar aos ambientes de saúde sobre a demanda de pacientes que estão necessitando ou podem vir a necessitar de atendimento médico. Ao mesmo tempo, identificar como garantir um registro do uso e da alocação dos funcionários nos setores em todo o tempo em que estiverem a disposição dos ambientes de saúde. Utilizando dessas informações, também identificar como esse modelo poderia prever possíveis situações em que haveria uma demanda de pacientes superior a quantidade de funcionários disponíveis, trazendo soluções para a realocação (remanejamento de pessoas dentro dos setores de um mesmo ambiente de saúde, ou entre diferentes ambientes médicos) e alocação (chamar pessoas que estariam em casa para o atendimento) dos mesmos para se ajustar a demanda. Além disso, este trabalho também tem como foco de pesquisa gerar *dashboards* com a situação *on-the-fly* dos ambientes de saúde de quaisquer regiões dentro de uma cidade inteligente para disponibilização para os gestores dos ambientes de saúde e os órgãos governamentais.

1.3 Objetivos

Buscando fornecer um modelo adaptável como alternativa à tradicional alocação estática de pessoal médico, este estudo propõe o HealCity. O HealCity funciona ajustando a alocação da equipe médica de ambientes de saúde inteligentes com base em uma abordagem de elasticidade proativa. Em particular, o HealCity usa sinais vitais como (i) frequência respiratória, (ii) temperatura corporal, (iii) frequência cardíaca, (iv) variabilidade da frequência cardíaca; e (v) saturação de oxigênio, para estimar a demanda de problemas de saúde dos pacientes. Tais estimativas permitem identificar situações em que a disponibilidade de pessoal dificilmente atenderá à demanda. A partir desses dados, o HealCity propõe o ajuste proativo de recursos humanos em ambientes de saúde inteligentes, adaptados à demanda de atendimento ao paciente. O HealCity propõe uma alocação eficiente do corpo médico, realocando esses profissionais e também alocando novos recursos humanos às áreas com maior necessidade. Em vez de ajustar a capacidade de atendimento dos ambientes médicos com base na chegada de pacientes em emergências, é possível ajustar essa capacidade antecipadamente, propondo inclusive filas de atendimento virtuais. Ao invés de o paciente ir ao ambiente de saúde e esperar em ambientes superlotados, assim que o modelo identifica distúrbios de saúde, deve automaticamente inserir aquele paciente em uma fila virtual. Em seguida, ajustar a capacidade de atendimento do ambiente de saúde de acordo com a demanda de emergência atual e a demanda futura nas filas virtuais. Por fim, o modelo deve notificar o paciente da necessidade de ir ao ambiente médico somente quando a capacidade de atendimento já estiver ajustada à demanda. Em tempos de COVID-19, o paciente tem exposição mínima a outros pacientes, caso a fila seja virtual com pessoas em casa esperando para receber atendimento médico. A ideia é sempre oferecer o melhor tempo de espera possível para os pacientes independente da carga de trabalho (quantidade deles dentro ou fora dos ambientes de saúde). Portanto, a seguinte afirmação define o objetivo

principal desta pesquisa:

Desenvolver um modelo que compreenda arquitetura, definição de players e funções, bem como novos algoritmos e conceitos, para ***controle de recursos humanos*** *em ambientes de saúde localizados em cidades inteligentes, utilizando-se de sensores para* ***medir sinais vitais*** *considerados estado da arte, para* ***registrar a relação entre a capacidade de atendimento*** *de pacientes dos hospitais com a* ***quantidade de funcionários*** *que estão prestando atendimento.*

Para completar o objetivo principal, a pesquisa deve cumprir vários objetivos secundários. Esses objetivos secundários apresentados a seguir, apoiam o desenvolvimento do modelo e as aplicações necessárias para a sua validação.

(i) Reduzir o tempo de espera por atendimento nos ambientes de saúde;

(ii) Otimizar o tempo dos profissionais de saúde, de forma a evitar a subutilização dos recursos humanos em determinado ambiente de saúde e a superutilização em outro;

(iii) Identificar quais são os ambientes de saúde mais próximos aos pacientes com capacidade para atendê-los; e

(iv) Permitir a visualização da situação dos ambientes de saúde de quaisquer regiões dentro de uma cidade inteligente por parte dos administradores desses ambientes e dos representantes governamentais através de *dashboards*.

1.4 Organização do Texto

Esta tese está organizada em sete capítulos. Após o capítulo de Introdução, o Capítulo 2 trará o Referencial Teórico utilizado para este trabalho, demonstrando o estado da arte das tecnologias que estão relacionadas ao tema proposto, apresentando os conceitos de Internet das Coisas e cidades inteligentes, bem como as tecnologias utilizadas para ambos os paradigmas. Além disso, este capítulo também apresentará um estudo sobre os processos hospitalares e técnicas de predição computacional, apresentando os conceitos de séries temporais, e modelos matemáticos. Por fim trará um estudo sobre a elasticidade de recursos computacionais.

O Capítulo 3 trará uma relação de Trabalhos Relacionados a este, apresentando a metodologia de pesquisa utilizada para seleção dos estudos. O capítulo apresentará um breve resumo dos trabalhos escolhidos e as principais características de cada um deles, apresentando uma analise comparativa entre eles, mostrando pontos interessantes de cada uma das abordagens e as lacunas identificadas.

O Capítulo 4 apresentará o modelo HealCity, proposto para tentar solucionar o problema levantando na Introdução. Trará uma visão da arquitetura necessária para o sistema e das decisões de projeto, bem como apresentará o conceito de Elasticidade Multinível de Recursos Humanos

em cidades inteligentes. Além disso, este capítulo também apresentará uma abordagem para gerenciamento de futuros pacientes em cidades inteligentes.

No Capítulo 5 será expressa a metodologia de avaliação do modelo, apresentando a implementação desenvolvida e o ambiente utilizado para a simulação do sistema. Também definirá os cenários de testes escolhidos, juntamente com os parâmetros, a infraestrutura e a carga de trabalho que foram utilizados, bem como as métricas que foram elaboradas para análise e avaliação do modelo proposto.

O Capítulo 6 trará a apresentação dos resultados encontrados para cada um dos cenários de testes propostos na metodologia de avaliação. Além disso, apresentará uma análise dos resultados encontrados em cada uma das métricas propostas, de forma a avaliar o modelo proposto. Por fim, trará uma discussão sobre os resultados obtidos.

O Capítulo 7 apresentará as considerações finais desta tese, bem como as lições aprendidas. Nele serão elencadas as contribuições do modelo, as limitações do HealCity e alguns possíveis trabalhos futuros que poderiam ser feitos para melhoria ou complementação do sistema proposto. Além de apresentar as produções científicas publicadas durante o período desta tese.

2 REFERENCIAL TEÓRICO

Este capítulo tem como objetivo apresentar os conceitos fundamentais para o desenvolvimento deste trabalho, sendo dividido em seis seções principais. Primeiramente, na Seção 2.1, serão apresentados os conceitos referentes ao paradigma Internet das Coisas de Saúde, importantes a este trabalho por apresentar a base relacionada aos sensores para captura de sinais vitais e localização dos pacientes. Após, na Seção 2.2, serão descritos os conceitos relacionados às cidades inteligentes, bem como será apresentado o modelo VitalSense e um pouco da legislação pertinente ao modelo. A seguir, na Seção 2.3, serão descritos os processos hospitalares utilizados na área da saúde, importantes para descrever como funciona o dia-a-dia hospitalar. A seguir, na Seção 2.4, serão apresentadas algumas técnicas computacionais de predição, que são úteis para prever e antecipar a demanda de uso das salas dos ambientes de saúde. Após, a Seção 2.5 apresentará os conceitos referentes a elasticidade de recursos computacionais, utilizada como base para a elasticidade de recursos humanos proposta neste trabalho. Por fim, a Seção 2.6 apresenta as considerações finais deste capítulo.

2.1 Internet das Coisas de Saúde

A Internet das Coisas é um paradigma que vem ganhando terreno rapidamente no cenário atual da comunicação computacional. A ideia se baseia na presença de uma variedade de coisas ou objetos ao nosso redor que estão conectados ao mundo virtual, cada um podendo interagir com os outros e cooperar para alcançar objetivos em comum (LOMBARDI; PASCALE; SANTANIELLO, 2021). Além disso, a Internet das Coisas torna possível para nossos dispositivos interagirem não apenas uns com os outros, mas com serviços e outras pessoas em uma escala global (AKEJU; BUTAKOV; AGHILI, 2018). O mundo está passando por uma transformação tecnológica sem precedentes, que evolui de sistemas isolados para "coisas" onipresentes, habilitadas para a internet, capazes de gerar e trocar grandes quantidades de dados valiosos. A Internet das Coisas não é mais um conceito e sim uma nova realidade que está mudando completamente a vida cotidiana, aumentando a produtividade dos negócios e melhorando a eficiência até de governos. Além disso, a IoT também promete revolucionar as aplicações da área da saúde, permitindo uma forma de cuidar dos pacientes mais personalizada, preventiva e colaborativa (REJEB et al., 2023).

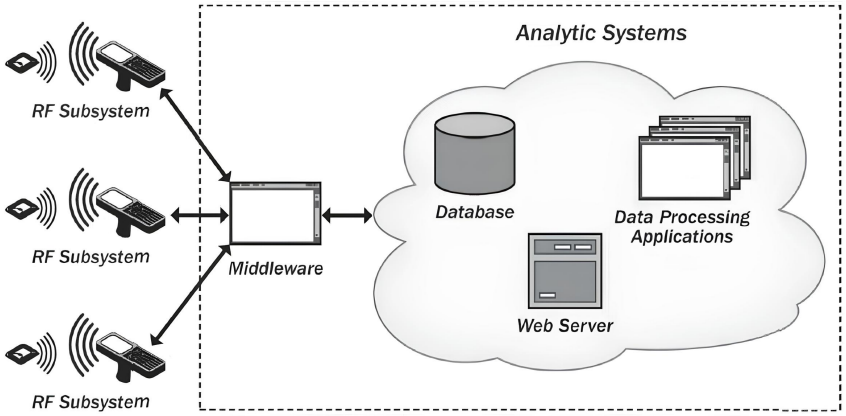
O avanço das tecnologias digitais de terceira (3G) e quarta geração (4G) propiciou o surgimento do ecossistema de digitalização inteligente e conectada nos sistemas de saúde, delineando um modelo de Internet das Coisas de Saúde, ou saúde 4.0, que incorpora elementos de Internet das Coisas, interoperabilidade, *blockchain*, aprendizado de máquina e *Big Data* (BERTONI et al., 2022). A necessidade de uma resposta pronta à crise desencadeada pela pandemia do COVID-19 acelerou a implementação do modelo de saúde 4.0 (BERTONI et al., 2022). Segundo Rodrigues et al. (2023), diversas tecnologias podem ser utilizadas para a saúde 4.0, dentre

elas as tecnologias RFID (Subseção 2.1.1) e *Real-Time Location Systems* (Subseção 2.1.2), bem como diversos sensores podem ser utilizados para captura de sinais vitais (Subseção 2.1.3).

2.1.1 Tecnologia RFID

Segundo Suresh e Chakaravarthi (2022), a tecnologia RFID é uma tecnologia de identificação automática, que funciona através de sinais de rádio, recuperando e armazenando dados remotamente. Conforme Kassim et al. (2022), ela é empregada em centenas, ou talvez até milhares, de sistemas com o objetivo de coletar dados sobre objetos de forma a identificar maneiras de resolver problemas do dia-a-dia. As *tags* suportam além da identificação, uma variedade de informações adicionais, que podem ir desde o nome da fabricante, e tipos de produtos, até a fatores ambientais como, por exemplo, a temperatura (WANT, 2006). Para que um sistema RFID funcione são necessários alguns componentes (conforme pode ser visto na Figura 3), dentre eles podemos citar os três principais: *Tags* RFID (*Transponder*), o Leitor RFID (*Transceiver*), e o *Middleware* (PEREIRA, 2012).

Figura 3 – Componentes do Sistema RFID.



Fonte: Moreira (2008).

As *tags* são pequenos circuitos que tem como objetivo atrelar dados lógicos a um objeto físico. Cada *tag* possui um mecanismo próprio para armazenar dados e uma forma de enviá-los, assim podemos dizer que elas possuem no mínimo três partes: uma antena, um chip semicondutor conectado a essa antena e alguma forma de encapsulamento (KASSIM et al., 2022). Elas podem ser divididas em dois principais grupos: as *tags* passivas, que são alimentadas diretamente pelo leitor e só são ativadas quando estão no raio de leitura, e as *tags* ativas que possuem fonte de energia própria, mantendo-se sempre em atividade. Já um leitor ou *reader* é o aparelho que permite comunicação, interpretação e leitura dos dados contidos em *tags* RFID que estão dentro do raio de alcance de sua antena (PEREIRA, 2012). Por fim, o *middleware* é o elemento responsável pela filtragem, agregação e distribuição de dados, leitura e gravação das informações contidas nas *tags*, e gestão dos leitores (PEREIRA, 2012). Através do uso do RFID se torna possível rastrear quando uma pessoa acessa ou sai de uma determinada região, sendo

utilizado para *Real-Time Location Systems* (RODRIGUES et al., 2023).

2.1.2 *Real-Time Location Systems*

De acordo com Bendavid et al. (2024), Sistemas de Localização em Tempo Real, ou *Real-Time Location Systems*, comumente conhecidos pela sigla RTLS, são sistemas para a identificação e rastreamento da localização de pessoas, produtos e equipamentos em tempo real ou quase em tempo real. Um RTLS geralmente consiste em leitores fixos (sensores de localização) recebendo sinais sem fio de pequenos crachás de identificação ou *tags* anexadas a objetos ou pessoas, para determinar onde essas entidades estão localizadas dentro de um ambiente interno ou externo. Para isso, cada *tag* deve transmitir periodicamente seu próprio ID que deve ser registrado para o ativo ou pessoa que está sendo representado. Dependendo da tecnologia utilizada, o sistema é capaz de identificar e localizar as *tags* nos ambientes e consequentemente as pessoas e objetos que elas representam. Assim, quando alguém necessita localizar algum desses objetos ou pessoas, é só realizar uma consulta no sistema (SELVI et al., 2022).

As informações de localização geradas pelos RTLS geralmente não incluem detalhes de navegação completos ou contínuos, como velocidade, direção de movimento ou orientação espacial de objetos e pessoas rastreados (BOULOS; BERRY, 2012). Os padrões para o RTLS incluem a série de normas ISO/IEC 24730 (INTERNATIONAL ORGANIZATION FOR STARDARTIZATION (ISO); INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC), 2014), entretanto a norma não cobre toda a gama de tecnologias RTLS existentes (BOULOS; BERRY, 2012). Ainda de acordo com Selvi et al. (2022) e Boulos e Berry (2012), emergências, hospitais, sistemas de saúde domiciliares, e até aplicações assistenciais para vida doméstica cotidiana podem ser potencialmente beneficiadas com o uso de uma solução RTLS apropriada.

Segundo Boulos e Berry (2012), em um RTLS, o software de localização se comunica com tags e sensores para determinar a localização das pessoas e objetos. O software retransmite essas informações para *middlewares* e aplicações especializadas. O *middleware* é responsável por intermediar a comunicação entre os principais componentes do RTLS (*tags*, sensores de localização e software) e uma variedade de aplicações capazes de exibir e explorar a localização em tempo real e as informações de status das entidades rastreadas. Essas aplicações podem variar de simples interfaces para consultar e exibir as informações sobre a localização das pessoas e objetos, até uma integração mais abrangente de sistemas comerciais ou empresariais, por meio de aplicações abertas para enriquecer esses sistemas com informações de localização, de forma a torná-los capazes de realizar uma variedade de tarefas e operações de gerenciamento. Um exemplo de uso seria em sistemas hospitalares ADT (*Admission, Discharge and Transfer*, ou Admissão, Descarga e Transferência), HIS (*Hospital Information Systems*, ou Sistemas de Informações Hospitalares), RIS (*Radiology Information Systems*, ou Sistemas de Informações Radiológicas), sistemas de sala de cirurgia e sistemas de gerenciamento de leitos.

De acordo com Boulos e Berry (2012), ao rastrear a localização física de um ativo ou de

um indivíduo, dependendo das necessidades das aplicações, pode ser necessário saber a posição absoluta (coordenadas absolutas, como latitude, longitude e altitude), posição relativa (distância em três dimensões com referência a um ponto fixo) ou posição simbólica (presença em uma área específica, por exemplo, o cirurgião está na sala de operações). Assim, soluções RTLS estão disponíveis para atender aos requisitos de diferentes aplicações, independentemente da precisão da localização, em diferentes resoluções:

- **Localização baseada em presença:** O RTLS retorna a localização da *tag* quando a mesma está presente em uma determinada área;
- **Localização no nível de sala:** O RTLS retorna em qual sala especifica a *tag* está localizada (Por exemplo, o sistema é capaz de informar a sala exata onde uma determinada enfermeira está localizada);
- **Localização no nível de sub-sala:** o RTLS é capaz de identificar a *tag* em uma parte específica da sala (Por exemplo, em um quarto de um hospital com vários pacientes, saber em qual leito a *tag* está localizada;
- **Localização em ponto específico:** A localização da *tag* é retornada por um ponto específico, como entrada ou saída de alas e locais. Ao monitorar os tempos em que as *tags* foram detectadas em pontos, é possível determinar a direção de movimento das mesmas;
- **Localização com precisão:** A localização exata da *tag* é identificada precisamente em um mapa e relatada como posição absoluta ou relativa.
- **Localização por associação:** A localização de uma *tag* é retornada com relação a proximidade com outra *tag* (um exemplo de uso seria se um paciente possui uma *tag* e um monitor cardíaco possui outra *tag*, é possível identificar quando o monitor está sendo usado para aquele paciente);

Todas as tecnologias RTLS compartilham o objetivo comum de determinar a localização de ativos e indivíduos com a precisão necessária para a aplicação final. Cada tecnologia terá sucesso à sua maneira, desde que tenha sido cuidadosamente adaptada para as aplicações. Para certos sistemas, o uso das soluções RTLS combinadas que incorporam tecnologias complementares, como infravermelho e RFID, podem fornecer níveis de precisão e flexibilidade inatingíveis por qualquer tecnologia concorrente (BOULOS; BERRY, 2012). Além do rastreamento da movimentação de pessoas um mesmo dispositivo IoT pode ser utilizado para monitoramento de sinais vitais, através de sensores embarcados (RODRIGUES et al., 2023).

2.1.3 Sensores para captura de sinais vitais

De acordo com Rodrigues et al. (2021), empregar vários dispositivos sensores para monitorar o estado de saúde das pessoas diariamente não apenas é possível, como é essencial para

monitorar a propagação de novas doenças. As pessoas podem usar dispositivos vestíveis para transmitir seus sinais vitais para a infraestrutura de saúde pública em tempo real. Por meio do emprego de *Fog* e *Cloud computing*, é possível fornecer vários serviços de atendimento aos pacientes em um novo paradigma em que o sistema público de saúde pode fornecer serviços inteligentes sob demanda. Por exemplo, um sistema de vigilância de saúde pode prever problemas de saúde de uma pessoa usando sensores de sinais vitais. Assim, o sistema de saúde pode auxiliar os pacientes antes que eles mesmos perceberem que precisam de ajuda médica.

Os diversos sensores disponíveis atualmente permitem desenvolver sistemas e plataformas para acompanhar a saúde das pessoas em tempo real. Dispositivos vestíveis diminuem a complexidade para os pacientes, oferecendo conectividade sem fio para a transmissão rápida de dados em tempo real. Além disso, pulseiras inteligentes e relógios inteligentes estão se tornando populares, pois estão amplamente disponíveis no mercado a um custo reduzido. Além de serem fáceis de usar, esses dispositivos não precisam de um especialista para configurá-los (RODRIGUES et al., 2021). Segundo Rosa Righi et al. (2023), destacam-se cinco sinais vitais mais marcantes detectados por sensores embarcados em dispositivos não invasivos, tais como o Fitbit Versa 3:

(i) frequência respiratória;

(ii) temperatura corporal;

(iii) frequência cardíaca;

(iv) variabilidade da frequência cardíaca; e

(v) saturação de oxigênio.

A *frequência respiratória* é a razão de respirações em um minuto. Geralmente, uma frequência respiratória normal fica entre 12-20 respirações por minuto. Segundo Rosa Righi et al. (2023), a falta de ar é um dos sintomas mais relevantes para doenças respiratória, tais como a COVID-19. Consequentemente, o estado do paciente é refletido diretamente nas medições de sua frequência respiratória. Portanto, a frequência respiratória é um parâmetro de saúde crucial para o monitoramento constante de infecções por COVID-19. Atualmente, os dispositivos vestíveis podem fornecer um monitoramento contínuo da frequência respiratória por meio de sensores. A maioria das estratégias requer a colocação de sensores no tórax, abdômen, pescoço ou nariz do paciente. Com o avanço e concepção de novas tecnologias para IoT, sensores cada vez menores e menos intrusivos estão sendo projetados. Mesmo que os sensores estejam se tornando leves e pequenos, sua colocação ainda pode desagradar no uso diário. Assim, surgem novas soluções em algoritmos para derivar a frequência respiratória de sensores ópticos embutidos em *smartwatches* e pulseiras inteligentes. Este é o caso dos sensores desenvolvidos pela *Fitbit*[1]. As pulseiras e relógios inteligentes da empresa podem medir a frequência respiratória

[1]https://www.fitbit.com

através da medição da frequência cardíaca do paciente. O período entre os batimentos cardíacos diminui durante a inspiração e aumenta durante a expiração. Assim, o sensor óptico dos batimentos cardíacos mede as pequenas diferenças entre esses períodos enquanto o paciente está dormindo para definir a frequência respiratória média daquela noite. O processo de aquisição de dados exige que o indivíduo use o dispositivo por pelo menos três horas durante o sono.

A *temperatura corporal* é essencial para monitorar possíveis anormalidades de saúde. Segundo Rosa Righi et al. (2023), a temperatura corporal acima de 37,3ºC caracteriza febre, indicando que o corpo está tentando combater alguma doença ou infecção. A temperatura corporal pode ser medida usando termômetros tradicionais de forma axilar, oral e retal. No entanto, soluções recentes empregam tecnologias menos intrusivas que podem medir a temperatura da pele. Diferentemente da temperatura central, a temperatura da pele varia frequentemente para regular e estabilizar a temperatura central. Mais recentemente, tornou-se comum o emprego de sensores de imagem e infravermelho para verificar rapidamente a temperatura corporal dos indivíduos de forma sem toque. Além disso, estudos e iniciativas estão constantemente desenvolvendo tecnologias vestíveis para monitorar a temperatura da pele. Em geral, tais estratégias aplicam adesivos impressos em nanotubos de carbono que fornecem detecção de temperatura mais precisa. No entanto, os sensores baseados em nanotubos de carbono requerem uma unidade de computação para adquirir dados dos sensores para disponibilizá-los para processamento. Isso impõe complexidade e uma barreira para o monitoramento de temperatura em tempo real. Como alternativa, alguns dispositivos portáteis são equipados com recursos de computação. Este é o caso da *Fitbit* que desenvolveu alguns dispositivos capazes de medir a temperatura basal da pele durante a noite. O indivíduo só deve usar seu dispositivo de pulso sem precisar de hardware adicional para coleta de dados.

A *frequência cardíaca* mede os batimentos por minuto da atividade cardíaca. Uma frequência cardíaca típica varia de 50 a 90 batimentos por minuto. Segundo Rosa Righi et al. (2023), o método tradicional para medir a frequência cardíaca é o eletrocardiograma. No entanto, esse método depende do paciente utilizar vários eletrodos presos ao peito e conectados a uma unidade central que processa os dados. Embora muito precisos, os eletrocardiogramas são impraticáveis para o monitoramento diário da frequência cardíaca. Assim, as iniciativas atuais exploram diferentes estratégias com fontes bioelétricas. As pulseiras inteligentes geralmente usam fotopletismografia para medir a frequência cardíaca. Quando o coração bate, os vasos capilares se expandem e se contraem com base nas mudanças de volume no sangue. A fotopletismografia é uma técnica óptica não invasiva capaz de medir as variações do volume sanguíneo na estrutura dos vasos capilares. O sensor utiliza um diodo emissor de luz (LED) que penetra na pele e percorre o tecido, sinalizando sua intensidade para um detector, por meio de reflexões.

A *variabilidade da frequência cardíaca* é o intervalo de tempo entre dois batimentos no ciclo cardíaco, e sua diminuição pode significar uma piorara no quadro clínico. De acordo com Rosa Righi et al. (2023), a variabilidade da frequência cardíaca varia entre as pessoas, e um valor alto representa uma resistência mais significativa ao estresse, enquanto um valor

baixo pode indicar doenças, estresse, depressão ou ansiedade. Valores reduzidos podem fornecer uma indicação precoce de que o indivíduo está sofrendo de algum tipo de infecção. As estratégias tradicionais medem a variabilidade da frequência cardíaca por meio do eletrocardiograma, sendo inviáveis para monitoramento diário. Pulseiras inteligentes são capazes de medir a frequência cardíaca, dessa forma, ao calcular a raiz quadrada média das diferenças sucessivas entre batimentos cardíacos normais, é possível determinar a variabilidade da frequência cardíaca. Portanto, os dispositivos que suportam frequência cardíaca podem fornecer também a variabilidade da frequência cardíaca.

Por fim, a *saturação de oxigênio* refere-se ao nível de oxigênio no sangue, que pode diminuir à medida que o pulmão é afetado e não pode realizar as trocas gasosas adequadamente. Em indivíduos com taxas abaixo de 95%, é um sinal de alerta e pode indicar falta de ar. De acordo com Rosa Righi et al. (2023), os sensores de saturação de oxigênio funcionam através da detecção da quantidade de oxigênio transportada pelas moléculas de hemoglobina através do sangue periférico. Em resumo, a saturação de oxigênio é a porcentagem de sangue oxigenado. Ao contrário de outros sinais vitais, os níveis de oxigênio geralmente permanecem constantes durante todas as atividades diárias, incluindo pratica de exercícios e o sono. Os métodos tradicionais para calcular a saturação de oxigênio dependem do uso de fotopletismografia compostos por sensores de luz vermelha e infravermelha aplicados nas extremidades do corpo. O método consiste em emitir sinais luminosos passando-os pelo sangue venoso. A maioria dos oxímetros disponíveis no mercado possui um design de clipe composto por duas partes: uma equipada com sensores para emitir luz e outra com um receptor e são colocados em algum dedo do paciente. Apesar de não invasivo, sua colocação pode causar algum desconforto ao indivíduo, pois pode interferir nas atividades das mãos. Neste contexto, pulseiras inteligentes aparecem como uma alternativa. Elas usam LEDs vermelhos e infravermelhos que, quando direcionados contra a pele, são refletidos em uma certa intensidade que estima a quantidade de oxigênio que o sangue possui. O sangue rico em oxigênio reflete mais luz vermelha do que a luz infravermelha. O oposto acontece com sangue mal oxigenado quando mais luz infravermelha é refletida. Dessa forma ao invés de colocar o dispositivo em um dedo, os pacientes podem apenar usar uma pulseira inteligente.

Ainda de acordo com Rodrigues et al. (2021), existem diversos dispositivos disponíveis para acompanhar a saúde das pessoas em tempo real. Embora a ideia pareça interessante, os fabricantes impõem diferenças entre o que pode ser obtido diretamente com os sensores integrados e o que pode ser obtido em seu aplicativo móvel. Assim, se um desenvolvedor planeja capturar dados de uma pulseira inteligente para um aplicativo de smartphone, estes sinais podem ser diferentes daqueles disponíveis no aplicativo do fabricante. Em sua pesquisa, que reuniu 148 dispositivos de 25 fabricantes diferente, Rodrigues et al. (2021) observou que apenas sete dos 148 dispositivos são capazes de monitorar os cinco parâmetros de sinais vitais em um único aparelho. A Tabela 1 apresenta os 7 dispositivos capazes de capturar os cinco sinais vitais já apresentados.

Tabela 1 – Lista de dispositivos com sensores capazes de capturar os sinais vitais apresentados.

Dispositivo	Modelo	Tecnologia de Comunicação
Embrace PLUS	Empatica	USB 2.0 e BLE
Fitbit Sense	Fitbit	Wi-Fi, BLE e NFC
Fitbit Versa 2	Fitbit	Wi-Fi, BLE e NFC
Fitbit Versa 3	Fitbit	Wi-Fi, BLE e NFC
Fitbit Charge 5	Fitbit	BLE e NFC
Fitbit Luxe	Fitbit	BLE
Fitbit Ionic	Fitbit	Wi-Fi, BLE e NFC

Fonte: Adaptada de Rodrigues et al. (2021).

Não ter todos os sensores em um único dispositivo requer o uso de mais de um dispositivo para capturar sinais de paciente, dessa forma dificultando o acesso e captura de sinais em tempo real, o que pode se tornar impraticável. Excluindo as principais empresas de tecnologia como Fitbit, Garmin e Samsung, a maioria dos dispositivos não fornece nenhuma maneira prática para obtenção dos dados em aplicações externas, só permitindo a visualização dos dados por meio dos aplicativos móveis proprietários das empresas.

2.2 Cidades Inteligentes

Como resultado do desenvolvimento de novas tecnologias, como a supracitada IoT, surgiram as cidades inteligentes (smart cities) (NASSEREDDINE; KHANG, 2024). Segundo Ziosi et al. (2022), as cidades abrigam a maior parte da população mundial e são os locais onde ocorre a maior parte do consumo e da produção. Em termos de números, cerca de 55% da população global residia em cidades em 2020, atingindo um pico de 85% em países como Austrália, Reino Unido e Holanda (ZIOSI et al., 2022). Nesse contexto surge o conceito de cidades inteligentes com a ideia de usar inovações tecnológicas para enfrentar grandes desafios urbanos e sociais (ZIOSI et al., 2022). Não existe uma definição única de uma cidade inteligente, e o termo tem significados diferentes para diferentes grupos de pessoas e varia de cidade para cidade e de estado para estado dependendo do nível de desenvolvimento, prontidão para mudanças e reformas, recursos e objetivos da cidade e população (NEVISTIC; BACIC, 2022). A União Internacional de Telecomunicações (ITU) analisou mais de cem conceitos diferentes para cidades inteligentes, resultando na seguinte definição:

> Uma cidade inteligente sustentável é uma cidade inovadora que utiliza tecnologias de informação e comunicação e outros meios para melhorar a qualidade de vida, a eficiência da operação dos serviços urbanos e a competitividade, garantindo ao mesmo tempo atender às necessidades das gerações presentes e futuras em aspectos econômicos, sociais e ambientais (INTERNATIONAL TELECOMMUNICATION UNION (ITU-T), 2014).

As cidades inteligentes usam novas tecnologias para várias aplicações, como transporte (ELASSY et al., 2024), coleta de lixo (RAI et al., 2024), reparos de ruas, eficiência administrativa, vigilância e muito mais. Nesse sentido, pesquisadores consideram que as cidades inteligentes têm um grande potencial para melhorar a eficiência em comparação com as cidades tradicionais (ZIOSI et al., 2022). De acordo com Syed et al. (2021), uma cidade inteligente é composta por vários componentes que são ilustrados na Figura 4. Os serviços inteligentes (*Smart City Services*) abrangem as atividades que sustentam a população de uma cidade, envolvendo tarefas municipais, como abastecimento de água, gestão de resíduos, controle e monitoramento ambiental, etc. Casas inteligentes (*Smart Homes*) envolvem o uso de sensores instalados pelas moradias das pessoas que fornecem informações sobre a moradia e seus ocupantes. A infraestrutura inteligente (*Smart Infrastructure*) ajuda a garantir que sua infraestrutura esteja em forma e utilizável, utilizando sensores para monitoramento da integridade estrutural das cidades. A agricultura inteligente (*Smart Agriculture*) é o uso de sensores embutidos em plantas e campos para medir vários parâmetros para ajudar na tomada de decisões e prevenir/doenças, pragas e etc.

Figura 4 – Componentes de uma cidade inteligente.

Fonte: Adaptada de Syed et al. (2021).

A energia inteligente (*Smart Energy*) leva em consideração o uso das tecnologias das cidades inteligentes para tornar as redes de geração de energia atuais e recém-instaladas mais observáveis, permitindo um melhor gerenciamento da geração e consumo de energia (SYED et al., 2021; AKRAM et al., 2024). A ideia de industria inteligente (*Smart Industry*) foca em tornar as industrias mais eficientes, aumentando a produtividade enquanto reduzem custos, através de manutenção preditiva, sensores para monitoramento/detecção de falhas e gerenciamento inteligente de produção e pessoas (SYED et al., 2021). Já a mobilidade inteligente (*Smart Transport*), foca em infraestruturas para fornecer serviços de transporte e logística para instituições e pessoas, considerando um sistema de transporte seguro, limpo, estável, integrado e

diversificado, reduzindo custos, tempo de transporte e poluição do ar (ATTARAN; KHEIBARI; BAHREPOUR, 2022). Por fim, a saúde inteligente (*Smart Health*) refere-se ao uso das tecnologias das cidades inteligentes para melhorar a disponibilidade e a qualidade dos cuidados de saúde da população (SYED et al., 2021; RODRIGUES et al., 2023).

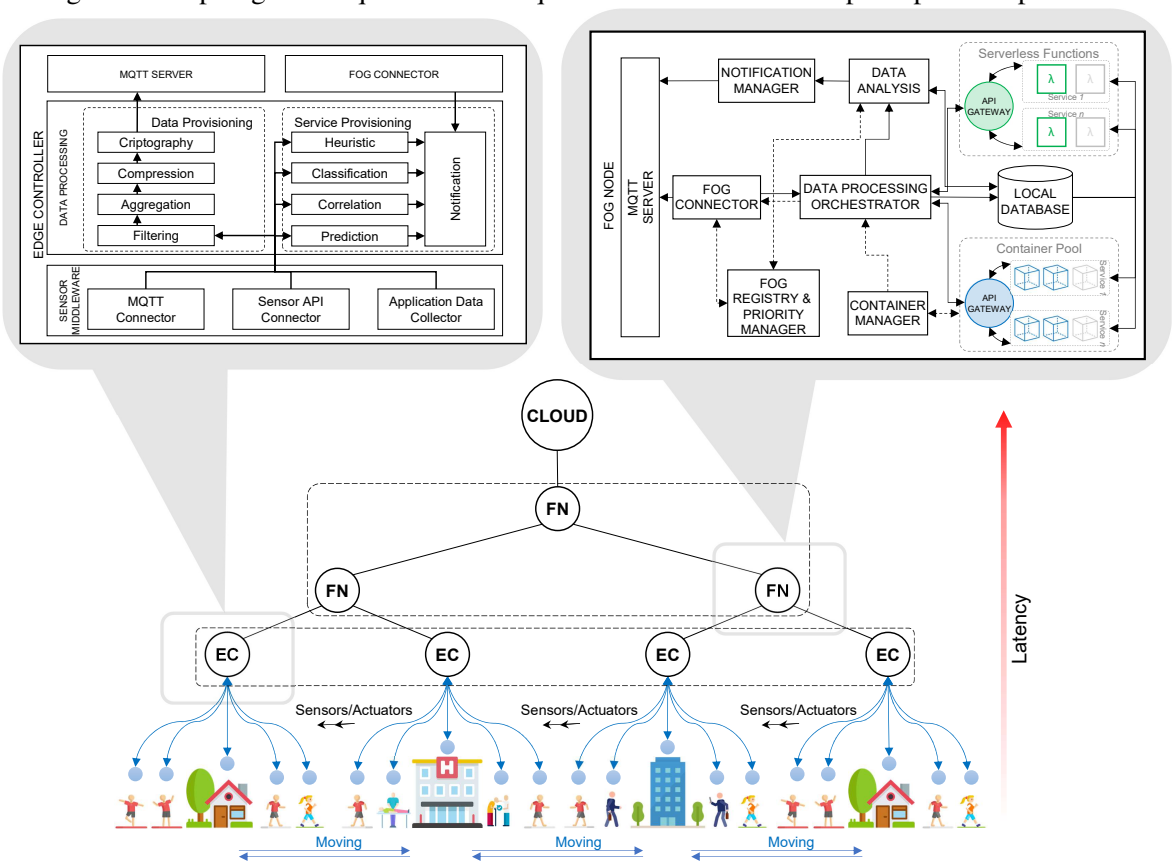
2.2.1 Modelo VitalSense

De acordo com Rodrigues et al. (2021), a falta de sensores acaba por impor desafios que exigem a combinação de dois ou mais sinais vitais para um diagnóstico. Neste contexto, é nítida a utilidade do uso dos dispositivos vestíveis com vários sensores de sinais vitais. Entretanto, esses dispositivos inteligentes para aplicações de saúde são relativamente novos. Estes dispositivos até já foram utilizados na área da saúde, sob autorização dos proprietários, mas as empresas ainda os projetam focando apenas em cuidados pessoais, *fitness* e esportes. No entanto, seu uso até o momento é focado apenas em estudos científicos, não possuindo dispositivos voltados para a área da saúde, com foco no monitoramento de doenças com o uso da detecção de sinais vitais.

Neste contexto, o MinhaHistoriaDigital (RODRIGUES; RIGHI, 2022) é um projeto de pesquisa brasileiro patrocinado pelo Programa Pesquisa para o SUS (PPSUS) e pela Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) que visou desenvolver uma solução para rastrear e capturar os sinais vitais de pessoas em tempo real. O projeto foi desenvolvido no âmbito da COVID-19, para oferecer uma estratégia para rastrear a COVID-19 e suas variações disseminadas, monitorando pacientes em suas casas e como eles respondem à infecção. Através do projeto foi desenvolvido o sistema VitalSense (RODRIGUES et al., 2023) que consiste na utilização de uma arquitetura *Edge-Fog computing* para processar os dados de saúde dos pacientes de acordo com sua localização. Dependendo da localização física do paciente, o sistema seleciona um nó *Fog* para processar seus dados que tenha o melhor tempo de resposta. Dessa forma o VitalSense visa fornecer um sistema de monitoramento de saúde em larga escala e em tempo real para uso em cidades inteligentes.

O VitalSense funciona através de algoritmos e uma arquitetura que emprega diferentes estratégias para melhorar a escalabilidade: (i) elasticidade de recursos; (ii) computação *serverless*; (iii) compressão de dados; e (iv) *sharding* para acesso eficiente aos dados. A arquitetura do modelo possui uma topologia hierárquica, onde os sensores dos pacientes na parte inferior fornecem os sinais vitais para o sistema. Empregando *Edge* e *Fog computing*, o VitalSense processa os dados localmente (Edge) ou remotamente (Fog). Essa estratégia visa encontrar o melhor tempo de resposta de acordo com a localização do paciente e a carga do sistema. Ambos os nós Edge e Fog fornecem serviços específicos para processamento dos dados e gerar alarmes de acordo com a saída dos serviços (RODRIGUES; RIGHI, 2022; RODRIGUES et al., 2023). A Figura 5 apresenta a arquitetura da solução detalhando os dois componentes principais: (i) Controlador *Edge* (EC); e (ii) Nó *Fog* (FN).

Figura 5 – Topologia da arquitetura hierárquica do VitalSense e seus principais componentes.



Fonte: Rodrigues e Righi (2022).

De acordo com Rodrigues e Righi (2022), os componentes do modelo se conectam em uma topologia hierárquica de acordo com sua localização física. Portanto, o VitalSense pode fornecer baixa latência e agregação de dados por região. Os pacientes utilizam sensores que fornecem seus sinais vitais para um controlador *Edge*. Por sua vez, os controladores *Edge* se conectam a um nó *Fog* específico de acordo com a sua localização física e latência para transmitir os dados processados. Os nós na *Fog* aplicam algoritmos de agregação e previsão de dados para produzir alarmes para situações críticas. Segundo Rosa Righi et al. (2023), por meio do emprego de *Fog* e *Cloud computing*, é possível fornecer diversos serviços de saúde aos pacientes sob demanda. Através do emprego de técnicas de probabilidade e estatística e métodos de aprendizado de máquina torna-se possível: (i) previsão de eventos; (ii) correlações causa-efeito ; (iii) classificação de dados; e (iv) reconhecimento de padrões.

2.2.2 Legislação Aplicável

Para implementar um modelo de gerenciamento de recursos humanos em cidades inteligentes, existem leis que devem ser cumpridas pelos ambientes de saúde quando o assunto é este. Conforme o Decreto-Lei nº 5.452 (BRASIL, 1943) e a Lei nº 13.467 (BRASIL, 2017), com relação a jornada de trabalho e movimentação de funcionários podemos elencar algumas dessas

disposições:

- Conforme Art. 59-A, é facultado às partes, mediante acordo individual escrito, convenção coletiva ou acordo coletivo de trabalho, estabelecer horário de trabalho de doze horas seguidas por trinta e seis horas ininterruptas de descanso, observados ou indenizados os intervalos para repouso e alimentação.

- Conforme Art. 59 e 61, a duração diária do trabalho poderá ser acrescida de horas extras, em número não excedente de duas, por acordo individual, convenção coletiva ou acordo coletivo de trabalho. Ocorrendo necessidade imperiosa, poderá a duração do trabalho exceder do limite legal ou convencionado.

- Conforme Art. 66, entre 2 (duas) jornadas de trabalho haverá um período mínimo de 11 (onze) horas consecutivas para descanso.

- Conforme Art. 67, Será assegurado a todo empregado um descanso semanal de 24 (vinte e quatro) horas consecutivas, o qual, salvo motivo de conveniência pública ou necessidade imperiosa do serviço, deverá coincidir com o domingo, no todo ou em parte.

- Conforme Art. 469, ao empregador é vedado transferir o empregado, sem a sua anuência, para localidade diversa da que resultar do contrato, não se considerando transferência a que não acarretar necessariamente a mudança do seu domicílio.

- Conforme Art. 2º, considera-se empregador a empresa, individual ou coletiva, que, assumindo os riscos da atividade econômica, admite, assalaria e dirige a prestação pessoal de serviço. Sempre que uma ou mais empresas, tendo, embora, cada uma delas, personalidade jurídica própria, estiverem sob a direção, controle ou administração de outra, ou ainda quando, mesmo guardando cada uma sua autonomia, integrem grupo econômico, serão responsáveis solidariamente pelas obrigações decorrentes da relação de emprego.

2.3 Processos Hospitalares

Para proposição de um modelo de gerenciamento de recursos humanos para a área da saúde, além do levantamento das tecnologias e conceitos necessários, também é preciso se atentar aos processos hospitalares envolvidos. Segundo Azevedo et al. (2010), hospitais podem ser vistos como portas de entrada para o sistema de saúde. Nesse sentido, um sistema de saúde poderia ser pensando como um circuito com vários pontos de entrada, onde existe um lugar adequado para se atender a necessidade de cada tipo de paciente. Assim, hospitais devem ser serviços de porta aberta para toda a população, não sendo aceitáveis qualquer medida que tente restringir o acesso a saúde. De acordo com Bittar (2000), quando se pensa em um hospital, podemos subdividi-lo em quatro grandes áreas: (i) Infra-Estrutura, (ii) Ambulatório e Emergência, (iii) Internação Clínica e Cirúrgica, e (iv) Diagnostico e Tratamento. Para cada uma dessa áreas, podemos

elencar diversas subáreas e diversos processos importantes para a realização do atendimento hospitalar.

2.3.1 Ambulatório e Emergência

Os processos relacionados à área Ambulatorial e de Emergência são extremamente importantes, por serem as alimentadoras da demanda que um hospital pode possuir. É uma área que pode contribuir não apenas com a qualidade do atendimento, mas também com a melhora da qualidade de vida do paciente, pois nesse processo situações que não necessitem de internação podem ser diagnosticadas, prevenidas e até tratadas. Possui como principais subáreas os ambulatórios clínico e cirúrgico, e os serviços de urgência e emergência (BITTAR, 2000).

O processo de triagem nos serviços do hospital deve ser bem organizado, não somente com relação a área física, material e equipamentos, mas também com pessoal competente e suficientemente treinado para analisar e definir a prioridade no atendimento de emergência a pacientes com enfermidades graves que correm risco de vida, ou que necessitam de assistência médica imediata (AZEVEDO et al., 2010). Segundo Coutinho, Cecílio e Mota (2012), os objetivos da triagem ou da classificação de risco são:

- Identificar os pacientes em situações de risco de vida;
- Determinar a área mais adequada para realizar o tratamento do paciente que se apresenta a um serviço de emergência;
- Reduzir o congestionamento do serviço de emergência, melhorando o fluxo de pacientes nas áreas de tratamento;
- Informar aos pacientes o tempo de espera estimado para atendimento, bem como o tipo adequado de serviço ao qual eles necessitam;
- Assegurar que pacientes que necessitem de uma prioridade de atendimento, sejam atendidos primeiro através da classificação;
- Contribuir com informações que ajudem a definir as causas do serviço, bem como sua complexidade;
- Garantir que os pacientes sejam reavaliados periodicamente;
- Priorizar o atendimento de pacientes sem realizar o diagnóstico.

O Sistema de Triagem de Manchester (GRUPO PORTUGUÊS DE TRIAGEM, 2010), conhecido pela sigla STM, é um sistema para gerenciamento de risco clínico, usado para gerenciar o fluxo de pacientes de forma segura, quando o sistema estiver com a sua capacidade de atendimento excedida. O STM possui cinco níveis de prioridade. Para cada nível é definido um

numero, cor, nome e tempo-alvo máximo aceitável para o atendimento médico. A Figura 6 traz uma representação visual dos níveis desse sistema.

Figura 6 – O Sistema de Triagem de Manchester.

Nível	Tipo de Urgência	Cor	Tempo de Espera
1	Emergente	Vermelho	Imediato
2	Muito Urgente	Laranja	10 Minutos
3	Urgente	Amarelo	60 Minutos
4	Pouco Urgente	Verde	120 Minutos
5	Não Urgente	Azul	240 Minutos

Fonte: Elaborada pelo autor, com base em Coutinho, Cecílio e Mota (2012).

Dessa forma, a finalidade dos serviços de emergência é oferecer atendimento imediato e de bom padrão aos pacientes, trazendo uma equipe de saúde treinada sobre as técnicas de atendimento, e capaz de atender os pacientes com foco na produção de saúde e qualidade de vida (AZEVEDO et al., 2010). Através do Sistema de Triagem de Manchester, enfermeiros são capazes de classificar os pacientes conforme o nível de urgência em que eles precisam ser atendidos. Assim, diminuindo o risco de erro por parte dos atendentes, priorizando o atendimento de pacientes em estado grave e melhorando o processo como um todo.

2.3.2 Exames de Diagnóstico e Tratamento

Os processos relacionados à área de Diagnóstico e Tratamento são aqueles relacionados ao atendimento realizado pelos médicos e enfermeiros a fim de identificar o problema de saúde que acomete o paciente. Além de identificar as patologias através dos mais variados tipos de exames, tem como objetivo definir qual o tratamento adequado ao paciente e, dependendo do caso, aplicá-lo.

Segundo Bittar (2000), nas últimas décadas a área de Exames de Diagnóstico e Tratamento foi uma área hospitalar que recebeu um desenvolvimento tecnológico enorme, contribuindo com a melhoria da qualidade dos diagnósticos e tratamentos. Em função desses avanços, houve um aumento significativo no número e na complexidade dos processos, exigindo uma tecnologia cada vez mais cara e recursos humanos altamente treinados. Diversos são os tipos de exames e processos relacionados a estes. Em um hospital geralmente se têm as seguintes subáreas: anatomia patológica, angiografia/hemodinâmica, banco de sangue, banco de tecidos, betaterapia, biologia molecular, cardiotocografia, dermatologia, diálise peritoneal, ecocardiografia, endoscopia, eletrocardiografia, eletroencefalografia, eletromiografia, ergometria, fisioterapia,

fonoaudiologia, hemodiálise, holter, litotripsia, medicina nuclear, microondoterapia prostática, neurofisiologia, oftalmologia, ortóptica, patologia clínica, pneumologia, potencial evocado, quimioterapia, radiologia, reprodução humana, radioterapia, reabilitação, ressonância magnética, tomografia, ultra-sonografia e urodinâmica.

2.3.3 Internação Clínica ou Cirúrgica

Os casos graves necessariamente serão levados a uma internação, com utilização de processos mais complexos do que exames e tratamentos simples, tendo procedimentos invasivos, como perfusões intravenosas, que consomem mais recursos do hospital e causam uma maior carga de trabalho, na qual o paciente permanece em atendimento por um tempo maior. Com isso, sendo sujeito a conflitos e possibilidades de observação do que ocorre no desenvolvimento de alguns processos. Em um hospital geralmente se têm as seguintes subáreas: anestesia, centro cirúrgico, centro obstétrico, centro de recuperação anestésica, internação clínico-cirúrgica, obstétrica e pediátrica, unidade neonatal, unidade de moléstia infecto-contagiosa, e serviço de controle da infecção hospitalar (BITTAR, 2000).

2.4 Técnicas de Predição

De acordo com Nisha e Sreekumar (2017), uma predição ou previsão computacional pode ser definida como uma declaração sobre um evento incerto que vai ocorrer no futuro. Para isso, é utilizado o conhecimento de uma série temporal do passado para prever o futuro. Uma previsão acurada ajuda a evitar possíveis riscos ou diminuir os riscos para problemas que ainda não ocorreram. A previsão de algum período específico pode ajudar a controlar as situações que fujam ao que se espera para determinado problema. Em séries temporais, a previsão parece significar estimar valores futuros através de valores passados de uma série. A predição computacional pode ser usada para planejamento estratégico (planejamentos de longo prazo), para finanças e contabilidade (orçamentos e controle de custos), na área de marketing (análise de vendas futuras e impacto de novos produtos) e também para as áreas de produção e operações.

A análise de séries temporais é considerada uma ferramenta extremamente poderosa para determinar modelos dinâmicos, visando definir e controlar variáveis de um sistema de forma mais apropriada. Um grande número de modelos de séries temporais está disponível para a previsão. Existem os modelos que utilizam da probabilidade e estatística, como o modelo *Auto-Regressive Integrated Moving Average* (ARIMA), e aqueles que utilizam de Inteligência Artificial, como os modelos que utilizam de Redes Neurais Artificiais e do algoritmo *Random Forest*. (NISHA; SREEKUMAR, 2017).

A previsão de séries temporais é uma área com bastante relevância, onde previsões de observações passadas são coletadas e analisadas, para desenvolver um modelo que descreve relações que não se situam claras e definidas, de forma que seja possível a predição de um determinado

evento no futuro (ROSSI, 2013). Nesse método de predição computacional, são utilizadas técnicas de probabilidade e estatística para realizar a inferência do futuro. Para isso, um modelo matemático é desenvolvido para representar o problema de previsão. Diversos são os modelos usando abordagens estatísticas, como o *Generalized Auto Regressive Conditional Heteroskedasticity* (GARCH), o *Naïve Bayes*, o *Wavelet analysis* e o *Support Vector Regression* (SVR), entretanto um dos mais utilizados é o modelo ARIMA (NISHA; SREEKUMAR, 2017).

2.4.1 Modelo Auto-Regressive Integrated Moving Average (ARIMA)

O modelo Auto-Regressivo Integrado de Médias Móveis, ou *Auto-Regressive Integrated Moving Average*, comumente conhecido apenas pela sigla ARIMA usa o conceito de atraso e deslocamento de informações passadas para prever padrões futuros (NISHA; SREEKUMAR, 2017). De acordo com Rossi (2013), esse tipo de modelo é muito popular por causa de suas propriedades estatísticas. Os modelos ARIMA possuem características flexíveis, que permitem ajustes e adaptações de seus parâmetros, fazendo com que diferentes tipos de séries temporais possam ser representadas através dele. Os modelos ARIMA também podem ser considerados bastante flexíveis, por terem a capacidade de representar séries temporais como o Auto-Regressivo puro (AR), Médias Móveis puro (MA), e a série combinada ARMA. Se o modelo for bem modelado, ele é mais do que capaz de fornecer bons resultados de previsões computacionais.

O modelo ARIMA é uma generalização dos modelos citados anteriormente por Rossi (2013):

- ARIMA (p, 0, 0) = AR (p);
- ARIMA (0, 0, q) = MA (q);
- ARIMA (p, 0, q) = ARMA (p, q).

De acordo com Neto (2011), a equação auto-regressiva de médias móveis, ARMA (p, q), pode ser descrita pela Equação 1:

$$y_t = \phi_1 y_{t-1} + ... + \phi_p y_{t-p} + u_t - \theta_1 y_{t-1} - ... - \theta_q u_{t-q} \tag{1}$$

onde ϕ e θ são os parâmetros do modelo proposto e u_t é o termo de erro do modelo. As variáveis y_t representam a parte Auto-Regressiva do modelo, conhecida como AR (p), enquanto que u_t representam a parte das Médias Móveis, conhecida como MA (q).

Conforme Neto (2011), na equação ARMA (p, q) explicada acima, ainda falta um termo de Integração *I* para que o modelo esteja de forma completa. Dessa forma, o modelo ARIMA (p,d,q) completo é obtido pela Equação 2:

$$y_t = \phi_1 y_{t-1} + ... + \phi_{p+d} y_{t-p-d} + u_t - \theta_1 y_{t-1} - ... - \theta_q u_{t-q} \tag{2}$$

Os modelos ARIMA são extremamente uteis na previsão de diferentes séries setoriais, pois podem representar séries estacionárias, através do uso do modelo ARIMA (p, 0, q), e também séries não-estacionárias, através do modelo ARIMA completo (p,d,q), sendo que o ajuste destes modelos sempre foi notável com relação ao desempenho para previsões (NETO, 2011). Uma das etapas mais importante do ciclo iterativo do método é a identificação do modelo ARIMA. Nessa fase o principal objetivo é encontrar os valores adequados para p, d e q e também determinar os parâmetros que vão ser úteis nas próximas etapas (ROSSI, 2013).

Depois de identificados os parâmetros, deve-se verificar se os mesmos foram bem representados. Uma das técnicas para fazer isso é o *superajustamento*, onde é estimado um modelo com parâmetros extras e examinado se estes são significativos. Logo após, verificar se a inclusão dele diminui significativamente a variância residual. Para isso é necessário analisar os resíduos do modelo ajustado. Então se a Equação 3:

$$\phi(B)W_t = \theta(B)a_t \tag{3}$$

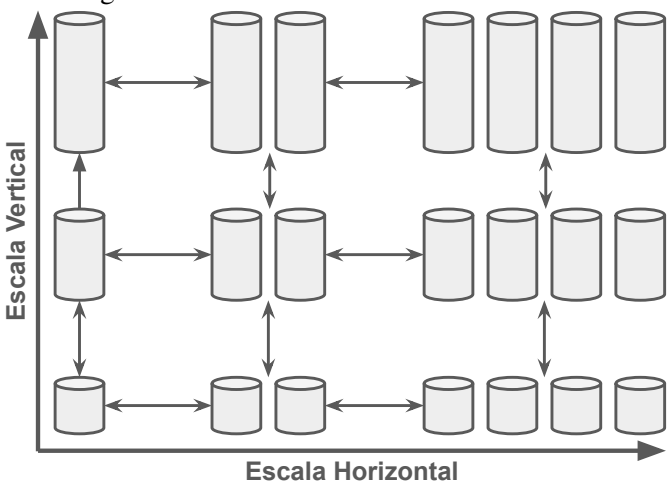
for verdadeira, com $W_t = \Delta^d Z_t$, então, os "erros verdadeiros" $a_t = \theta^{-1}(B)\phi(B)W_t$ serão ruído branco. Após os modelos já identificados, estimados e diagnosticados para que representem adequadamente os dados observados, eles estão prontos para realizar o processo de previsão conforme a necessidade (ROSSI, 2013).

2.5 Elasticidade de Recursos

Al-Dhuraibi et al. (2017) definem o conceito de elasticidade como a habilidade de um sistema ser capaz de adicionar ou remover recursos computacionais durante o uso e processamento dos dados, de forma a se adaptar a variação de carga de trabalho em tempo real. Estes recursos podem ser núcleos de CPU, memória RAM, GPU, instâncias de Máquinas Virtuais (conhecidas como *Virtual Machines* ou VM's), entre outros. Ainda de acordo com Al-Dhuraibi et al. (2017), elasticidade é uma propriedade dinâmica da Computação em Nuvem (conhecida como *Cloud computing*). Existem dois tipos de elasticidade na literatura: horizontal e vertical. A elasticidade vertical consiste na capacidade do sistema de aumentar ou diminuir as características dos recursos computacionais, tais como Tempo de CPU, núcleos, memória e rede. Já a elasticidade horizontal consiste na capacidade do sistema de adicionar ou remover instâncias dos recursos computacionais associados a aplicação que os usa. A Figura 7 ilustra ambos os conceitos, onde no eixo horizontal é possível ver o aumento da quantidade de nós computacionais, através da elasticidade horizontal, e no eixo vertical é possível ver o aumento de tamanho dos nós, representando o aumento dos recursos dos mesmos, através da elasticidade vertical.

Conforme a pesquisa de Rostirolla et al. (2018), o conceito de elasticidade pode ser estendido para o gerenciamento de consumo de energia em cidades inteligentes, de forma que os recursos de uma cidade podem ser ligados e desligados automaticamente, conforme a neces-

Figura 7 – Elasticidade Vertical e Horizontal.



Fonte: Elaborada pelo autor, com base em Al-Dhuraibi et al. (2017).

sidade dos usuários. Com base nisso, pode-se inferir que o conceito de elasticidade apesar de proposto para a alocação dinâmica de recursos na computação em nuvem pode ser estendido para outras áreas de aplicação. Para Al-Dhuraibi et al. (2017) e Rostirolla et al. (2018) escalabilidade é a habilidade de um sistema de sustentar crescentes cargas de trabalho, fazendo uso dos mais diversos recursos adicionais. Assim, para ter uma melhor compreensão e ilustrar o assunto, Al-Dhuraibi et al. (2017) resumiram o conceito de elasticidade na seguinte equação:

$$Elasticidade = \underbrace{\text{escalabilidade} + \text{automação}}_{\text{auto-escalabilidade}} + \text{otimização}$$

Com base nessa fórmula, Al-Dhuraibi et al. (2017) dizem que a elasticidade é construída em cima do conceito de escalabilidade, podendo ser considerada uma automação desse processo, no entanto, visa otimizar o mais rápido possível os recursos em um determinado momento de um sistema qualquer. Outro termo bastante associado com a elasticidade é a eficiência, que tenta definir de forma adequada como um recurso pode ser eficientemente utilizado junto ao processo de escalonamento de um sistema. Dessa forma, quanto menor a quantidade de recursos utilizados para realizar uma determinada tarefa em um tempo aceitável, melhor é a eficiência do sistema escalável.

A elasticidade idealmente deve ser invisível para o usuário do sistema, devido a dinamicidade do provisionamento dos recursos. A elasticidade é extremamente dependente do tempo do sistema. Portanto, a velocidade para realizar os ajustes é um ponto extremamente importante para a elasticidade. O tempo de reação é o intervalo de tempo entre o momento em que uma reconfiguração foi acionada ou solicitada, até que esta adaptação tenha sido concluída. Sistemas elásticos podem ser classificados de várias formas, dependendo de seus atributos, que podem ser: (i) Configuração, (ii) Escopo, (iii) Objetivo, (iv) Método, (v) Arquitetura, (vi) Provedor e (vii) Modo de Operação. (AL-DHURAIBI et al., 2017).

2.5.1 Configuração

Segundo Al-Dhuraibi et al. (2017), a configuração de um sistema elástico refere-se a alocação específica de CPU, memória, rede e armazenamento. Nesse contexto, a configuração representa a reserva inicial de recursos para o sistema. Durante a primeira aquisição do serviço, o usuário geralmente escolhe entre uma lista de pacotes de recursos, conforme a necessidade.

Figura 8 – Configuração da Elasticidade de Recursos.

Fonte: Elaborada pelo autor, com base em Al-Dhuraibi et al. (2017).

De acordo com Al-Dhuraibi et al. (2017) e conforme pode ser visto na Figura 8, a configuração pode ter dois modos de operação: (i) Rígido ou (ii) Configurável. No modo rígido, os recursos são oferecidos e provisionados em capacidades constantes. Nesse modo, as instâncias de máquinas virtuais possuem limites predefinidos de recursos (CPU, memória, entre outros). O problema com a configuração rígida é que os recursos alocados raramente atendem à demanda, portanto, há sempre um provisionamento inferior ou excessivo ao que realmente é necessário. Já no modo configurável, o cliente é capaz de escolher os recursos específicos de cada uma das máquinas virtuais. Para ambos os modos, os usuários podem reservar os recursos de acordo com os seguintes métodos de reserva:

- **Reserva sob demanda:** Os recursos são reservados imediatamente ou as solicitações são rejeitadas se não houver recursos disponíveis suficientes.
- **Reserva antecipada:** Os clientes enviam solicitações iniciais para reservar recursos e uma cobrança de preço fixo é necessária para iniciar a reserva. Os recursos devem estar disponíveis em um horário pré-especificado.
- **Reserva de melhor esforço:** As solicitações de reserva são enfileiradas e atendidas adequadamente conforme o esforço necessário.
- **Reserva baseada em leilão:** Configurações específicas de recursos são reservadas dinamicamente assim que seus preços são inferiores ao valor do lance oferecido pelo cliente.

2.5.2 Escopo

Para Al-Dhuraibi et al. (2017) e conforme pode ser visto na Figura 9, as ações de elasticidade podem ser aplicadas em dois níveis, de **infraestrutura** ou de **aplicação/plataforma**. As ações de elasticidade são responsáveis por executar as decisões tomadas pelo sistema de controle de elasticidade, de forma a dimensionar os recursos conforme a necessidade e a estratégia

estabelecida. As infraestruturas de nuvem são baseadas na tecnologia de virtualização, que podem ser máquinas virtuais ou contêineres. Além disso, com relação a aplicação/plataforma, as mesmas podem ser de uma ou várias camadas.

Figura 9 – Escopo da Elasticidade de Recursos.

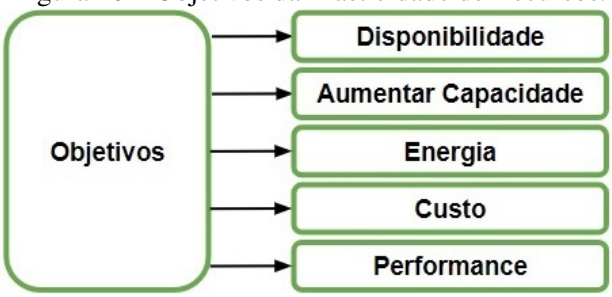
Fonte: Elaborada pelo autor, com base em Al-Dhuraibi et al. (2017).

A maioria das soluções de elasticidade existentes são dedicadas a aplicações de uma camada onde o gerenciamento de elasticidade é realizado apenas em uma camada, principalmente a de negócios. Quando o controle de ação de elasticidade está no nível de aplicação/plataforma, ele é denominado elasticidade embarcada. A maioria das soluções existentes é dedicada a aplicações baseados em servidor. No entanto, existem aplicações que têm particularidades de comportamento de execução diferentes, como aplicações científicas. Devido à natureza de uma aplicação científica, como paralelismo e modelos, uma solução de elasticidade não pode ser generalizada. A solução de elasticidade deve considerar a estrutura interna e o comportamento de uma aplicação científica, portanto, para ter uma solução elástica confiável, ela deve ser incorporada no código fonte da aplicação. Na elasticidade embarcada, as aplicações elásticas são capazes de ajustar seus próprios recursos de acordo com os requisitos de tempo de execução ou devido a mudanças no fluxo de execução (AL-DHURAIBI et al., 2017).

2.5.3 Objetivo

A elasticidade de recursos pode ter diferentes propósitos, como melhorar o desempenho, aumentar a capacidade de recursos, economizar energia, reduzir custos e garantir a disponibilidade. Com relação aos objetivos da elasticidade, há diferentes perspectivas: Os provedores da *Cloud* IaaS (*Infrastructure as a Service* ou Infraestrutura como Serviço) tentam maximizar o lucro minimizando os recursos enquanto oferecem uma boa Qualidade de Serviço (*Quality of Service* ou QoS), os provedores de PaaS (*Platform as a Service* ou Plataforma como Serviço) procuram minimizar o custo que pagam à nuvem e os clientes buscam aumentar sua Qualidade de Experiência (*Quality of Experience* ou QoE) e minimizar seus pagamentos. As soluções de elasticidade não podem cumprir os propósitos de elasticidade de diferentes perspectivas ao mesmo tempo, cada solução normalmente lida com uma perspectiva. No entanto, algumas soluções tentam encontrar uma maneira ideal de equilibrar alguns dos objetivos contraditórios (AL-DHURAIBI et al., 2017). A Figura 10 apresenta todos os objetivos encontrados para sistemas de elasticidade de recursos por Al-Dhuraibi et al. (2017) em sua pesquisa.

Figura 10 – Objetivos da Elasticidade de Recursos.

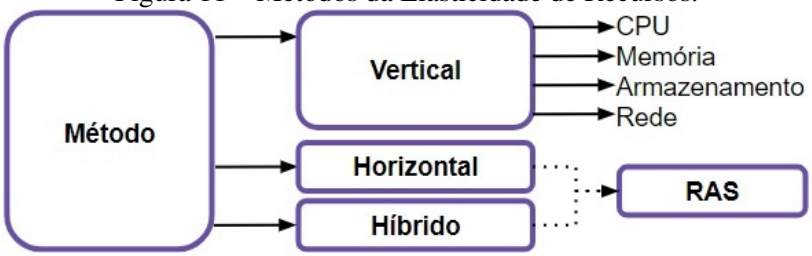


Fonte: Elaborada pelo autor, com base em Al-Dhuraibi et al. (2017).

2.5.4 Método

Segundo Al-Dhuraibi et al. (2017), para implantar soluções de elasticidade, deve ser implementado um dos seguintes métodos: escalabilidade horizontal, escalabilidade vertical ou escalabilidade híbrida. Técnicas de escalabilidade horizontais e verticais têm suas vantagens e desvantagens. A elasticidade horizontal é fácil de implementar e é amplamente utilizada no meio comercial. No entanto, ela pode levar a uma utilização ineficiente dos recursos, devido ao fato de fornecer instâncias fixas ou estáticas, que na maioria das vezes não são capazes de se ajustar exatamente a demanda requerida pela aplicação. Já a elasticidade vertical permite redimensionar as instâncias, entretanto não é amplamente utilizada, tendo poucos sistemas comercias que a suportem. Como dito anteriormente, na escalabilidade horizontal, instâncias são adicionadas e removidas conforme técnicas de balanceador de carga e na escalabilidade vertical, os recursos são redimensionados em tempo de execução como pode ser visto na Figura 11.

Figura 11 – Métodos da Elasticidade de Recursos.



Fonte: Elaborada pelo autor, com base em Al-Dhuraibi et al. (2017).

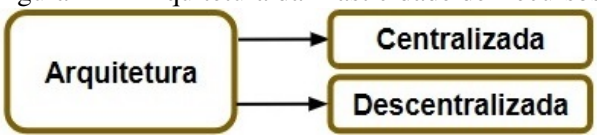
A migração também pode ser considerada para permitir o redimensionamento vertical quando não há recursos suficientes disponíveis, ou quando é necessário transferir uma instância para uma máquina física com menos recursos de forma a melhorar o desempenho. Antes de executar o processo de migração ou replicação de VMs, uma Estratégia de Alocação de Recursos (Resource Allocation Strategy ou RAS) é utilizada para decidir onde a nova instância será alocada ou criada, em qual servidor da nuvem. Existem ainda sistemas que utilizam uma estrutura que combinam o redimensionamento vertical, adicionado e removendo recursos de VMs existentes, e redimensionamento horizontal, adicionando novas VMs conforme a necessidade

(AL-DHURAIBI et al., 2017).

2.5.5 Arquitetura

Para Al-Dhuraibi et al. (2017) e como pode ser visto na Figura 12, a arquitetura das soluções para gerenciamento de elasticidade de recursos pode ser centralizada ou descentralizada. Na arquitetura centralizada existe apenas um controlador de elasticidade, responsável por realizar o redimensionamento automático, provisionando e desprovisionando recursos. Já na arquitetura descentralizada, existem vários controladores de elasticidade, responsáveis pelo provisionamento de recursos em diferentes plataformas em nuvem. Nesse modelo existe ainda um árbitro, considerado um ponto chave na arquitetura descentralizada, responsável por alocar recursos para os controladores nos diferentes componentes do sistema.

Figura 12 – Arquitetura da Elasticidade de Recursos.

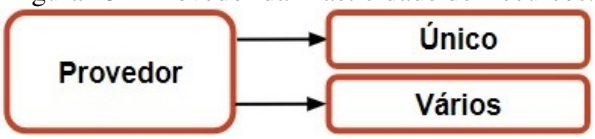


Fonte: Elaborada pelo autor, com base em Al-Dhuraibi et al. (2017).

2.5.6 Provedor

Conforme Al-Dhuraibi et al. (2017) e como pode ser visto na Figura 13, as soluções de elasticidade de recursos podem ser aplicadas a um ou a vários provedores da nuvem. Um único provedor de nuvem pode ser público ou privado, estando presente fisicamente em uma ou várias regiões ou centros de dados. Nesse contexto, várias nuvens significam que existem mais de um provedor de nuvem. Provedores de nuvem podem incluir nuvens híbridas que podem ser privadas ou públicas. A maioria das soluções e propostas para elasticidade de recursos suporta apenas um único provedor de nuvem.

Figura 13 – Provedor da Elasticidade de Recursos.



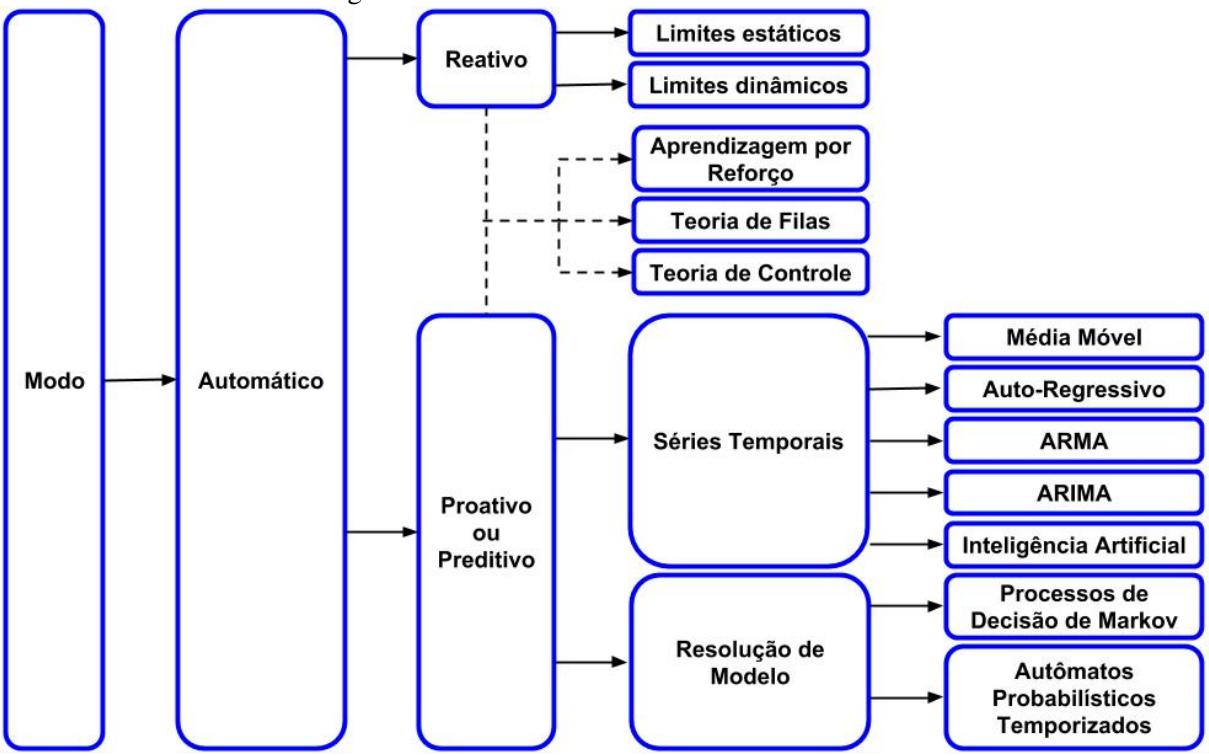
Fonte: Elaborada pelo autor, com base em Al-Dhuraibi et al. (2017).

2.5.7 Modo de Operação

Modo de operação refere-se às interações necessárias para realizar as ações de elasticidade no sistema. Normalmente, ações de elasticidade são executadas de forma automática. Ações de

escalabilidade podem ser alcançadas por intervenção manual do usuário, através de um modo manual ou programável, onde as ações de elasticidade são geralmente executadas através de uma Interface de Programação de Aplicações (Application Programming Interface, comumente conhecida como API). A política manual é usada em alguns sistemas comerciais de nuvem onde o usuário é responsável por realizar o monitoramento do ambiente virtual e por executar todas as ações de redimensionamento. Este modo apesar de ser atrelado ao conceito de escalabilidade, não pode ser considerado como um modo de elasticidade, uma vez que viola o conceito de automação, necessário para o sistema ser considerado elástico. Dessa forma, um sistema elástico possui apenas um modo, o modo automático, onde todas as ações de redimensionamento são realizadas automaticamente, podendo ser classificado em dois sub-modos: (i) Reativo e (ii) Proativo ou Preditivo, conforme pode ser visto na Figura 14 (AL-DHURAIBI et al., 2017).

Figura 14 – Modos da Elasticidade de Recursos.



Fonte: Elaborada pelo autor, com base em Al-Dhuraibi et al. (2017).

Na elasticidade reativa, as ações de elasticidade são disparadas conforme regras ou *thresholds* predefinidos, fazendo com que o sistema reaja disparando ações para adaptar as alterações no sistema de acordo com a carga (carga de trabalho ou utilização de recursos). Nesse modo, existem duas abordagens principais, baseadas em (i) Limites Estáticos e em (ii) Limites Dinâmicos. (AL-DHURAIBI et al., 2017).

Nos limites estáticos as ações de elasticidade são disparadas para aumentar ou diminuir recursos quando uma determinada condição é atendida. Essa política depende de limites ou requisitos predeterminados, onde as condições são baseadas nas medidas de um ou de um conjunto de métricas de recursos, como uso de CPU, utilização de memória, tempo de resposta

entre outros. Geralmente dois ou mais limites são usados para cada métrica, assim os valores medidos são comparados com as métricas para realizar o processo de escalabilidade. Um exemplo poderia ser: se a utilização da CPU for maior que 80% e essa situação durar por 5 minutos, o recurso deverá ser redimensionado. Já nos limites dinâmicos, ocorre o contrário, os limites não são estáticos e sim adaptativos, isto é, eles mudam dinamicamente de acordo com o estado das aplicações que os hospedam (AL-DHURAIBI et al., 2017).

De acordo com Al-Dhuraibi et al. (2017); Ahamed et al. (2023), na elasticidade proativa ou preditiva, as ações de elasticidade são disparadas com base em técnicas de previsão, antecipando as necessidades futuras da aplicação e acionando ações de elasticidade com base nessa antecipação prevista. Nesse modo, existem duas abordagens principais, baseadas em (i) Análise de Séries Temporais e (ii) Mecanismos de Resolução de Modelos. A análise de séries temporais, é usada para identificar padrões de repetição na carga de trabalho do sistema ao longo do tempo e para tentar prever os valores futuros dessa carga de trabalho. Assim, essa abordagem é responsável por fazer uma estimativa do uso futuro de um determinado recurso e da carga de trabalho da aplicação e, com base nessa predição, o controlador de elasticidade executará as ações necessárias conforme seu método de redimensionamento. Geralmente, a análise de séries temporais possui dois objetivos principais: prever valores futuros da série temporal com base nas últimas observações e identificar os padrões repetidos, se encontrados, de forma a usá-los para prever valores futuros. Para alcançar este objetivo, existem várias técnicas, como Média Móvel, Auto-Regressivo, série combinada ARMA, modelo ARIMA (apresentados na Subseção 2.4.1) e Inteligência Artificial. Já os mecanismos de resolução de modelos são abordagens baseadas na verificação probabilística de modelos ou estruturas de modelagem matemática para estudar os diversos comportamentos do sistema e antecipar seus estados futuros, como Processos de Decisão de Markov e Autômatos Probabilísticos Temporizados.

2.6 Considerações Parciais

Este capítulo apresentou conceitos pertinentes para a compreensão do restante do trabalho. Primeiramente foram apresentados os conceitos relacionados a Internet das Coisas de Saúde, os quais são fundamentais para a coleta de dados necessária ao modelo. Após, foram apresentados os conceitos relacionados as cidades inteligentes, locais para onde este trabalho está sendo proposto. A seguir, foram apresentados os processos hospitalares, um item fundamental para entendimento de um modelo de aplicação para a área da saúde. Após, foram apresentadas técnicas computacionais para predição noções importantes para o modelo preditivo de demanda de pacientes proposto. Por fim, foram retratados os conceitos relacionados a elasticidade de recursos da computação em nuvem, fundamentais para o modelo de elasticidade de recursos humanos propostas pelo modelo. Estes conceitos compõem a base necessária para compreensão das tecnologias utilizadas e as decisões tomadas para o modelo proposto.

3 TRABALHOS RELACIONADOS

Este capítulo tem como objetivo apresentar os trabalhos relacionados ao modelo proposto. Com base no tema de pesquisa, foram analisados alguns trabalhos relacionados que tentam resolver problemas parecidos, ou que servem como base para possíveis tentativas de solução. Desa forma, as próximas Seções descrevem algumas abordagens que focam na falta de recursos humanos para atender a demanda de pacientes ou que utilizam Internet das Coisas para identificar o estado de saúde ou localização de pacientes, a fim de obter uma visão do atual estado da arte dessas tecnologias. Os artigos foram classificados em ordem crescente por ano de publicação e depois por ordem alfabética dos nomes dos autores. Por fim, as iniciativas serão comparadas e analisadas para identificar as atuais lacunas na literatura da área.

3.1 Metodologia de Pesquisa e Seleção dos Trabalhos

De acordo com os estudos de Roehrs et al. (2017), a definição da estratégia de pesquisa é um dos passos mais importantes de uma revisão da literatura. Nesse processo foram definidas as palavras-chaves de pesquisa e a definição do escopo de pesquisa. Essas palavras foram definidas de forma a se obter os resultados de pesquisa focados no tema deste trabalho. Portanto, a *String de Pesquisa* localizada na Caixa de Texto 1 foi definida para a busca e seleção dos artigos nas bibliotecas acadêmicas eletrônicas.

Caixa de Texto 1: String de Pesquisa

((health OR healthcare OR hospitalar OR clinical) AND ((Smart City OR Smart Cities OR Internet of Things OR IoT OR sensors) OR (human resource OR human resource management OR hr OR hr management OR nurse management OR resource)))

Fonte: Elaborada pelo autor.

De forma a se cobrir um grande número de estudos, foram selecionadas várias bibliotecas científicas eletrônicas para serem utilizadas como base para pesquisa dos trabalhos relacionados, sendo elas: *ACM Digital Library*, *Google Scholar*, *IEEE Xplore Digital Library*, *IET Digital Library*, *PubMed*, *Elsevier B. V. ScienceDirect* e *Wiley Online Library*. De acordo com Roehrs et al. (2017), estes portais são algumas das bibliotecas que cobrem os jornais e conferências mais relevantes da área da ciência da computação. O período de pesquisa abrangeu artigos publicados a partir de 2018. Após a realização da busca, de forma a reduzir e elencar apenas os trabalhos relevantes, optou-se por selecionar trabalhos que atendessem a um dos seguintes critérios:

- Critério 1: Trabalho aborda de alguma forma a falta de recursos humanos para atender a demanda de pacientes;
- Critério 2: Trabalho propõe estratégias para gerenciamento de recursos humanos de forma a atender a demanda de pacientes;

Na Tabela 2 são apresentados todos os Trabalhos Relacionados selecionados, trazendo: (i) Um identificador para cada um deles, (ii) os nomes dos autores em formato de referência bibliográfica, juntamente com o ano da publicação, e (iii) a *publisher* da publicação. Os trabalhos foram ordenados em ordem ascendente pelo ano de publicação e, após, por ordem alfabética dos nomes dos autores.

Tabela 2 – Trabalhos selecionados.

Identificador	**Autores (ano)**	**Publisher**
TR1	Graham et al. (2018)	IEEE
TR2	Franke et al. (2018)	ECSDEV
TR3	Oueida, Aloqaily e Ionescu (2019)	Springer
TR4	Fischer et al. (2020)	Elsevier
TR5	Apornak et al. (2021a)	Taylor & Francis
TR6	Apornak et al. (2021b)	Taylor & Francis
TR7	Hafezalkotob et al. (2022)	Elsevier
TR8	Liao, Chu e Ho (2022)	MDPI
TR9	Duma e Aringhieri (2023)	Elsevier
TR10	Lazebnik (2023)	Elsevier
TR11	Rasidi et al. (2024)	UniMAP

Fonte: Elaborada pelo autor.

3.2 Using Data Mining to Predict Hospital Admissions From the Emergency Department (GRAHAM et al., 2018)

Graham et al. (2018) estudaram a superlotação de emergências hospitalares e suas consequências negativas para os pacientes. Eles propuseram um sistema para prever a chegada de pacientes um departamento de emergência de um hospital através do uso de técnicas de mineração de dados. Três algoritmos foram utilizados para construir modelos preditivos: regressão logística, árvores de decisão e Gradient Boosted Machines (GBM). O estudo descobriu que o GBM superou as árvores de decisão e a regressão logística na previsão da chegada de pacientes. Segundo os autores, as informações do modelo podem ajudar os gestores hospitalares a planejar os recursos com antecedência, evitando gargalos no fluxo de pacientes. No entanto, embora o modelo possa ser usado para apoiar o planejamento e a tomada de decisões, os autores deixam claro que as decisões individuais de admissão ainda requerem o julgamento dos gestores.

3.3 Sustainable Approach for a Smart Human Resource Management in Healthcare (FRANKE et al., 2018)

Franke et al. (2018) focaram na insuficiência de pessoal nos hospitais e propuseram uma solução para gerenciar a disponibilidade de recursos humanos e as necessidades hospitalares. Os autores propuseram a criação de uma rede hospitalar regional onde diferentes clínicas de uma mesma área tenham um cadastro único para contratação e disponibilização de recursos hu-

manos. Este sistema emprega tecnologia de informação inteligente para contratar enfermeiros para necessidades específicas. Os hospitais podem solicitar pessoal adicional para preencher lacunas no horário de trabalho, conforme necessário. A abordagem proposta visa otimizar processos e aumentar a eficiência e a eficácia através da utilização de pessoal flutuante para atuar em diferentes locais, reduzindo assim a subutilização de recursos.

3.4 A Smart healthcare reward model for resource allocation in smart city (OUEIDA; ALOQAILY; IONESCU, 2019)

Oueida, Aloqaily e Ionescu (2019) focaram seus estudos no uso e gerenciamento de recursos de saúde em cidades uma vez que sistemas de saúde otimizados são particularmente críticos, devido às populações altamente concentradas nas cidades. Isso pode levar a infecções e surtos de doenças, que podem se tornar situações epidêmicas e exigir contenção, o que é muito caro. Dessa forma, os autores propuseram um modelo para fornecer melhor entrega e utilização de recursos. O modelo baseado em recompensa de qualidade foi desenvolvido para estudar e reagir aos fatores de satisfação dos sistemas de saúde e propõe um algoritmo baseado em otimização chamado Algoritmo de Recompensa Máxima, que aprimora o uso e a entrega de recursos de saúde. O algoritmo foi testado com vários experimentos e simulações e provou que pode fornecer confiabilidade, eficiência e alcançar uma melhoria de desempenho de 50,1% a 77,2%.

3.5 ElHealth: Using Internet of Things and Data Prediction for Elastic Management of Human Resources in Smart Hospitals (FISCHER et al., 2020)

Fischer et al. (2020) estudaram o problema da superlotação de filas de pacientes em ambientes de saúde, ocasionadas pela ineficiência na alocação dos profissionais de saúde. Nesse contexto, os autores propuseram um modelo com foco em IoT capaz de monitorar o uso dos ambientes de saúde pelos pacientes e adaptar a alocação de recursos humanos para melhor atender as necessidades dos pacientes. O modelo emprega técnicas de predição para antecipar quando a demanda de uma determinada sala excederá sua capacidade e propõe ações para alocar profissionais de saúde de acordo. Através da implementação, os autores concluíram que através do uso de IoT e predição é possível gerenciar o uso de recursos humanos em ambientes de saúde, analisando os dados capturados para identificar possíveis padrões, e antecipar situações que fujam aos interesses dos gestores para que elas sejam tratadas e resolvidas de antemão.

3.6 Optimizing human resource cost of an emergency hospital using multi-objective Bat algorithm (APORNAK et al., 2021a)

Apornak et al. (2021a) focaram seus estudos na otimização do uso de recursos humanos na área de saúde. Neste contexto, os autores propuseram um método de avaliação de custos

com foco na otimização dos recursos humanos em um departamento de emergência aplicando um método flexível baseado no algoritmo Bat. Através dos estudos, foi possível identificar a alocação ótima de recursos humanos no departamento de emergência estudado, com uma expectativa de redução de mais de 24,3% nos custos de recursos humanos para o departamento de emergência. Os autores salientam que a pesquisa foi concentrada no custo dos recursos humanos, e que a sua aplicação pode diminuir a mão de obra no hospital durante o processo de otimização. Nessas circunstâncias, os pacientes podem ser afetados pela falta de cuidadores, dependendo do tipo e grau de sua doença.

3.7 Human resources optimization in hospital emergency using the genetic algorithm approach (APORNAK et al., 2021b)

Assim como em seu trabalho anterior, Apornak et al. (2021b) focaram mais uma vez seus estudos na otimização do uso de recursos humanos na área de saúde. Neste novo trabalho, os autores aplicaram um algoritmo genético com base em 27 variáveis de decisão e uma função de aptidão que combinou os salários diretos anuais e os pacientes atendidos por um grupo composto pelo pessoal essencial em área de especialização em três turnos de trabalho. Aplicando o algoritmo genético através do simulador MATLAB, os autores foram capazes de encontrar a função de aptidão que representa a quantidade mínima de recursos para atender o número máximo de pacientes do hospital após apenas 500 gerações. A solução ótima de alocação de pessoas proposta possui uma função de aptidão cerca de 36% menor que o estado atual do hospital. Além disso, a solução ótima inclui a identificação das competências e habilidades para atender a demanda potencial para casos de emergência, sugerindo a substituição de determinados funcionários por aqueles mais qualificados para atendimento da demanda.

3.8 A cooperative robust human resource allocation problem for healthcare systems for disaster management (HAFEZALKOTOB et al., 2022)

De acordo com Hafezalkotob et al. (2022), após um evento catastrófico, os sistemas de saúde desempenham um papel fundamental para salvar vidas dos indivíduos durante o período de resposta ao desastre. Neste contexto, os autores desenvolveram uma abordagem cooperativa baseada na teoria dos jogos para aliviar os impactos negativos de tais eventos, minimizar os níveis normais de serviço hospitalar e reduzir despesas indesejadas. Os autores propuseram uma formulação linear robusta para aumentar o nível de serviço dos recursos limitados após um desastre, permitindo comportamentos colaborativos entre diferentes ambientes de saúde, incluindo transferência de recursos humanos, leitos e pacientes entre hospitais. Como resultado, os autores mostraram que os modelos lineares propostos garantem que os hospitais sempre tenham a tendência de participar de coalizões maiores e que os gestores podem considerar a colaboração de hospitais como uma solução para lidar com desastres.

3.9 An Analysis of Waiting Time for Emergency Treatment and Optimal Allocation of Nursing Manpower (LIAO; CHU; HO, 2022)

De acordo com Liao, Chu e Ho (2022), as pessoas normalmente procuram atendimento de emergência quando sofrem de uma doença aguda ou desconhecida, o que leva a um grande número de pacientes aguardando para atendimento médico em Taiwan. Assim, a má alocação de mão de obra de enfermagem e as cargas de trabalho excessivas dos prestadores de serviços de emergência tornaram-se problemas sérios para as instituições médicas de Taiwan. Neste contexto, os autores aplicaram um método baseado na teoria das filas para tratar a sobrecarga de trabalho para prestadores de serviços de emergência e a má alocação de mão de obra de enfermagem, visando fornecer subsídios para uma alocação mais flexível de recursos e sugestões para melhoria nos processos. Como resultado, os autores afirmaram que o estudo analisou as causas da sobrecarga do serviço de emergência e identificou soluções para melhorar a utilização da mão de obra de enfermagem. Dessa forma, um modelo de tempo de espera baseado na teoria das filas é capaz de analisar os tempos de espera dos pacientes e estimar alocação ótima da mão de obra de enfermagem necessária para atendimento.

3.10 Real-time resource allocation in the emergency department: A case study (DUMA; ARINGHIERI, 2023)

Conforme Duma e Aringhieri (2023), a sobrelotação é um fenômeno que afeta os serviços de urgência em todo o mundo e, para lidar com o problema, são necessárias estratégias para alocação de recursos em tempo real, uma vez que as chegadas e as atividades a realizar só são conhecidas ao longo do tempo. Assim, os autores propuseram um algoritmo de alocação online baseado em priorização, combinado a uma ferramenta de previsão, para melhorar o desempenho de uma emergência, aliviando a superlotação, levando em consideração o estado da emergência em tempo real e a previsão das próximas atividades fornecidas por um modelo *ad hoc* de mineração de processos. Os resultados mostraram que, levando em consideração as prováveis atividades subsequentes, as decisões de alocação de recursos para a execução das atividades podem causar melhorias no tempo de permanência na emergência e na utilização de recursos.

3.11 Data-driven hospitals staff and resources allocation using agent-based simulation and deep reinforcement learning (LAZEBNIK, 2023)

De acordo com Lazebnik (2023), a alocação de pessoal e recursos hospitalares é um desafio crítico nos sistemas de saúde, pois envolve o equilíbrio entre as exigências dos pacientes, a disponibilidade de recursos e a necessidade de fornecer cuidados de saúde de em ambientes com recursos limitados. As abordagens tradicionais baseiam-se no planeamento manual e em ajustes *ad hoc*, que podem ser morosos e geralmente conduzir a resultados abaixo do ideal. Neste contexto, os autores usaram uma simulação baseada em *deep reinforcement learning* e *fully connected neural networks* para resolver o problema de alocação de pessoal e recursos

hospitalares. Como resultado, os autores mostraram que o modelo proposto é capaz de alcançar melhores resultados em termos de sucesso do tratamento do paciente e custo-benefício em comparação com algoritmos anteriores de alocação de recursos. Estes resultados sugerem que a abordagem tem potencial para melhorar a eficiência e eficácia da alocação de pessoal e recursos hospitalares nos sistemas de saúde.

3.12 Estimation of New Resource Allocation in Hospital's (or Medical Care) Inpatient Department using Discrete Event Simulation (RASIDI et al., 2024)

De acordo com Rasidi et al. (2024), a superlotação, os longos tempos de espera e a escassez de enfermeiros e camas são problemas significativos em hospitais. Neste contexto, os autores apresentam um modelo de simulação para o setor de internação de um hospital público localizado em Kelantan, utilizando Simulação de Eventos Discretos para identificar gargalos no processo e o número ideal de recursos necessários para atender às demandas. Assim, um modelo de melhoria foi desenvolvido com base nos resultados da otimização e demonstrou melhorar significativamente o tempo de espera e a taxa de utilização de enfermeiros e leitos. Os autores deixam claro que os resultados do estudo podem ser usados para auxiliar na tomada de decisões e aumentar a eficiência e eficácia do departamento de internação.

3.13 Análise

Nesta revisão da literatura, buscou-se identificar uma amostra de estudos que permitisse obter uma visão do atual estado da arte das tecnologias relacionados ao uso de Internet das Coisas e cidades inteligentes para apoio na área da saúde. A Tabela 3 apresenta uma comparação entre os artigos coletados, mostrando algumas lacunas em cada uma das abordagens, relacionando algumas de suas principais características, mostradas a seguir:

- **Problema:** Refere-se ao problema estudado pelos autores;
- **Proposta:** Refere-se a proposta de solução ao problema apresentada;
- **Técnica:** Apresenta qual a técnica utilizada pelos autores;
- **Recursos Humanos:** Refere-se as propostas para tratar a falta de recursos humanos nos ambientes de saúde;
- **Métricas:** Referem-se as métricas utilizadas pelos autores para avaliação do trabalho.

Com base nos artigos selecionados, pode-se identificar que existem atualmente várias frentes e abordagens quando o objetivo é prever problemas relacionadas à área de saúde utilizando os conceitos de Internet das Coisas e gestão de recursos humanos. Através dos artigos pesquisados foi possível constatar que não só é possível utilizar a tecnologia para isso, como já

Tabela 3 – Comparação entre os trabalhos relacionados.

Trabalho	Problema	Proposta	Técnica	Recursos Humanos	Métricas
TR1 (2018)	Superlotação de emergências	Prever admissões em um hospital	Uso de mineração de dados	Não propõe soluções, apenas fornece dados para tomada de decisão	Acurácia
TR2 (2018)	Falta de recursos para o atendimento dos pacientes	Compartilhamento de funcionários entre hospitais	Uso de uma rede de hospitais com uma *pool* de RH	Propõe uma rede para remanejamento de RH entre hospitais	Não propôs métricas
TR3 (2019)	Insatisfação de médicos, pacientes e gestores	Identificar os níveis de satisfação e propor ajustes nos recursos	Modelo para cálculo de satisfação	Identifica a necessidade de adicionar mais RH	Tempo de espera e pessoas aguardando
TR4 (2020)	Falta de recursos para o atendimento dos pacientes	Alocação adaptativa de RH com base em previsões de demanda	Uso de sensores para identificar a localização dos pacientes	Propõe a alocação adaptativa de RH usando predição e elasticidade	Tempo de espera e quantidade de RH
TR5 (2021)	Otimização do custo de RH para atendimento	Identificar o uso mínimo de RH para atender a demanda	Uso do algoritmo Bat	Propõe o ajuste dos RH de forma a diminuir custos	Custo de RH
TR6 (2021)	Otimização do uso de RH para atendimento	Organizar a escala de RH	Uso de Algoritmos Genéticos	Propõe o ajuste dos turnos de trabalho dos RH	Função fitness
TR7 (2022)	Falta de recursos para atendimento em situações de desastres	Realocação de recursos de saúde	Uso de teoria dos jogos e formulação linear robusta	Propõe remanejamento de RH entre hospitais	Custo de RH e perda de pacientes
TR8 (2022)	Falta de enfermeiros para o atendimento de pacientes	Identificar quantos enfermeiros são necessários para o atendimento	Uso da teoria de filas	Identifica a quantidade necessária de enfermeiros	Tempo de espera
TR9 (2023)	Otimização da alocação de recursos	Alocação de recursos dinamicamente usando predição	Uso de Hybrid Activity Trees	Propõe uma alocação online de RH utilizando predição	Utilização de recursos e tempo de permanência
TR10 (2023)	Pessoal hospitalar e alocação de recursos	Alocação adaptativa de recursos com base na demanda	Uso de Deep Reinforcement Learning e Fully connected Neural Networks	Propõe uma alocação ótima de RH utilizando predição	Sucesso de tratamento e custo-eficiência
TR11 (2024)	Superlotação de internações e falta de recursos	Identificar quantos enfermeiros e camas são necessários para o atendimento	Uso de Simulação de Eventos Discretos	Identifica a quantidade necessária de enfermeiros	Utilização de recursos e Tempo de espera

Fonte: Elaborada pelo autor.

está sendo utilizado em diversas abordagens na comunidade científica. Quando leva-se em consideração a falta de recursos humanos para atender a demanda de pacientes, alguns artigos encontrados se concentram apenas em prever a demanda futura dos pacientes ou a quantidade futura de médicos disponíveis, não propondo soluções para o problema, deixando para outros a responsabilidade de procurar por soluções ao problema. Para aqueles que propõem soluções, as abordagens não são em *on-the-fly*, sendo necessária a intervenção de gestores na tomada de decisão. E para a única abordagem em *on-the-fly* identificada, a mesma não leva em conta o monitoramento das condições de saúde das pessoas, apenas leva em conta quando os pacientes

já estão no ambiente médico. Nesse contexto, podemos listar as seguintes lacunas na área:

- Nenhum dos modelos analisou o estado de saúde dos pacientes fora dos ambientes de saúde para prever o futuro *uso de recursos* em cidades inteligentes;
- Esses modelos carecem de *soluções* para resolver proativamente problemas de carga de recursos humanos *antes que eles aconteçam*;
- Embora existam modelos capazes de estimar a demanda futura de um ambiente hospitalar, nenhuma dessas abordagens propõe um sistema para ajustes de recursos humanos *on-the-fly* em múltiplos ambientes de saúde;
- Faltam propostas para cidades inteligentes na área da saúde que relacionem a entrada de pacientes nos ambientes de saúde com os *sinais vitais* da população das cidades. Embora existam várias abordagens para a captura de sinais vitais na literatura, e para gerenciamento de recursos humanos, não foram encontradas abordagens capazes de integrar essas informações em um modelo unificado de gestão de recursos humanos.

Com base na revisão da literatura, fica evidente que melhorar a eficiência no uso dos recursos humanos faz diferença na qualidade do atendimento ao paciente. O problema da insuficiência de recursos humanos para atendimento médico não é algo recente e pode-se notar que vários estudos mostram que no futuro ainda será um problema. Portanto, estratégias para adequar antecipadamente a capacidade de atendimento às necessidades dos ambientes de saúde são extremamente necessárias. O uso da Internet das Coisas juntamente com o conceito de cidades inteligentes e técnicas de previsão podem auxiliar na proposição de futuras soluções ou na automatização de processos na área da saúde. Entretanto, o potencial dessas tecnologias continua a ser largamente subutilizado, uma vez que é possível aproveitar as mesmas para propor soluções, otimizar e maximizar a utilização dos recursos humanos já existentes. É possível perceber que falta um modelo unificado para receber a demanda real de pacientes e distribuir entre vários ambientes de saúde, propondo o encaminhamento adequado das pessoas levando em conta a capacidade de atendimento destas unidades médicas.

3.14 Considerações Parciais

Este capítulo apresentou um levantamento de trabalhos relacionados a esta tese. Considerando os trabalhos encontrados, foi possível encontrar várias lacunas a serem exploradas. Dentre as frentes na literatura que abordam a falta de recursos humanos para atender a demanda de pacientes, identificou-se a oportunidade para um modelo multinível de elasticidade para gerenciamento de recursos humanos *on-the-fly* para múltiplos ambientes de saúde. Além disso, foi possível ver uma oportunidade para uso de uma abordagem capaz de identificar o estado de saúde de pessoas fora dos ambientes de saúde, e utilizar estes dados para saber quando os mesmos precisarão se deslocar para atendimento e preparar os ambientes de saúde de antemão *on-the-fly*.

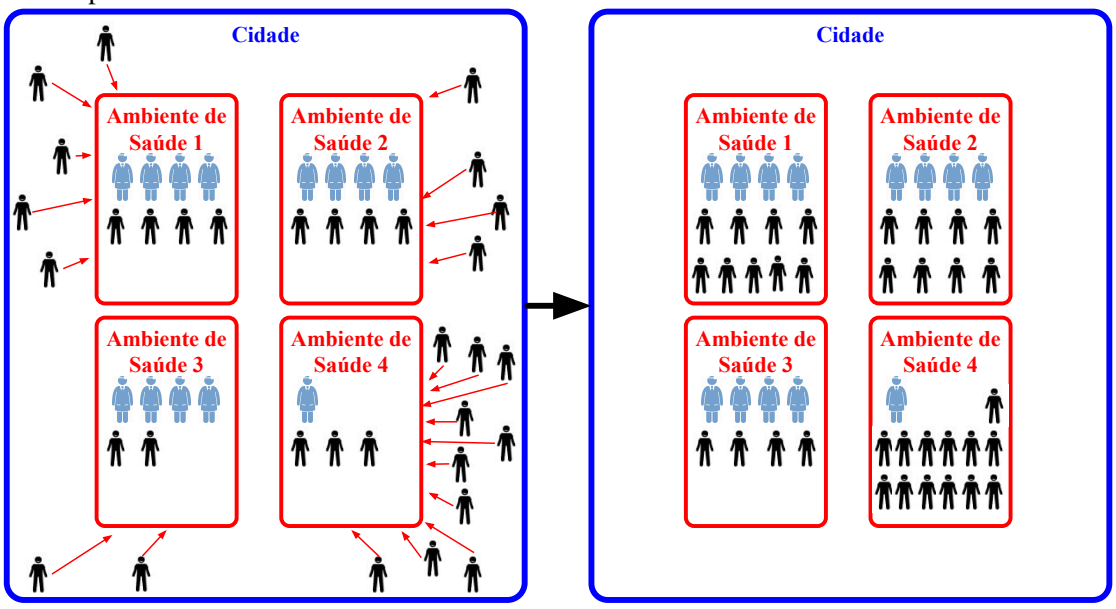
4 MODELO HEALCITY

Este capítulo tem como objetivo apresentar o modelo HealCity para gerenciamento de recursos humanos baseado no fluxo de pacientes dentro e fora dos ambientes de saúde de uma cidade inteligente. As próximas seções detalham o modelo proposto, apresentando as motivações e caso de uso do modelo (Seção 4.1), trazendo as principais decisões de projeto (Seção 4.2), a arquitetura proposta (Seção 4.3), o conceito de Elasticidade Multinível de Recursos Humanos em cidades inteligentes utilizando a abordagem proativa proposta (Seção 4.4), e, por fim, a abordagem proposta para gerenciamento das pessoas fora dos ambientes de saúde que necessitam de cuidados médicos (Seção 4.5).

4.1 Motivação e Caso de Uso

Com base no atual estado da arte e nas lacunas identificadas nos trabalhos relacionados, pode-se perceber que não apenas é possível realizar um controle do estado de saúde de pacientes em ambientes médicos, bem como nos mais variados ambientes de uma cidade inteligente. Através do uso da Internet das Coisas, é possível não apenas mensurar o estado de saúde das pessoas na cidade, como também identificar a localização desses cidadãos pessoas dentro ou fora de qualquer ambiente de uma cidade inteligente utilizando-se da tecnologia. Como visto anteriormente, um dos principais problemas dos ambientes médicos é a superlotação de pacientes que acaba por criar grandes filas para atendimento. E por mais simples ou complicada que o problema de saúde de um determinado paciente pode ser, se a equipe médica não reagir a tempo

Figura 15 – Cenário hipotético em que não existe qualquer estratégia de ajuste de recursos humanos ou indicação de ambiente de saúde adequado. Neste cenário, os pacientes tendem a se dirigir ao ambiente de saúde mais próximo e o número de recursos humanos em cada ambiente de saúde mantém-se constante.

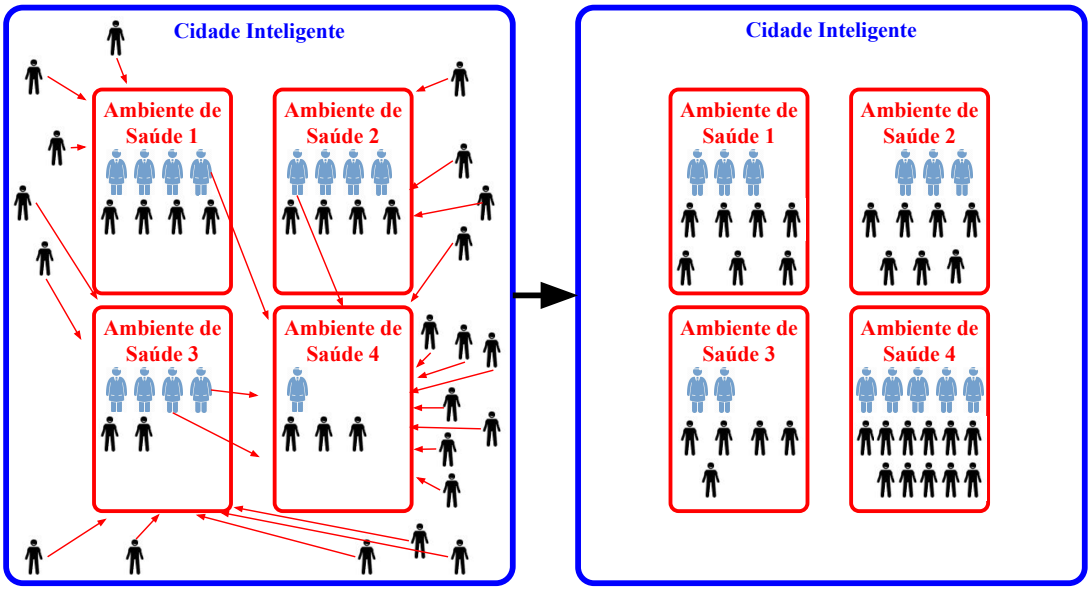


Fonte: Elaborada pelo autor.

não existem garantias sobre o futuro da saúde desse paciente. A revisão da literatura indica que a maioria das estratégias centra-se na otimização da utilização de recursos humanos, mas não avalia o estado de saúde dos pacientes fora dos ambientes de saúde ou a sobrelotação de pacientes em determinadas áreas, desconsiderando as potenciais vantagens que uma distribuição eficaz de recursos de saúde poderia proporcionar aos pacientes.

Dessa forma, pode-se imaginar um cenário hipotético, conforme expresso na Figura 15. onde existem quatro ambientes de saúde, onde três deles possuem três recursos humanos para atendimento e o último possui apenas um atendente. Além dos pacientes que já estão nos ambientes de saúde, existem mais pacientes chegando, aumentando a carga de trabalho de forma desproporcional, impactando o tempo de atendimento no ambiente mais despreparado para a demanda. Neste cenário hipotético, supõe-se que cada atendimento leva 10 minutos sem exceção, e cada atendente pode realizar um atendimento por vez. Dessa forma, o ambiente de saúde 1 atenderia todos os pacientes em 30 minutos, o ambiente de saúde 2 levaria 20 minutos, o ambiente de saúde 3 levaria 10 minutos e o ambiente de saúde 4 levaria 130 minutos. Assim uma abordagem capaz de ajustar a capacidade de atendimento à demanda de pacientes poderia melhorar consideravelmente o fluxo no ambiente de saúde mais impactado.

Figura 16 – Cenário hipotético em que existe uma estratégia de ajuste de recursos humanos e a indicação de ambiente de saúde adequado para um novo paciente. Neste cenário, os pacientes se dirigirem ao ambiente de saúde mais preparado para o seu atendimento e o número de recursos humanos em cada ambiente de saúde pode variar.

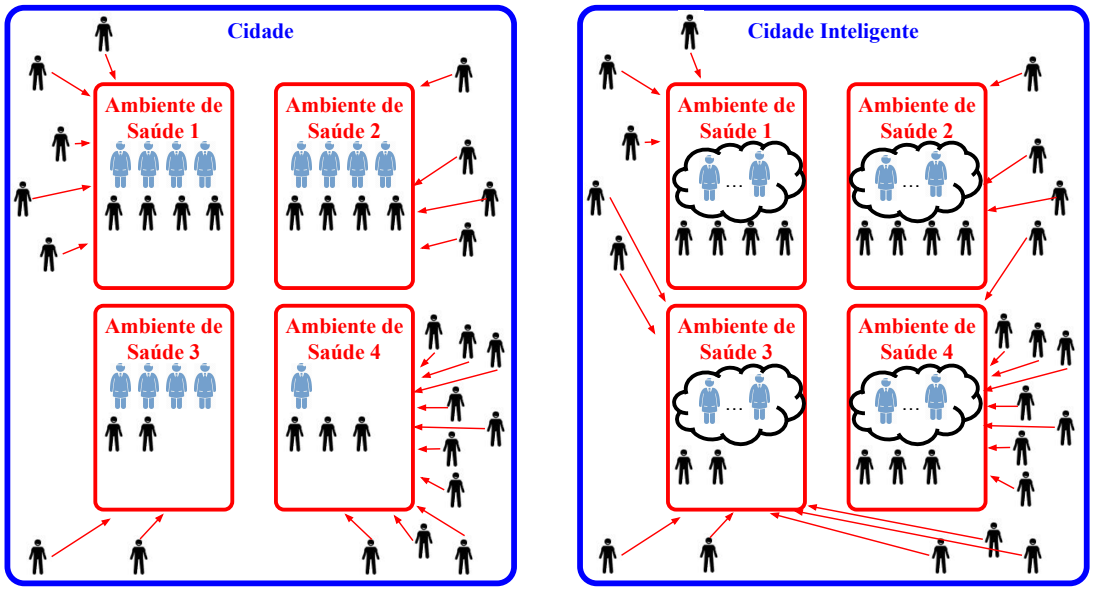


Fonte: Elaborada pelo autor.

Assim, se um modelo de gerenciamento de recursos humanos tivesse este cenário como entrada e tomasse algumas ações de movimentação de recursos humanos (mover um atendente dos ambientes de saúde 1 e 2 e mais dois atendentes do ambiente de saúde 3 para o ambiente de saúde 4), bem como indicar o melhor lugar para cada paciente ir, o resultado poderia ser bem diferente, conforme ilustrado na Figura 16. Dessa forma, mantendo os mesmos parâmetros do

cenário anterior, todos os ambientes de saúde levariam os mesmos 30 minutos para atender a demanda total de pacientes da cidade inteligente. Assim fica evidente que uma estratégia adequada para gerenciamento de recursos humanos em ambientes de saúde pode melhorar consideravelmente o tempo de espera por atendimento de saúde em um cenário multi-hospitais. Assim, para os cidadãos fica a impressão de uma *nuvem de recursos humanos*, onde a quantidade de atendentes é provisionada conforme a necessidade, como pode ser visto na Figura 17. Dessa forma é possível ver que o problema observado nos ambientes de saúde está em consonância com as lacunas científicas observadas no estado-da-arte, conforme apresentado no Capítulo 3.

Figura 17 – Comparação entre os cenários propostos, onde no segundo cenário, existe a impressão de uma *nuvem de recursos humanos*, onde a quantidade de atendentes é provisionada conforme a necessidade.



Fonte: Elaborada pelo autor.

Com base neste contexto, foi proposto o HealCity, um modelo multinível para gestão eficiente de recursos humanos baseado no fluxo de pacientes *dentro* e *fora* dos ambientes de saúde. Em particular, o HealCity adapta o conceito de elasticidade horizontal da *Cloud computing* ao contexto de gestão de pessoas, permitindo o ajuste da capacidade de atendimento do ambiente de saúde à demanda dos pacientes. Isso envolve alocação, desalocação e realocação dinâmicas de profissionais de saúde com base nos requisitos dos ambientes de saúde e nas necessidades das cidades inteligentes. HealCity consolida dados de várias fontes: chegadas e necessidades de pacientes (usando sensores espalhados pela cidade inteligente, sensores conectados às pessoas e um conjunto de dados de cidade inteligente), movimentação de pacientes (usando sensores) e disponibilidade da equipe médica (usando sensores e um conjunto de dados). Usando esses dados, é possível monitorar ativamente a demanda dos pacientes dentro e fora dos ambientes de saúde, facilitando a antecipação das demandas futuras dos pacientes. Assim, o HealCity calcula uma alocação otimizada de recursos humanos, o que contribui para reduzir o tempo de espera dos pacientes, um fator crítico em tempos de pandemia ou surtos virais.

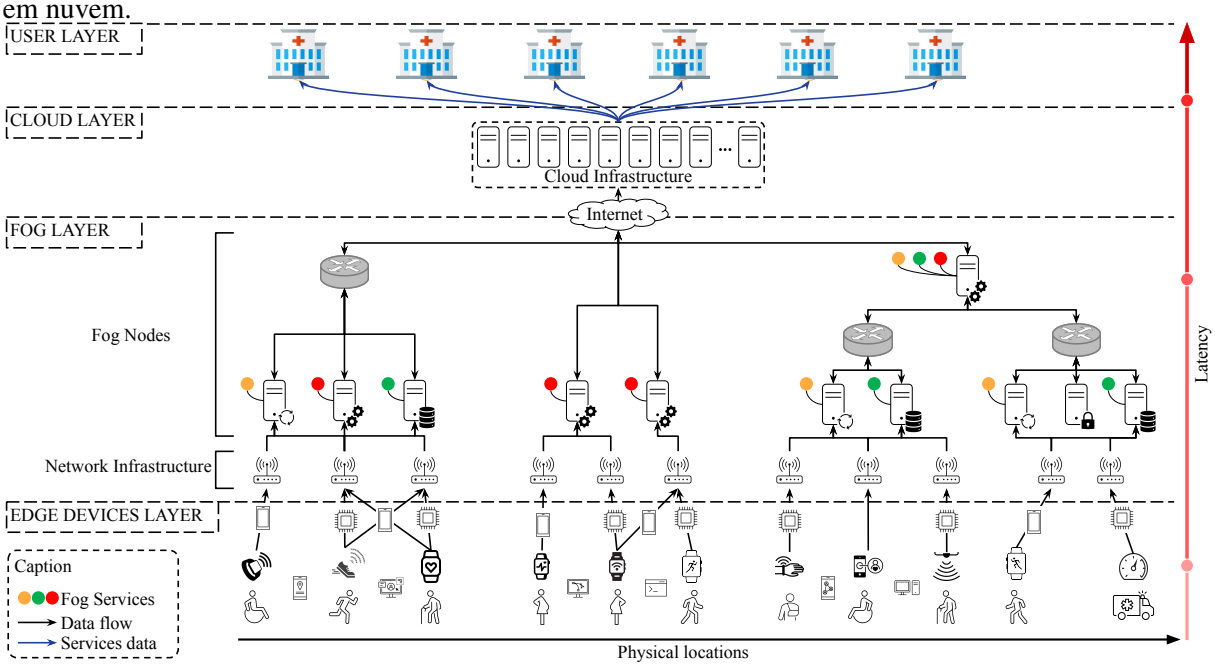
4.2 Decisões de Projeto

Conforme já relatado na Seção 2.2, a definição de uma cidade inteligente pode ser extremamente ampla. Neste contexto, o escopo para este modelo foca no quesito *Saúde Inteligente* e em *Casas Inteligentes* já apresentados anteriormente. Para o modelo HealCity, cidades inteligentes são as cidades que possuem uma infraestrutura de tecnologia da informação com serviços *Edge*, *Fog* e *Cloud* para processamento dos dados de saúde dos cidadãos. Neste contexto, na cidade inteligente proposta, todas as casas devem ser inteligentes no que tange a existência de controladores *Edge* para recepcionar dados capturados pelos sensores de saúde dos seus moradores. Além disso, lugares estratégicos da cidade também devem possuir esses mesmos controladores *Edge* para o monitoramento de saúde dos cidadãos não ser limitado às suas próprias casas. A infraestrutura em *Fog* deve ser regional para diminuir a latência na recepção dos dados dos controladores *Edge*.

O modelo parte da premissa que cada pessoa deve possuir um dispositivo vestível vinculado a si no sistema de forma a ser capaz de identificá-lo e deve vestir o mesmo durante todo o seu dia na cidade inteligente, dentro ou fora de ambientes de saúde. Esses dispositivos devem possuir sensores capazes de acompanhar o estado de saúde das pessoas em tempo real. Dessa forma, o sistema pode monitorar ininterruptamente o estado de saúde dos pacientes nas cidades inteligentes. Devem existir diversos leitores espalhados pelos ambientes da cidade capazes de receber dados dos sensores junto aos cidadãos. Estes leitores, devem ser capazes de ler dados em tempo real e devem comunicar-se com um controlador *Edge* próximo ao usuário, para diminuir a latência de comunicação. Conforme levantado na Subseção 2.1.3, esses sensores devem ser capazes de identificar variações em sinais vitais como: (i) frequência cardíaca, (ii) variação da frequência cardíaca, (iii) frequência respiratória, (iv) temperatura e (v) saturação de oxigênio (ROSA RIGHI et al., 2023). Além de capturar os sinais vitais das pessoas, esses leitores devem ser capazes de identificar a localização das pessoas que passarem por eles. Assim dados que indiquem alguma anomalia nesses sinais vitais são comprimidos e então encaminhados para a nuvem da cidade, responsável por analisar estes dados e identificar a necessidade do atendimento de saúde pelo paciente. Para ilustrar isso, a Figura 18 apresenta uma visão de uma estrutura hierárquica baseada em árvore de cidades inteligentes que tem como foco a coleta de sinais vitais e saúde móvel, onde os sinais vitais capturados em diversos ambientes são enviados aos ambientes de saúde após serem processados na infraestrutura em nuvem.

O funcionário responsável pela recepção de pacientes nos ambientes de saúde deve ser capaz de realizar o vínculo do sensor vestível de cada pessoa a um prontuário assim que o mesmo dar entrada no ambiente de saúde. Assim pode-se identificar quando e onde determinado paciente está logo que ele entra no ambiente médico, juntamente com o tempo que ele se mantém em cada uma das salas enquanto está sendo atendido, além de continuar medindo o estado de saúde do mesmo. Dessa forma, conforme explicado na Subseção 2.1.2 esse modelo propõe um RTLS com precisão de localização no nível de sala, dentro dos ambientes de saúde utilizando da tec-

Figura 18 – Estrutura hierárquica baseada em árvore da cidade inteligente com foco no monitoramento dos parâmetros de saúde dos pacientes. As pessoas usam sensores que transmitem sinais vitais para uma infraestrutura *Fog-Cloud* que fornece informações de saúde diretamente para os ambientes de saúde. Nesta estrutura, os cidadãos estão no nível mais baixo, interagindo com dispositivos de ponta, enquanto os ambientes de saúde estão no nível mais alto, interagindo com dados já processados pela infraestrutura em nuvem.



Fonte: Rosa Righi et al. (2023)

nologia RFID, e em ponto específico, nos demais ambientes da cidade utilizando da localização dos controladores Edge que realizam a captura dos sinais vitais. Assim é possível identificar em qual sala determinado paciente está dentro de um ambiente médico, ou em que ponto da cidade o mesmo está. Conforme descrito na Subseção 2.1.1, a escolha do RFID se baseia na sua capacidade de permitir identificação automática, evitando a existência do erro humano nos processos de identificação. Além disso, segundo Boulos e Berry (2012), para certas aplicações, o uso de soluções RTLS que incorporem tecnologias complementares como RFID, podem fornecer níveis de precisão e flexibilidade inatingíveis por qualquer tecnologia concorrente. O funcionamento se baseia na capacidade do sistema de marcar as entradas e saídas de cada paciente em cada uma das salas existentes no ambiente do ambiente de saúde e registrar as mesmas, gerando marcações fidedignas a movimentação das pessoas. Com base nisso, o sistema deve ser capaz de identificar o uso de cada um dos ambientes por parte dos pacientes durante seu processo de atendimento. Do ponto de vista dos usuários, o HealCity deve ser transparente ao paciente, no sentido que o mesmo não precisa informar nenhuma condição relacionada a sua movimentação pelos ambientes de saúde, ou sobre seus sinais vitais, sendo uma atividade realizada automaticamente pelo sistema.

Assim, o modelo proposto fornece uma infraestrutura de monitoramento abrangente para cidades inteligentes que pode ser usada para melhorar vários aspectos da saúde pública. O He-

alCity usa duas abordagens diferentes para adquirir informações do paciente. Na primeira, o modelo utiliza um RTLS dentro do próprio ambiente médico para identificar a localização e movimentação dos pacientes. Na outra abordagem, o modelo adquire dados através do serviço VitalSense (RODRIGUES et al., 2023) para cidades inteligentes, responsável por informar quem são as pessoas da cidade e sua localização, que têm sinais vitais alterados e, consequentemente, que podem vir a precisar de atendimento médico. Conforme já detalhado na Subseção 2.2.1, o modelo VitalSense é uma solução escalável baseada em *Fog* para monitoramento em tempo real e processamento de dados de pacientes usando dispositivos vestíveis (RODRIGUES et al., 2023). Além disso, os nós do VitalSense criam um par de chaves RSA 2048 para criptografia de dados na camada de *Fog*. Os dados são criptografados por controladores Edge usando a chave pública e descriptografados por nós *Fog* usando a chave privada. Além disso, o VitalSense utiliza algoritmos de criptografia homomórfica nesta camada para garantir a segurança de todos os dados em trânsito. Os dados brutos podem ser enviados para serviços externos de terceiros e nuvens públicas capazes de realizar operações aritméticas nos dados sem acesso direto. Essas entidades não podem visualizar os dados reais e devem solicitar a descriptografia de seus resultados computacionais de um nó *Fog*. Desta forma, o serviço prestado pela VitalSense é útil para este trabalho, pois os seus nós *Fog* aplicam algoritmos de agregação e previsão de dados para produzir alarmes para situações críticas de saúde e, simultaneamente, estabelecer um ambiente seguro para o processamento de dados dos pacientes.

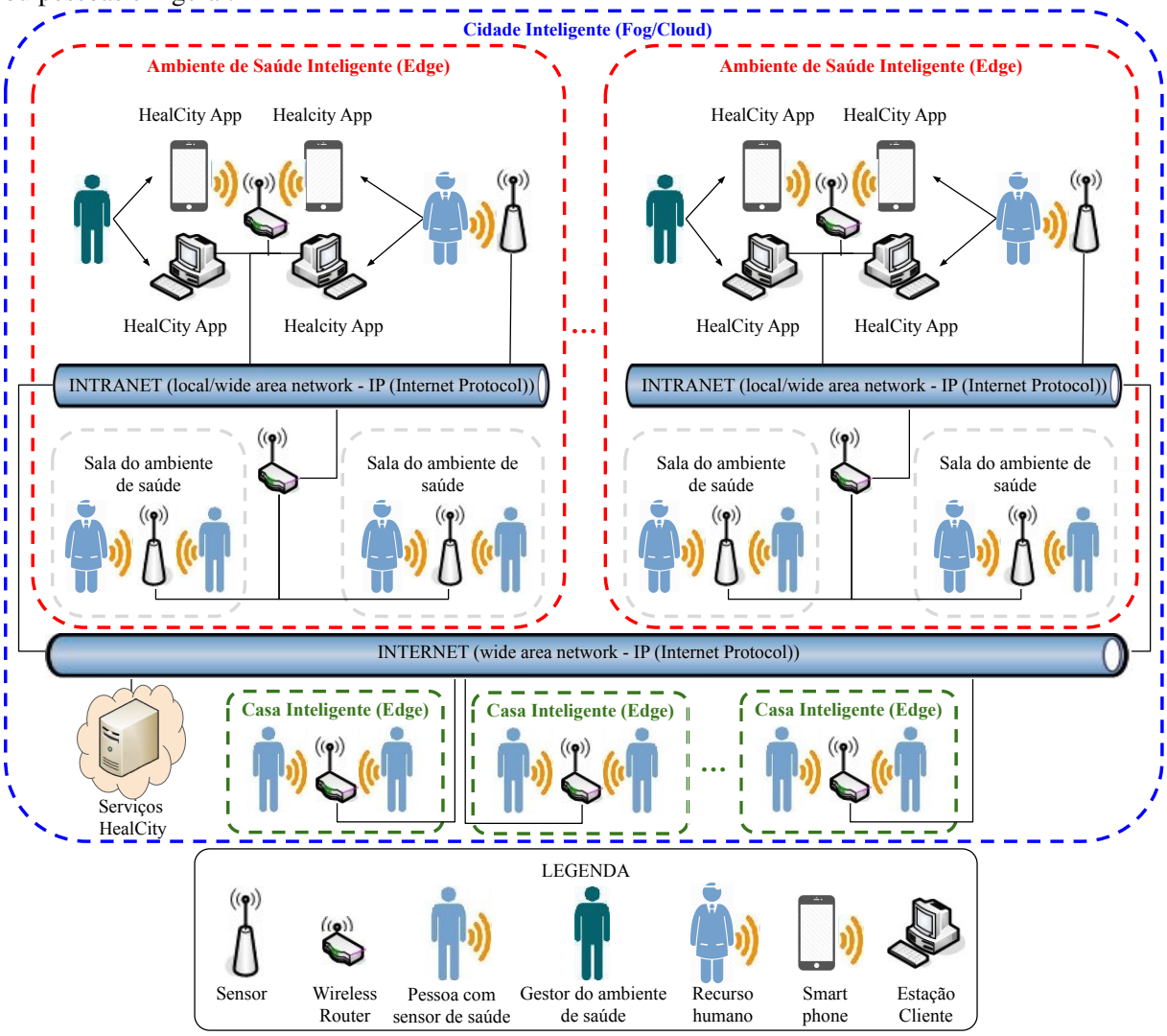
Com relação à estratégia de comunicação entre os ambientes de saúde, no modelo proposto os ambientes de saúde não se comunicam diretamente um com o outro, e sim através do serviço HealCity, responsável por agregar as informações fornecidas pelos diferentes ambientes de saúde, e fornecer os dados já tratados aos envolvidos. O detalhamento da arquitetura será mostrado na próxima seção. Com relação a estratégia computacional de predição, o modelo HealCity faz uso da abordagem baseada em probabilidade e estatística, através de uma implementação do modelo ARIMA, conforme detalhado na Subseção 2.4.1. Para a comunicação entre os leitores RTLS com a aplicação proposta, o modelo parte do uso do padrão EPCGlobal. Para isso, definiu-se que o *middleware* utilizado pela aplicação deve implementar os tipos de dados propostos no EPCGlobal. Optou-se por essa estratégia para permitir que o os dados capturados pelos leitores RTLS na aplicação possam ser utilizados por outras abordagens que façam uso do padrão.

4.3 Arquitetura

A arquitetura do modelo HealCity se baseia na existência de dois principais serviços: (i) *um serviço de tomada de decisão*, responsável pelas decisões de gerenciamento de recursos humanos, e (ii) *um serviço web*, responsável pela camada de visualização, através do aplicativo *HealCity App*. O primeiro serviço é responsável pelo processamento dos dados, tratando os registros de passagem, sinais vitais capturados e realizando as predições da demanda de paciente e definindo a alocação de recursos humanos para o atendimento. Já o segundo serviço serve como

interface de uso do sistema por parte dos gestores dos ambientes de saúde, recursos humanos, pacientes e pessoas em geral na cidade inteligente. A Figura 19 apresenta os componentes e a visão da rede no modelo proposto. Os leitores e sensores devem ficar instalados em todos os ambientes da cidade inteligente e estar ligados logicamente a *Fog*, podendo trocar informações com ela. Os dispositivos clientes conectados a essa rede devem apenas ter acesso a interface *HealCity App*, enquanto que o serviço que controla os dados recebidos deve ser executado em partes nos controladores Edge nos ambientes de saúde, em parte na *Fog*, e em parte na *Cloud*. O modelo parte do pressuposto que a interface será utilizada através de um aplicativo e deve funcionar em qualquer máquina cliente que tenha a permissão para acesso. Assim, sempre que um usuário carregar uma página do sistema, o servidor web interpreta as requisições, acionando a aplicação e coletando os dados a serem exibidos no banco de dados.

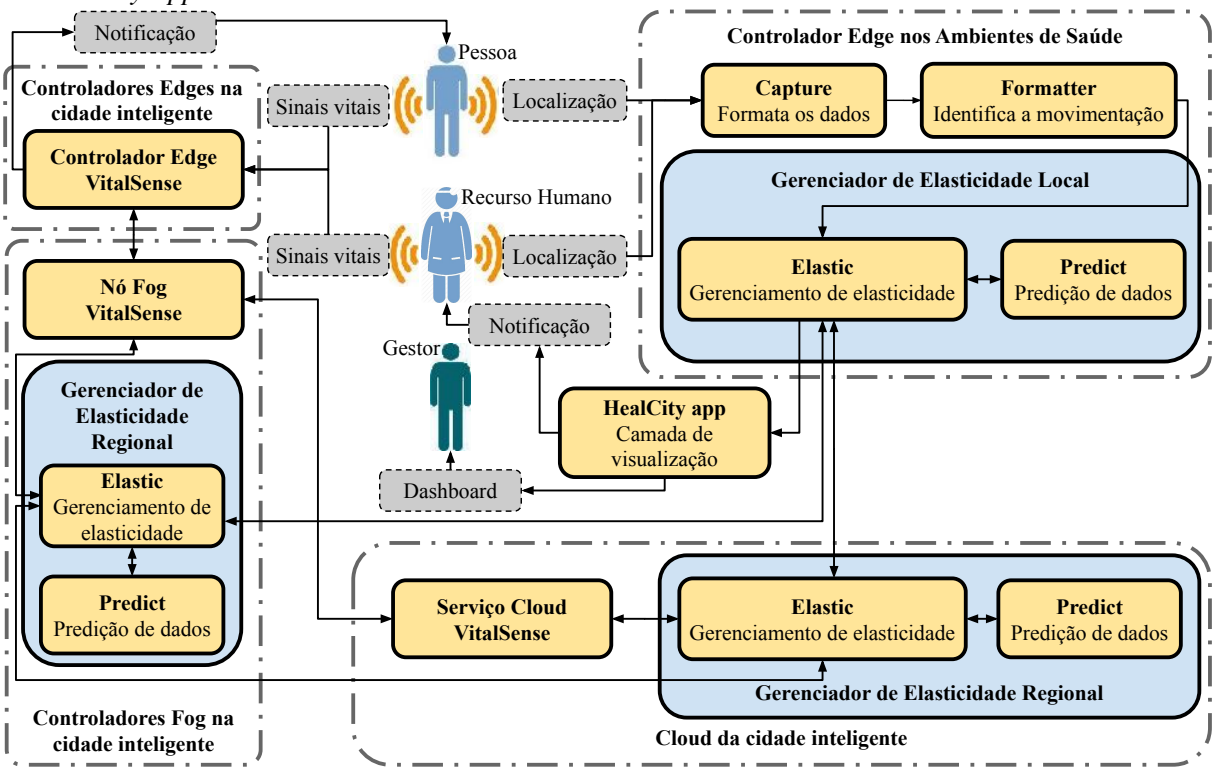
Figura 19 – Componentes e visão da rede no modelo HealCity com um (i) serviço web (HealCity App); (ii) serviço de tomada de decisão (Serviços HealCity); (iii) uma rede de sensores para captar os sinais vitais e a localização dos cidadãos; e (iv) gestores dos ambientes de saúde, recursos humanos, pacientes ou pessoas em geral.



Fonte: Elaborada pelo autor.

O modelo HealCity é baseado em uma estrutura modular responsável pelo tratamento da informação desde a aquisição inicial de dados pelos sensores até o resultado mostrado no *HealCity App*. Cada módulo possui uma função específica, possuindo uma informação de entrada e um resultado específico de saída que pode ser utilizado como entrada de outros módulos. A Figura 20 apresenta o detalhamento da arquitetura proposta, juntamente com os módulos propostos (*Capture, Formatter, Elastic, Predict e HealCity App*) e os tipos de dados e informações que são trocadas entre eles, trazendo uma visão gráfica desde as entradas do modelo até a saída do mesmo.

Figura 20 – Detalhamento da arquitetura e fluxo de dados do modelo HealCity, ilustrando a trajetória dos dados começando nas pessoas que fornecem sinais vitais e localização ao VitalSense e ao módulo *Capture*, respectivamente. Esses dados são posteriormente processados em vários módulos, culminando na exibição das notificações de elasticidade para os recursos humanos, ou dos *dashboards* aos gestores no *HealCity App*.



Fonte: Elaborada pelo autor.

Capture é o módulo responsável por recepcionar os dados capturados pelos sensores espalhados pelo ambiente de saúde e realizar o pré-processamento dos mesmos. Para tanto, esse módulo deve ser capaz de se comunicar com o *middleware* RTLS. Para garantir a inclusão das informações assim que as mesmas são capturas pelos sensores, a aplicação deve verificar constantemente com o *middleware* sobre a chegada de novos dados dos sensores. Nessa etapa, a partir dos identificadores da *tag* e do sensor, devem ser identificados o paciente e o sensor físico a qual pertencem. Com base nisso, deve ser criada uma marcação de passagem pelo sensor (*Marking*) e o resultado final deve ser armazenado na base de dados. Caso o sistema capture

muitas marcações de passagens de uma única vez, o módulo deve ser capaz de processá-los pouco-a-pouco, evitando sobrecargas na aplicação em função do volume de dados. O módulo em questão é executado no controlador Edge existente nos ambientes de saúde em função da latência.

Formatter é o módulo responsável por realizar o processamento dos dados do tipo *Marking*. Nesta etapa, devem ser identificadas as passagens dos pacientes pelas salas do ambiente de saúde. Cada par de marcações identifica a saída de um lugar e a entrada em outro, e um novo tipo de dados chamado *Pass* deve ser criado. Nele é armazenado a informação do lugar por onde o paciente passou, a data e hora de entrada e saída, o tempo de permanência neste lugar e o paciente que realizou essa passagem. Além disso, esse módulo deve ser capaz de armazenar a localização atual do paciente e o tempo de permanência. Em função do escopo do módulo e da necessidade de haver pares de marcações para identificar a movimentação dos pacientes, a sua execução ocorre sempre que são identificados novos pares de marcações, não sendo executado sobre marcações individuais. Visto a necessidade dos dados serem rapidamente processados pela aplicação, esse módulo se mantém em processo de busca constante por novas marcações, sendo executado no controlador Edge existente nos ambientes de saúde em função da latência.

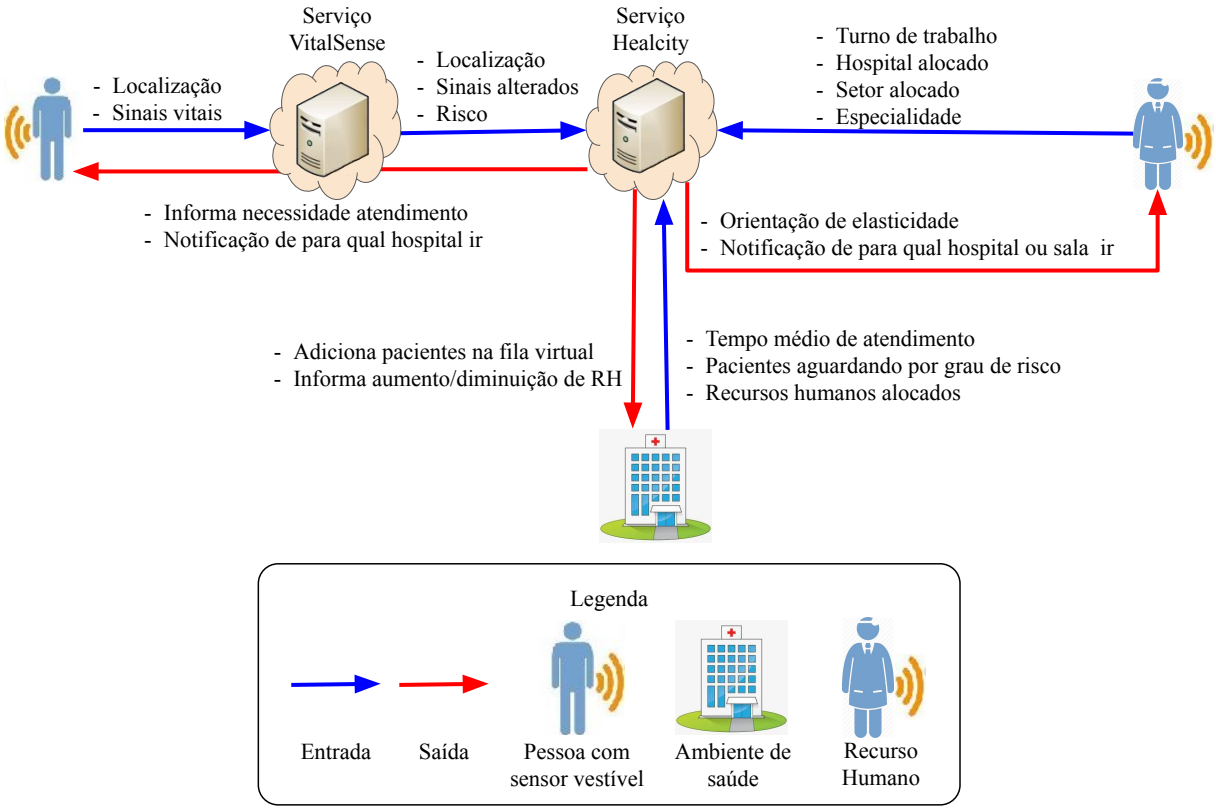
Predict é o módulo responsável por identificar os padrões de movimentação dos pacientes pelo ambiente médico. Com base nos registros de passagens, geradas anteriormente, deve ser identificado o caminho que os pacientes percorrem durante sua movimentação pelo ambiente de saúde, e o tempo de permanência em cada um dos ambientes. Dessa forma, esse módulo deve ser capaz de identificar os padrões relacionados à chegada de pacientes nesses ambientes, os padrões relacionados ao tempo de permanência e espera por atendimento e realizar previsões sobre as chegadas que irão acontecer no futuro. Essa previsão deve ser armazenada de tal forma que possa ser utilizada pelo gerenciador de elasticidade. É o modulo com maior carga de processamento do sistema. A sua execução não ocorre de forma constante como nos módulos anteriores em função de que as chegadas de pacientes nos ambientes de saúde geralmente não ocorrem de forma constante, sendo necessário o processamento em intervalos de tempo pré-definidos. Este módulo pode ser executado no controlador *Edge*, na *Fog*, e na *Cloud*, conforme a necessidade. Em um primeiro momento espera-se que as previsões individuais de cada ambiente de saúde possam ser executadas em seu próprio controlador Edge e as previsões conforme regiões nos nós da *Fog* ou na *Cloud*. Além disso, este módulo é abstrato o suficiente para receber qualquer lista de dados normalizada e aplicar a previsão para valores futuros. Assim, o *Predict* também é capaz de analisar uma série temporal de sinais vitais de determinado paciente e prever os sinais futuros, assim podendo prever uma tendência de piora nos sinais vitais.

Elastic é o módulo responsável pelo gerenciamento de elasticidade do sistema. Ele deve ser capaz de verificar a alocação de recursos humanos em cada um dos ambientes, verificar as passagens atuais de pacientes e comparar os dados com as previsões feitas pelo módulo Predict. Com base nessa comparação, o módulo deve ser capaz de gerar uma alocação dinâmica e automática dos recursos humanos para melhor atender a demanda futura de pacientes. Vale

ressaltar que o sistema gera notificações para os recursos humanos se realocarem, mas a realocação efetiva depende das pessoas cumprirem o que foi indicado pela aplicação. O módulo Elastic juntamente com o módulo Predict são considerados a parte mais importante e o núcleo do modelo proposto. Ao contrário dos módulos anteriores, esse módulo solicita predições ao módulo Predict para tomar ações de elasticidade, ou seja, ele realiza a análise dos recursos com base nas previsões feitas pelo módulo anterior. Na Seção 4.4 serão detalhados os algoritmos e a forma como o módulo realiza o gerenciamento elástico dos recursos humanos no ambiente médico. Assim como o módulo anterior, o Elastic pode ser executado no controlador *Edge*, na *Fog*, e na *Cloud*, conforme a necessidade. Espera-se que as decisões de elasticidade individuais de cada ambiente de saúde possam ser executadas em seu próprio controlador Edge e as decisões de elasticidade a nível de cada região na *Fog* e na *Cloud*.

HealCity app é o aplicativo responsável pela camada de visualização da informação processada pelos outros módulos do serviço HealCity. Através desse módulo os recursos humanos e gestores podem visualizar as informações geradas automaticamente pelo sistema, bem como devem ser capazes de acompanharem o tempo de permanência dos pacientes nas filas, tempo médio de atendimento e espera, e visualizarem as notificações de elasticidade geradas.

Figura 21 – Entradas e Saídas do modelo HealCity.

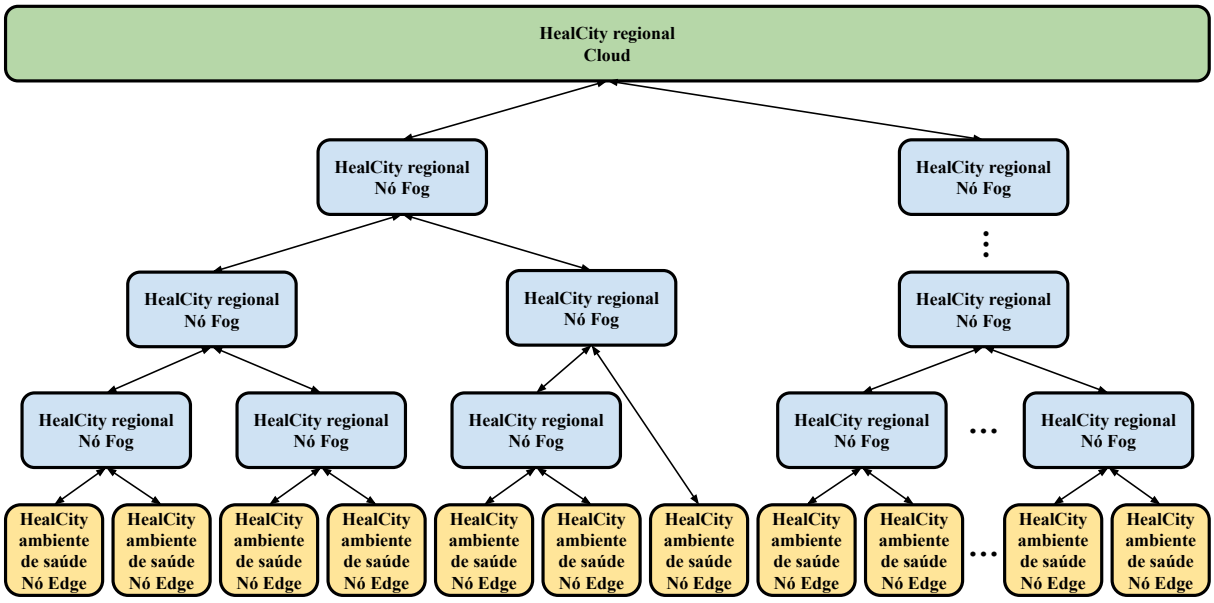


Fonte: Elaborada pelo autor.

Assim, como entrada do modelo, conforme mostrado na Figura 21, os pacientes fornecem sua localização e sinais vitais ao sistema. Os serviços VitalSense tratam os dados, filtrando os

dados apenas dos pacientes com sinais vitais alterados, fornecendo seu risco clínico e localização. Os recursos humanos fornecem ao sistema os seus turnos de trabalho, o ambiente de saúde e o quarto alocado e as especialidades que atendem. Por fim, os ambientes de saúde disponibilizam tempo médio de atendimento em seus quartos, pacientes aguardando de acordo com o nível de risco clínico e recursos humanos alocados em cada quarto. O HealCity pode utilizar esta informação como entrada para gerar e comunicar ações de elasticidade aos recursos humanos, indicando o ambiente de saúde ou quarto específico ao qual serão atribuídos. Além disso, o HealCity pode mostrar aos ambientes de saúde as mudanças necessárias em sua estrutura de atendimento e as previsões de chegadas de pacientes. Por fim, o modelo HealCity deve produzir um resultado que informe às pessoas com sinais vitais alterados para procurarem cuidados médicos e especifique o ambiente de saúde adequado às suas necessidades. O serviço VitalSense é responsável pela comunicação com pessoas fora dos ambientes de saúde. Assim, as notificações emitidas indicando a necessidade de determinado cidadão dirigir-se ao atendimento médico, são enviadas pelo modelo ao VitalSense, que possui seus procedimentos internos para notificar os cidadãos. Dessa forma, o escopo do modelo HealCity foca nos ambientes de saúde, e faz uso do serviço VitalSense para ter acesso a dados externos à estes ambientes.

Figura 22 – Solução hierárquica escalável proposta no modelo HealCity, onde podemos adicionar mais ambientes de saúde em qualquer nó da Fog e quantos nós Fog forem necessários.



Fonte: Elaborada pelo autor.

HealCity propõe uma solução hierárquica escalável na qual os nós Edge nos ambientes de saúde fazem seu processamento próprio e tomam decisões de alocação localmente. Eles se comunicam com um nó da *Fog* superior na hierarquia. Os nós *Fog* não precisam saber se o que está abaixo é um nó Edge de algum ambiente de saúde ou se é outro nó da *Fog*, e nem sabem se o nível acima na hierarquia é o último. Dessa forma, podemos adicionar mais ambientes de saúde em qualquer nó da *Fog* e quantos nós *Fog* forem necessários. Digamos que temos quatro

ambientes de saúde no mesmo bairro. Podemos adicionar um nó da *Fog* para esta vizinhança. Identificamos que dois ambientes de saúde estão mais próximos e os outros dois mais distantes. Podemos criar um nó *Fog* para atender apenas esses dois ambientes de saúde e vincular esse nó da *Fog* e os outros dois ambientes de saúde a outro nó da *Fog*, por exemplo. Assim, no último nível, na *Cloud*, executamos o mesmo algoritmo. A única diferença é que não existe um nível acima, portanto as decisões são finais. A Figura 22 mostra a solução hierárquica escalável proposta no modelo HealCity.

4.4 Elasticidade Multinível de Recursos Humanos em Cidades Inteligentes

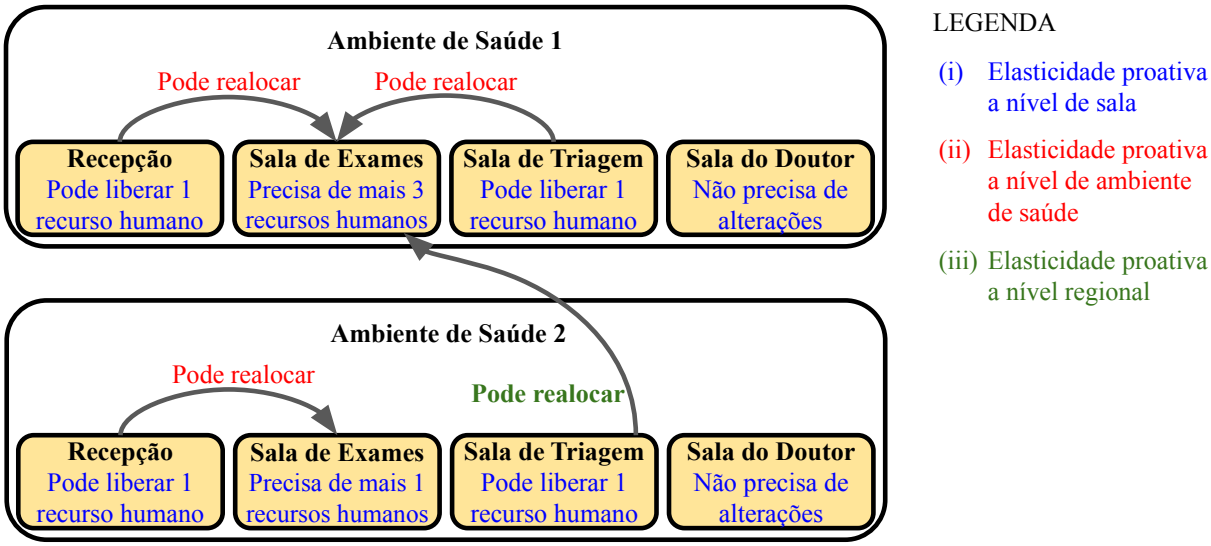
O modelo HealCity utiliza o termo elasticidade de uma forma diferente do conceito de elasticidade da computação em nuvem. No contexto do modelo proposto, elasticidade refere-se a habilidade do sistema de um ambiente de saúde ser capaz de alocar/desalocar/realocar recursos humanos capazes de atender pacientes durante o período de atendimento dos ambientes de saúde de uma cidade inteligente, de forma a se adaptar a variação da demanda de pacientes em diferentes ambientes *on-the-fly*. Para a elasticidade de recursos humanos proposta nesse modelo, é necessário adaptar alguns conceitos da elasticidade da computação em nuvem. Assim, no contexto de recursos humanos, os conceitos de elasticidade devem ser:

- **Alocação**, que refere-se a capacidade do sistema de requisitar profissionais da saúde que não estão em atendimento no ambiente de saúde para se deslocarem para o mesmo e direcioná-los para o atendimento da demanda de pacientes da cidade inteligente;
- **Realocação (ou migração)**, que refere-se a capacidade do sistema de migrar profissionais que estão atendendo em determinado ambiente de saúde para algum outro ambiente, ou outro ambiente de saúde, onde mais recursos humanos são necessários;
- **Desalocação**, que refere-se a capacidade do sistema de liberar os recursos humanos não mais necessários para o atendimento da demanda de pacientes da cidade inteligente.

Para realizar os procedimento de alocação, desalocação e realocação de recursos humanos de forma elástica, o modelo faz uso de uma abordagem multinível para prever a demanda futura de pacientes e do uso dos ambientes de saúde da cidade inteligente. Com base nessa abordagem, o modelo considera a elasticidade de recursos humanos de forma diferente (i) a **nível de sala**, onde nossa abordagem precisa estimar a utilização futura de cada sala e identificar se o número de atendentes é adequado para atender a demanda dos pacientes (conforme será discutido a seguir, na Subseção 4.4.1); (ii) a **nível de ambiente de saúde**, onde o modelo HealCity precisa garantir que o número de atendentes seja adequado para atender a demanda de pacientes em todos as salas do ambiente de saúde, permitindo a movimentação dos atendentes entre os salas se necessário (conforme será detalhado adiante na Subseção 4.4.2), e (iii) a **nível regional**, onde o modelo HealCity deverá verificar se há atendentes suficientes para atender a demanda dos

pacientes nas diferentes regiões da cidade inteligente, com movimentação de atendentes acontecendo entre os ambientes de saúde, e também propondo a movimentação dos pacientes que ainda não foram para qualquer ambiente de saúde (conforme será detalhado na Subseção 4.4.3). Um diagrama desses três níveis é apresentado na Figura 23.

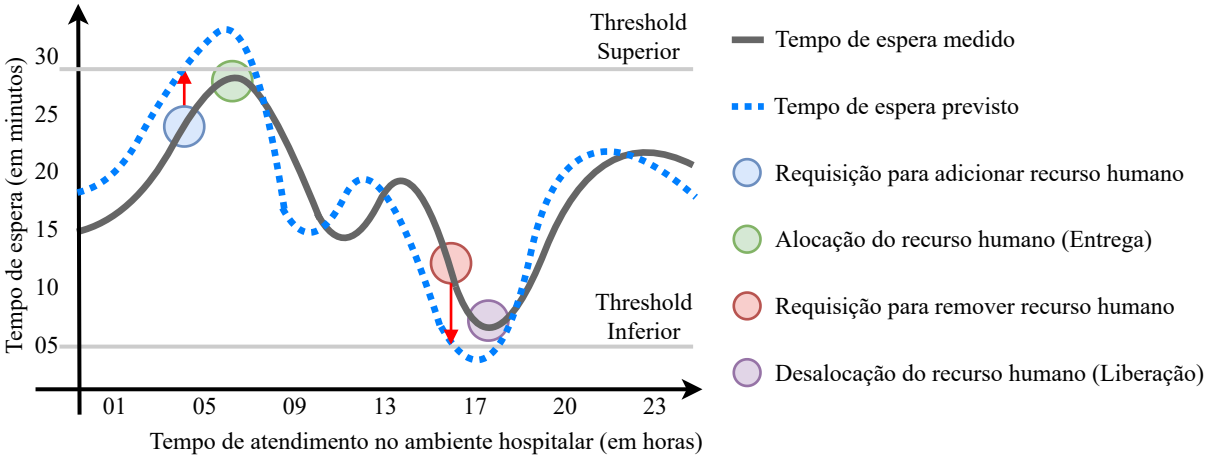
Figura 23 – Exemplo de Elasticidade Multinível de Recursos Humanos em cidades inteligentes com (i) elasticidade proativa a nível da sala, (ii) elasticidade proativa a nível de ambiente de saúde e (iii) elasticidade proativa a nível regional.



Fonte: Elaborada pelo autor.

O modelo HealCity adapta a estratégia de elasticidade utilizando *thresholds* superior e inferior da elasticidade de recursos para o contexto de pessoas, com base na demanda atual e prevista de pacientes. A Figura 24 demonstra a utilização de *thresholds* aonde em um certo momento a previsão do sistema informa que o *threshold* superior será alcançado e os recursos humanos devem ser alocados, e logo em seguida a previsão informa que o *threshold* inferior será atingido e os recursos humanos devem ser liberados.

Figura 24 – Elasticidade preditiva baseada em tempo de espera adotado pelo HealCity.



Fonte: Elaborada pelo autor.

4.4.1 Elasticidade Proativa a Nível de Sala

No nível da sala, o HealCity deve reconhecer os requisitos específicos de uma determinada sala. Nesta fase, o gerenciador da elasticidade de recursos humanos não deve preocupar-se com outras salas nos ambientes de saúde ou na cidade, mas apenas determinar quantos atendentes são necessários para satisfazer a sua própria demanda futura de atendimento. Para tanto, é gerada uma série temporal para o número de pacientes que chegam em cada momento e para o tempo de atendimento para servir de entrada ao modelo de predição ARIMA. Consequentemente, quando o modelo detecta tempos de espera que não estão de acordo com os limites estabelecidos, o HealCity deve calcular a quantidade de recursos de saúde necessários para atender a demanda de pacientes, reconhecendo assim a necessidade de ajustes naquela sala em específico. O modelo HealCity utiliza uma alocação paralela de recursos humanos, inspirada em estratégias semelhantes usadas em sistemas elásticos (AL-DHURAIBI et al., 2017) e computação de alto desempenho (ROSA RIGHI et al., 2016). Para calcular o número de recursos de saúde necessários para atender a demanda, o HealCity propõe alguns formalismos matemáticos que serão descritos na sequência. A Tabela 4 apresenta algumas nomenclaturas que serão utilizadas nestes formalismos matemáticos propostos no modelo.

Tabela 4 – Nomenclaturas para os formalismos matemáticos do modelo HealCity.

Nomenclatura	**Significado**
r	Sala do ambiente de saúde
t_n	Instante de tempo n específico
t_f	Instante de tempo final
t_i	Instante de tempo inicial
f_i	Instante de tempo inicial futuro
f_f	Instante de tempo final futuro
a	Atendentes alocados
$size(x)$	Tamanho de um vetor x
$CV(r, t_i, t_f)$	Vetor de atendimentos
$CDT(x[i])$	Tempo de Duração de um Atendimento
$ACT(r, t_i, t_f)$	Tempo Médio de Atendimento
$NA(r, t_n)$	Número de Atendentes
$ANA(r, t_i, t_f)$	Número Médio de Atendentes
$NWP(r, t_i)$	Número de Pacientes Aguardando
$NIP(r, t_n)$	Número de Pacientes Chegando
$ENP(r, t_i, t_f)$	Número Estimado de Pacientes
$ECT(r, t_i, t_f)$	Tempos Estimado de Atendimento
$HRES(r, t_i, t_f)$	Speedup Elástico de Recursos Humanos
$PHRES(r, a, f_i, f_f)$	Speedup Elástico Proativo de Recursos Humanos
$ATV(r)$	Vetor de Tempos de Alocação
$ADT(x[i])$	Tempo de Alocação de um Atendente
$EAT(r)$	Tempo Estimado de Alocação em uma sala r específica
$Cost(o, d)$	Custo para Alocação ou Realocação
$C(r, t_i, t_f)$	Confiança
$FNIP(r, f_i, f_f)$	Número Futuro de Pacientes Chegando
$PNIP(r, f_i, f_f)$	Número Previsto de Pacientes Chegando

Fonte: Elaborada pelo autor.

O Conjunto de Atendimentos (*Care Vector*) da sala "r" entre os tempos "t_i" e "t_f" pode ser encontrado através da função $CV(r, t_i, t_f)$. Além disso, a função $size(x)$ é responsável por retornar o tamanho de um vetor qualquer x. Com o uso dessa duas funções, é possível determinar o Tempo Médio de Atendimento (*Average Care Time*) na sala "r" do ambiente de saúde entre os tempos "t_i" e "t_f", definido pela Equação 4, onde $CDT(x[i])$ refere-se ao Tempo de Duração de um Atendimento (*Care Duration Time*) $x[i]$ que já ocorreu naquela sala e $x[]$ é o conjunto de atendimentos que ocorreram na sala.

$$ACT(r, t_i, t_f) = \frac{1}{size(x)} \sum_{i=0}^{size(x)-1} CDT(x[i]) \tag{4}$$

A Equação 4 tem como resultado um valor numérico de tempo. Um exemplo para a equação seria uma sala qualquer "r", entre os tempos 1 e 5, onde o resultado poderia ser definido como: $ACT(r, 1, 5) = 15$ minutos. Com o uso dessa equação, é possível estimar o tempo médio que levará um atendimento em uma determinada sala do ambiente de saúde. Em função da elasticidade de recursos humanos, em diferentes instantes de tempo existem diferentes quantidades de atendentes alocados para atendimento em cada uma das salas dos ambiente de saúde. Assim, O Número Médio de Atendentes (*Average Number of Attendants*) na sala "r" do ambiente de saúde entre os tempos "t_i" e "t_f" é definido pela Equação 5, onde $NA(r, t_n)$ refere-se ao Número de Atendentes (*Number of Attendants*) alocados para atendimento na sala "r" no instante de tempo "t_n".

$$ANA(r, t_i, t_f) = \frac{1}{t_f - t_i} \sum_{t_n=t_i}^{t_f-1} NA(r, t_n) \tag{5}$$

Um exemplo seria a sala r, com diferentes alocações de atendentes entre os tempos 0 e 7 conforme pode ser visto na Figura 25. Se fosse necessário identificar o número médio de atendentes dessa sala entre os tempos 1 e 5, por exemplo, a Equação 5 poderia ser aplicada da seguinte forma:

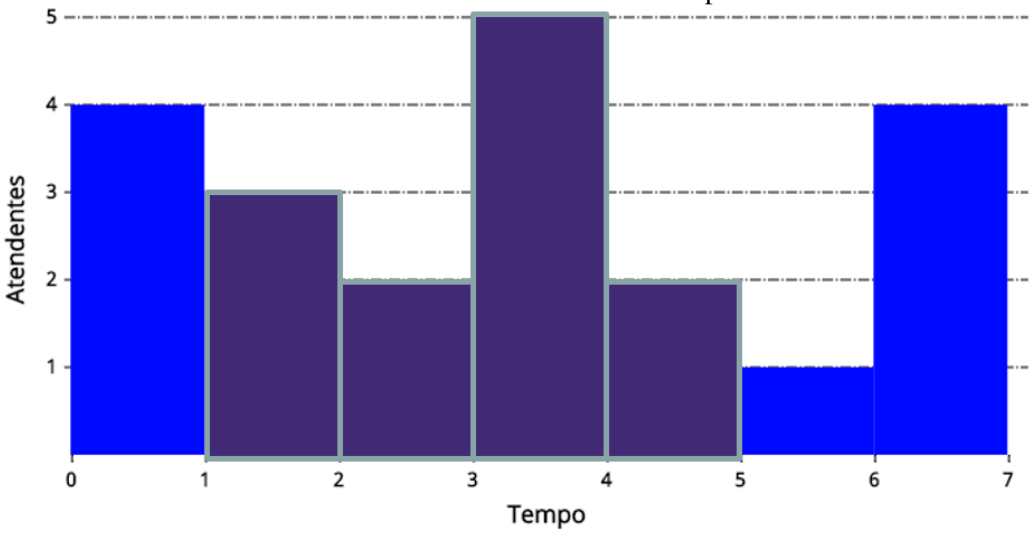
$$ANA(r, 1, 5) = \frac{1}{5-1} \sum_{i=1}^{5-1} NA(r, i) = \frac{1}{4}(NA(r, 1) + NA(r, 2) + NA(r, 3) + NA(r, 4))$$

Na situação hipotética exposta na Figura 25, no instante de tempo 1 haviam 3 atendentes alocados para atendimento, no instante 2 haviam 2 atendentes, no instante 3 haviam 5 atendentes e no instante de tempo 4 haviam novamente 2 atendentes. Dessa forma, seguindo a equação, nesse intervalo de tempo estariam alocados em média 3 atendentes, conforme pode ser visto continuando o cálculo:

$$ANA(r, 1, 5) = \frac{3 + 2 + 5 + 2}{4} = \frac{12}{4} = 3 \text{ atendentes}$$

A mesma ideia da função anterior pode ser adequada para realidade de pacientes, pois em

Figura 25 – Calculando o Número Médio de Atendentes em uma sala do ambiente de saúde entre os instantes de tempo 1 e 5, onde a cada instante de tempo a quantidade de atendentes alocados é alterada, chegando a um número médio de 3 atendentes alocados nessa sala hipotética entre os instantes de tempo.



Fonte: Elaborada pelo autor.

diferentes instantes de tempo existem diferentes quantidades de pacientes aguardando atendimento em cada uma das salas dos ambiente de saúde. Assim, o Número Estimado de Pacientes (*Estimated Number of Patients*) aguardando para atendimento na sala "r" do ambiente de saúde entre os tempos "t_i" e "t_f" é definida pela Equação 6, onde $NWP(r, t_i)$ refere-se ao número de pacientes aguardando (*Number of Waiting Patients*) atendimento na sala "r" no instante de tempo "t_i" e $NIP(r, t_n)$ refere-se ao número de pacientes chegando (*Number of Incoming Patients*) na sala "r" para atendimento no instante de tempo "t_n".

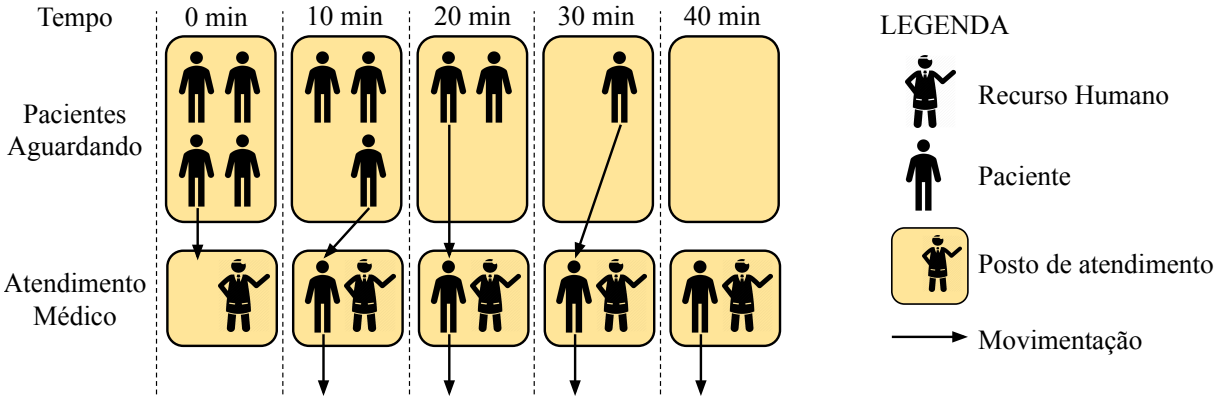
$$ENP(r, t_i, t_f) = NWP(r, t_i) + \sum_{t_n=t_i+1}^{t_f-1} NIP(r, t_n) \tag{6}$$

Utilizando as equações propostas anteriormente, o modelo se torna capaz de chegar ao Tempo Estimado para Atendimento (*Estimated Care Time*) de todos os pacientes aguardando, ou seja, o tempo que um novo paciente entrando na fila precisa esperar para ser atendido. O $ECT(r, t_i, t_f)$ é definido pela Equação 7, onde $ACT(r, t_i, t_f)$ refere-se ao tempo médio para os atendimentos na sala "r" entre os tempos "t_i" e "t_f" e $ENP(r, t_i, t_f)$ refere-se ao número estimado de pacientes (*Estimated Number of Patients*) que estão esperando na fila para atendimento na sala "r" entre os tempos "t_i" e "t_f". Um exemplo seria a sala r_1, entre dois tempos t_i e t_f que resultassem em uma quantidade média de 4 pacientes e um tempo médio de atendimento de 10 minutos conforme pode ser visto na Figura 26.

$$ECT(r, t_i, t_f) = ACT(r, t_i, t_f) \cdot ENP(r, t_i, t_f) \tag{7}$$

Na situação hipotética exposta na Figura 26, no instante de tempo $0min$ haviam 4 paci-

Figura 26 – Calculando o Tempo Estimado para Atendimento em uma sala do ambiente de saúde com 4 pacientes aguardando e tempo médio de atendimento de 10 minutos, chegando ao Tempo Estimado de Atendimento de 40 minutos.



Fonte: Elaborada pelo autor.

entes aguardando e nenhum em atendimento pelo médico, ou seja, nesse instante o primeiro paciente foi encaminhado ao atendimento. Já no instante $10min$, 3 pacientes estão aguardando, o primeiro paciente termina seu atendimento e vai embora, sendo assim o segundo paciente é encaminhado ao atendimento. No instante $20min$, 2 pacientes estão aguardando, o segundo paciente termina seu atendimento e vai embora, e o terceiro paciente é chamado para o atendimento. No instante $30min$, 1 paciente está aguardando, o terceiro paciente é liberado, e o último paciente é chamado para o atendimento. Assim, no instante $40min$, o último paciente é liberado, de forma que todos os pacientes foram atendidos em 40 minutos. Aplicando-se a Equação 7, chega-se ao mesmo resultado:

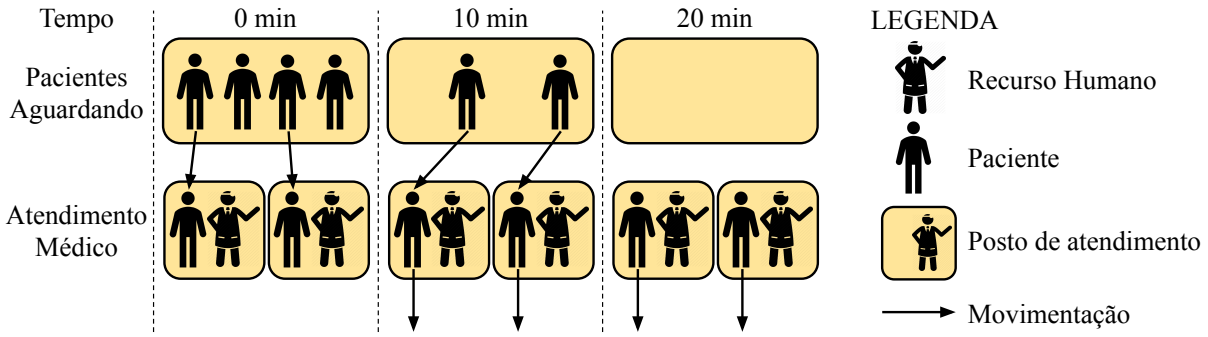
$$ECT(r, t_i, t_f) = ACT(r, t_i, t_f) \cdot ENP(r, t_i, t_f) = 10 \cdot 4 = 40 \text{ minutos}$$

Sabendo o $ECT(r, t_i, t_f)$, é possível analisar qual o tempo médio que levou para o atendimento de todos os pacientes aguardando na sala r entre os tempos "t_i" e "t_f". Entretanto, esse valor refere-se a uma sala do ambiente de saúde com um único atendente alocado para atendimento, mas na maioria dos casos haverá mais de um profissional da saúde atuando naquela sala, tornando-se necessário identificar qual o tempo médio com diferentes quantidades de atendentes. Nesse contexto, o modelo faz uso de uma alocação paralela de recursos humanos, tal como a alocação paralela de máquinas virtuais utilizada nos sistemas elásticos (AL-DHURAIBI et al., 2017) ou a utilização de processadores em paralelo na computação de alto-desempenho (ROSA RIGHI et al., 2016). Dessa forma, baseado no Speedup Elástico Reativo proposto por Rosa Righi et al. (2016), o modelo utiliza a Equação 8 para Speedup Elástico de Recursos Humanos (*Human Resources Elastic Speedup*).

$$HRES(r, t_i, t_f) = \frac{ECT(r, t_i, t_f)}{ANA(r, t_i, t_f)} \tag{8}$$

Um exemplo seria a mesma sala r anterior, entre dois tempos t_i e t_f com uma quantidade

Figura 27 – Calculando o tempo estimado de atendimento em uma sala do ambiente de saúde usando alocação paralela de atendentes, com 4 pacientes aguardando, tempo médio de atendimento de 10 minutos e 2 atendentes, chegando ao tempo estimado de atendimento, através do Speedup Elástico de Recursos Humanos, de 20 minutos.



Fonte: Elaborada pelo autor.

média de 4 pacientes, um tempo médio de atendimento de 10 minutos e com a alocação de 2 profissionais de saúde, conforme pode ser visto na Figura 27. Nessa situação hipotética, no instante de tempo $0min$ haviam 4 pacientes aguardando e nenhum em atendimento pelo médico, ou seja, nesse instante os dois primeiros pacientes foram encaminhado ao atendimento. Já no instante $10min$, 2 pacientes estão aguardando, e os dois primeiros pacientes são liberados, sendo assim os dois últimos pacientes são encaminhados ao atendimento. Assim, no instante $20min$, os dois últimos pacientes são liberados, de forma que todos os pacientes foram atendidos em apenas 20 minutos. Aplicando-se a Equação 8, chega-se ao mesmo resultado:

$$HRES(r, t_i, t_f) = \frac{ACT(r, t_i, t_f) \cdot ENP(r, t_i, t_f)}{ANA(r, t_i, t_f)} = \frac{10 \cdot 4}{2} = \frac{40}{2} = 20 \text{ minutos}$$

A função $HRES(s, ti, tf)$ retorna o tempo estimado de atendimento da sala "r" entre os tempos "t_i" e "t_f", considerando a alocação paralela de atendentes nesse período de tempo, através da função $ANA(r, t_i, t_f)$. Assim, com o aumento do número médio de atendentes alocados, o tempo estimado de atendimento acaba diminuindo, de forma inversamente proporcional. Um problema da elasticidade reativa é que as ações de elasticidade são realizadas após o limite superior ser atingido, causando um estado de sobrecarga no ambiente de saúde durante todo o período de movimentação dos profissionais. Assim, uma alternativa para este problema é o uso da elasticidade proativa (RIGHI et al., 2019). Assim, antecipando o momento em que o limite superior será atingido, a movimentação das pessoas pode ocorrer antecipadamente, minimizando ou evitando sobrecargas dos pacientes no ambiente de saúde. Neste contexto, foi utilizada a Equação 9 para Speedup Elástico Proativo de Recursos Humanos (*Proactive Human Resources Elastic Speedup*) através da Equação 9:

$$PHRES(r, a, f_i, f_f) = \frac{ECT(r, f_i, f_f)'}{a}, \tag{9}$$

onde a é o número de atendentes alocados entre os tempos futuros f_i e f_f, e $ECT(r, f_i, f_f)'$ é uma previsão do tempo de atendimento futuro para esta sala usando ARIMA e uma análise dos sinais vitais alterados de pessoas em suas casas que precisarão de atendimento médico em um futuro próximo. Podemos calcular ECT' como:

$$ECT(r, f_i, f_f)' = ACT(r, f_i, f_f)' \cdot ENP(r, f_i, f_f)',$$

onde $ACT(r, f_i, f_f)'$ e $ENP(r, f_i, f_f)'$ são predicções do tempo médio de atendimento e número futuro de pacientes na sala r, respectivamente. Assim, para cada sala r sendo calculada, geramos uma série temporal de $ACT(r, t_i, t_f)$ que ocorreu no passado, e a usamos para prever $ACT(r, f_i, f_f)'$. Além disso, para cada sala, exceto a sala de recepção, também geramos uma série temporal para $NIP(r, t_i, t_f)$ e podemos prever futuras entradas de pacientes e encontrar $ENP(r, f_i, f_f)'$. Para a sala de recepção, ao invés de gerar uma série temporal para $NIP(r, t_i, t_f)$, analisaremos efetivamente as pessoas com sinais alterados em suas residências para definir quais chegarão ao ambientes de saúde para encontrar $ENP(r, f_i, f_f)'$, conforme será detalhado na Seção 4.5.

Algoritmo 1: Elasticidade Proativa a Nível de Sala

Entrada: Sala r, a atendentes, tempo inicial futuro f_i, tempo final futuro f_f
Saída: Quantidade de atendentes a serem liberados ou alocados

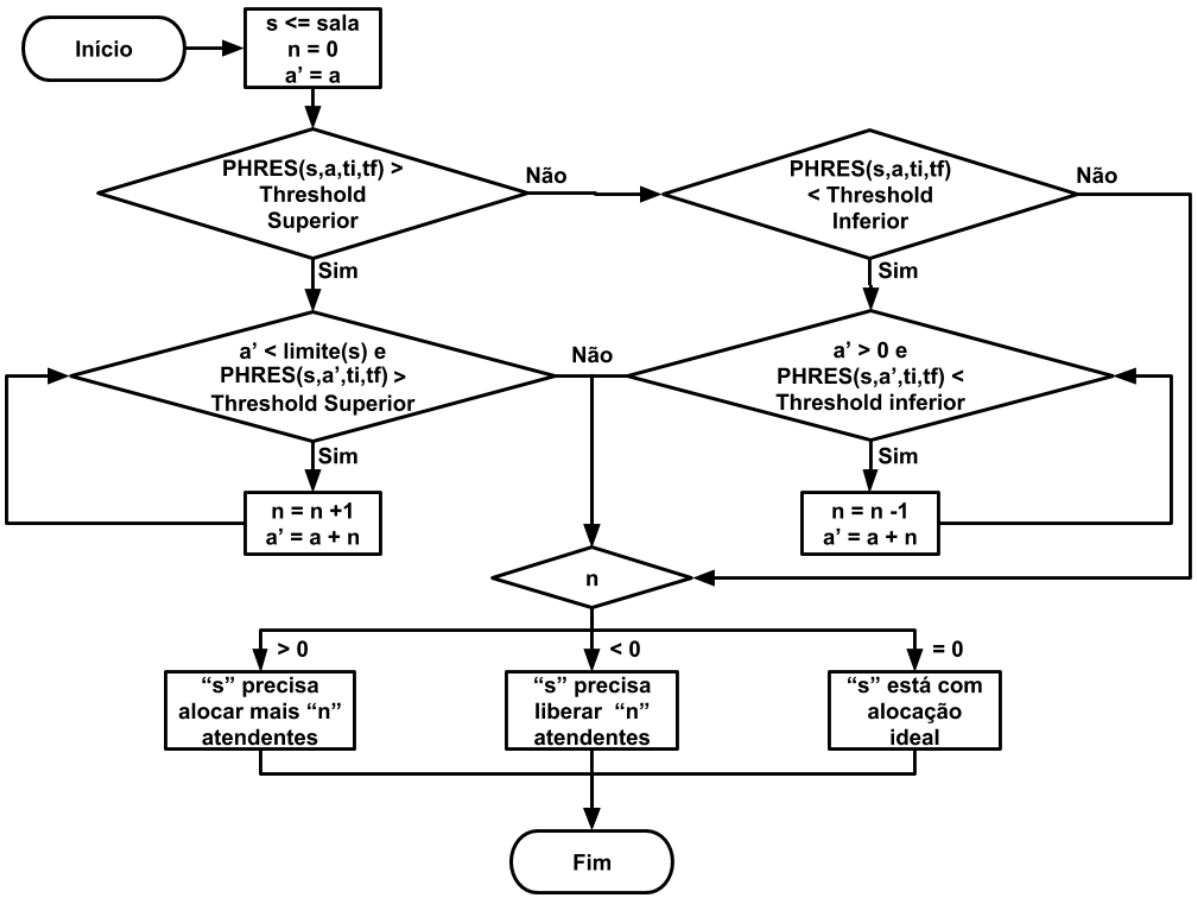
```
1  início
2      upper ← Threshold Superior de tempo de espera na sala r;
3      lower ← Threshold Inferior de tempo de espera na sala r;
4      n ← 0;
5      a' ← a;
6      se PHRES(r, a, f_i, f_f) > upper então
7          enquanto a' < limit(r) e PHRES(r, a', f_i, f_f) > upper faça
8              n ← n + 1;
9              a' ← a + n;
10         fim
11     senão se PHRES(r, a, f_i, f_f) < lower então
12         enquanto a' > 0 e PHRES(r, a', f_i, f_f) < lower faça
13             n ← n − 1;
14             a' ← a + n;
15         fim
16     fim
17     retorna n;
18 fim
```

Usando as equações acima, o HealCity pode prever o tempo de espera de qualquer sala de um ambiente de saúde. Variando o atributo a na equação $PHRES$, com o aumento e diminuição do número de profissionais de saúde em atendimento, o HealCity consegue identificar quantos atendentes seriam necessários para ajustar o tempo de espera de qualquer sala aos limi-

tes definidos por um gestor do ambiente de saúde. O Algoritmo 1 apresenta o pseudo-código proposto para verificar a necessidade de alocar ou desalocar recursos humanos em qualquer sala r em um ambiente de saúde inteligente com complexidade $O(n)$.

Assim, o HealCity se torna capaz de filtrar e realizar previsões complementares com o aumento e diminuição do número de atendentes sempre que necessário. Nesse ponto, o algoritmo é capaz realizar previsões do tempo estimado de espera das salas com alertas, com o aumento e a diminuição do número de profissionais de saúde em atendimento de forma a identificar quantos atendentes seriam necessários para adequar o tempo da sala aos limites propostos pela gestão do ambiente de saúde. Pode-se dizer que a predição a nível de sala é proativa, ocorrendo a todo momento, e é capaz de prever a super ou subutilização de cada uma das salas do ambiente de saúde através da utilização do Algoritmo 1. Dessa forma, o algoritmo proposto pode ser representado através do fluxograma de processos apresentado na Figura 28.

Figura 28 – Fluxograma de processos a nível de sala para verificar a possibilidade de alocar ou liberar recursos humanos em uma sala qualquer do ambiente de saúde inteligente.



Fonte: Elaborada pelo autor.

4.4.2 Elasticidade Proativa a Nível de Ambiente de Saúde

A nível de ambiente de saúde, o modelo HealCity tem a tarefa de gerenciar requisições específicas das salas em toda a extensão do ambiente de saúde em análise. O objetivo é garantir que cada sala tenha o número necessário de atendentes, considerando apenas a realocação de profissionais de saúde entre diferentes setores ou a desalocação de recursos humanos que já não são necessários. Dessa forma, o modelo precisa testar diferentes alocações para os atendentes para garantir que todas as salas identificadas na etapa anterior (Subseção 4.4.1) tenham atendentes suficientes para minimizar a superlotação. Para tanto, uma vez que o algoritmo a nível de sala analisou as estimativas de tempo de espera e identificou a existência de uma sala com demanda crescente de pacientes, e ao mesmo tempo uma outra sala com demanda decrescente, o algoritmo a nível de ambiente de saúde é acionado. O objetivo do algoritmo é verificar e propor as movimentações de pessoas necessárias para o ajuste do tempo de espera. Para redistribuir tais atendentes de saúde entre diferentes salas médicas, o modelo utiliza algumas estratégias conhecidas de outros contextos da computação científica e as adapta à necessidade da elasticidade proativa de recursos humanos. O Algoritmo 2 apresenta o pseudocódigo para elasticidade proativa em nível de ambiente de saúde com complexidade $O(2n^2)$.

Algoritmo 2: Elasticidade Proativa a Nível de Ambiente de Saúde

Entrada: Lista h de salas do ambiente de saúde, vetor v com todos os atendentes do ambiente de saúde, tempo inicial futuro f_i, tempo final futuro f_f

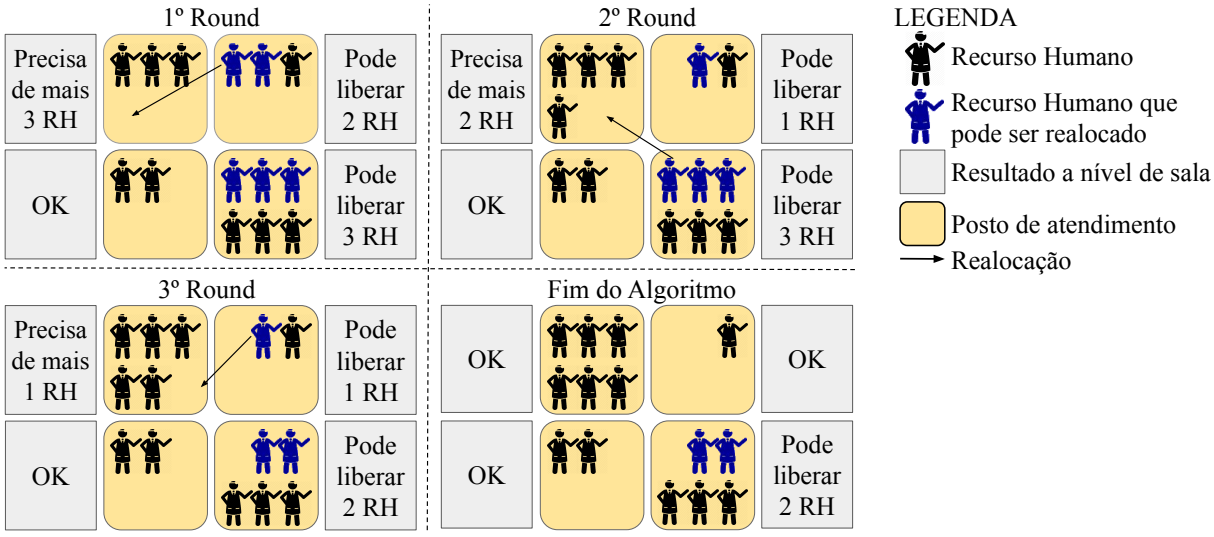
Saída: Vetor l de salas e quantidade de atendentes a serem alocados ou desalocados

1 **início**
2 $l \leftarrow$ novo vetor de salas e quantidade de atendentes a serem alocados ou desalocados;
3 **para cada Sala** r **na lista** h **de salas do ambiente de saúde faça**
4 $a \leftarrow$ número de atendentes alocados em r;
5 $q \leftarrow$ executa Algoritmo 1 para *Elasticidade Proativa a Nível de Sala* usando r, a, f_i and f_f como Entrada;
6 $l.add(r, q)$;
7 **fim**
8 ordena l, quantidade de atendentes disponíveis;
9 $l \leftarrow$ executa *Algoritmo 4 de Desalocação de Recursos Humanos* usando l e os atendentes alocados em v como Entrada;
10 ordena l, quantidade de atendentes disponíveis;
11 **para cada Sala** r **na lista** l **faça**
12 $l_r \leftarrow$ ordena l, quantidade de atendentes disponíveis com a especialidade da sala r;
13 executa *Algoritmo 3 de Human Resources List Scheduling* usando r e l_r como Entrada;
14 **fim**
15 $h \leftarrow$ salas do vetor l;
16 **retorna** l;
17 **fim**

No modelo HealCity, cada sala pode possuir uma especialidade exigida aos recursos humanos que nela estão alocados. Paralelamente, cada profissional de saúde tem uma lista com todas as suas especialidades. O processo de realocação de recursos humanos é realizado apenas entre profissionais que possuam a especialidade de sala de destino exigida. Isso é necessário porque em uma sala de exames laboratoriais é necessário um profissional de enfermagem acostumado a exames de sangue por exemplo, e mesmo que tenhamos técnicos de raio-x disponíveis para realocação, eles não conseguem melhorar o atendimento na referida sala. Um ponto a ser observado é que naquelas salas onde necessitam de uma especialidade que nenhum outro profissional em atendimento possui, a única alternativa é a a alocação de novos recursos humanos.

A seguir, é discutido primeiramente o conceito de realocação, seguido dos procedimentos de desalocação. Observa-se que, embora a desalocação apareça primeiro no algoritmo (Linha 10 do Algoritmo 2), na verdade o procedimento se baseia nos recursos humanos alocados durante a iteração anterior do algoritmo. A primeira opção do HealCity deve ser sempre a possibilidade de realocação dos recursos humanos já alocados no atendimento. A realocação é priorizada pois é a opção que traz menos custos, uma vez que realiza o ajuste do atendimento sem o acréscimo de atendentes. Dessa forma, para realizar-se a realocação de recursos humanos, em um primeiro momento pensou-se em adaptar o algoritmo *Round Robin* utilizado para escalonamento de processos, conforme Silberschatz, Galvin e Gagne (2012), para o contexto elástico de recursos humanos. Nesse algoritmo vislumbrado, todas as salas devem estar em uma espécie de fila circular no sistema, e sempre que uma sala necessitar de mais atendentes, o gerenciador de elasticidade dever procurar na próxima sala da fila se há algum atendente disponível. Caso exista, este é encaminhado para a sala que necessita. Caso não exista, ou a sala de origem ainda necessite de mais atendentes, o algoritmo deve verificar na sala sequente e assim sucessivamente até a sala de origem obter todos os atendentes da qual necessita.

Figura 29 – Exemplo de realocação através do algoritmo de *Round Robin* vislumbrado, com uma fila de 4 salas, e 14 atendentes, onde foi previsto que a Sala 1 precisa alocar 3 atendentes.



Fonte: Elaborada pelo autor.

A Figura 29 ilustra um cenário de execução para o algoritmo de *Round Robin* vislumbrado, onde a Sala 1 precisa de mais 3 atendentes e as Salas 2 e 4 possuem atendentes disponíveis. Seguindo a lógica da fila circular, no 1º *Round* um atendente da Sala 2 é realocado para a Sala 1, logo após no 2º *Round* um atendente da Sala 4 é realocado e, por fim, no 3º *Round* um atendente da Sala 2 é transferido para a Sala 1. Após uma análise da proposta de algoritmo, pode-se encontrar um possível problema no mesmo, onde em função da alocação ser circular, sempre um atendente será retirado de uma sala diferente por *Round*. Dessa forma, uma sala que possui 2 atendentes para serem liberados, que é o caso da Sala 4 no exemplo anterior, acaba por ser mais impactada do que a Sala 2, com 3 atendentes disponíveis, pois ao final da execução do algoritmo, a Sala 4 ficou no limite mínimo de pessoal, enquanto a Sala 2 ficou com 2 atendentes para liberar ainda, de forma que o processo de movimentação de recursos humanos se torna desbalanceado. Assim, no próximo ciclo de monitoramento, as chances da Sala 4 precisar de um de seus atendentes de volta é maior do que a Sala 2.

Pensando nisso, para realizar-se a realocação de recursos humanos, de forma balanceada, foi decidido adaptar o algoritmo *List Scheduling* dinâmico, também utilizado para escalonamento de processos segundo Wang e Sinnen (2018), onde todas as salas do ambiente de saúde estão em uma espécie de lista ordenada pela quantidade de atendentes disponíveis para realocação. Dessa forma, sempre que uma sala r necessitar de mais atendentes, o gerenciador de elasticidade deve procurar na primeira sala da lista l_r se há algum atendente disponível com a especialidade da sala r. Caso exista, este é encaminhado para a sala que necessita, e a lista deve ser ordenada novamente. Se a sala de origem ainda necessita de mais atendentes, o algoritmo deve procurar no primeiro da lista novamente, e assim sucessivamente até a sala de origem obter todos os atendentes da qual necessita. O Algoritmo 3 apresenta o pseudo-código proposto para esse processo com complexidade $O(n)$.

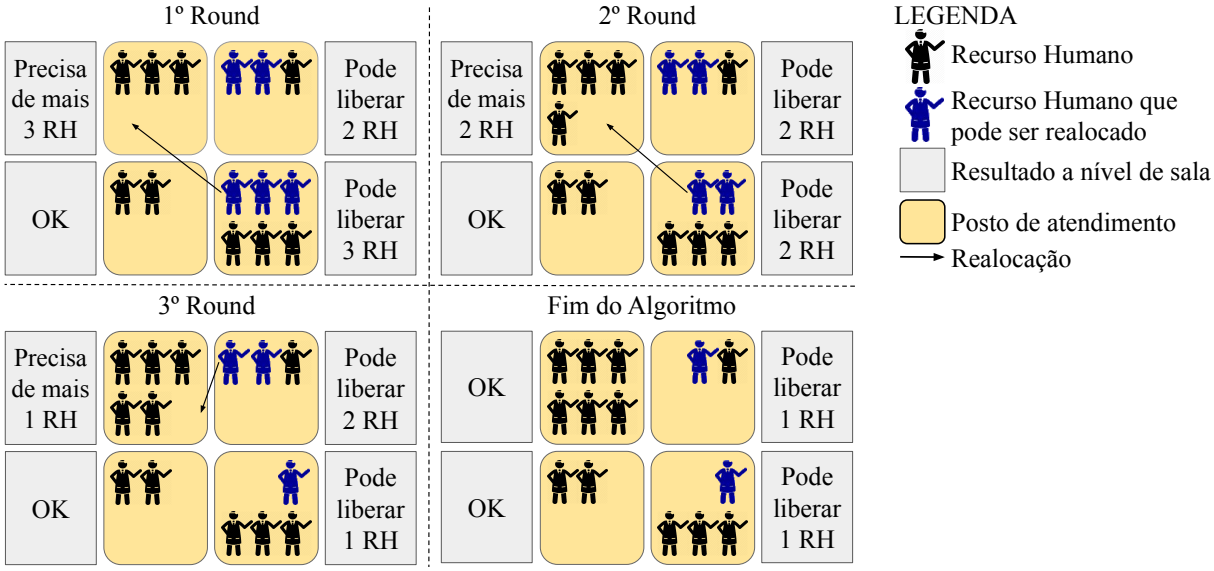
Algoritmo 3: Realocação através do algoritmo de *Human Resource List Scheduling*

Entrada: Sala r que precisa de atendentes, e a lista ordenada l_r com todas as salas
Saída: Situação final da sala r

```
1  início
2      next ← primeira sala da lista l_r;
3      enquanto s precisa de atendentes e há atendentes disponíveis em l_r com a
         especialidade da sala r faça
4          r recebe um atendente de next;
5          l_r ← a lista l_r reordenada;
6          se r ainda precisa de atendentes então
7              next ← primeira sala da lista l_r;
8          senão
9              retorna r está com a alocação adequada;
10     fim
11     retorna r ainda precisa de atendentes;
12 fim
```

A Figura 30 ilustra o mesmo cenário apresentado anteriormente, para execução do algoritmo proposto, onde a Sala 1 precisa de mais 3 atendentes e as Salas 2 e 4 possuem atendentes disponíveis. Seguindo a lógica de *List Scheduling*, no 1º *Round* a Sala 4 é a primeira da lista, com 3 atendentes disponíveis, e fornece um atendente para a Sala 1. Já no 2º *Round* apesar de ambas as salas estarem com a mesma quantidade de atendentes disponíveis a Sala 4 continua sendo a primeira da lista, agora com 2 profissionais da saúde para serem liberados, e fornece um atendente para realocação. Por fim, no 3º *Round*, a Sala 2 tornou-se a primeira da lista, com 2 recursos humanos disponíveis, contra apenas 1 da Sala 2, e então fornece um atendente para a Sala 1. Após uma análise desse algoritmo utilizando uma lista ordenada, pode-se notar que o mesmo não apresenta o mesmo problema da alocação por fila circular, uma vez que existe um balanceamento de carga entre as salas do ambiente de saúde. Dessa forma, ambas as salas que possuíam atendentes disponíveis para liberação no exemplo ilustrado na Figura 30 ficaram com a mesma quantidade de recursos humanos disponíveis ao final da execução. Assim, no próximo ciclo de monitoramento, ambas as salas possuem as mesmas chances de necessitarem de novos profissionais para atendimento. Então, para o processo de realocação de recursos humanos optou-se pelo uso do algoritmo de *List Scheduling* adaptado, nesse trabalho nomeado de *Human Resource List Scheduling* (HRLS). Um ponto importante a ser destacado é que a realocação só ocorre quando não vai haver degradação significativa do tempo de espera com este procedimento. Assim, um profissional só é realocado se a sua saída da sala não faça com que o tempo de espera da mesma supere o *threshold* superior definido pelo gestor do ambiente de saúde.

Figura 30 – Exemplo de realocação através do algoritmo *List Scheduling* adaptado, com uma lista de 4 salas, ordenada pela quantidade de recursos humanos disponíveis para elasticidade, e 12 atendentes, onde foi previsto que a Sala 1 precisa alocar mais 5 profissionais da saúde.



Fonte: Elaborada pelo autor.

Um problema da elasticidade, tanto na mecânica, quanto na computação em nuvem, é a

histerese, isto é, a tendência do sistema de retornar ao estado anterior na ausência do impulso que ocasionou a mudança. No contexto da elasticidade de recursos humanos, a *histerese* pode ocorrer quando em um determinado instante de tempo uma sala fornece atendentes para outra, e no instante seguinte, ela precisa deles de volta, e a sala destino os fornece. Um exemplo, seria a Sala 1 precisar de mais dois profissionais para atendimento e a Sala 2 fornecer esses atendentes para a mesma. Como o tempo de espera da Sala 1 vai melhorar, o impulso que ocasionou a mudança deixa de existir. Já como a Sala 2 teve uma quantidade menor de funcionários em atendimento, provavelmente o tempo de espera vai piorar e a mesma vai precisar de novos atendentes. Nesse contexto, a Sala 1 vai acabar devolvendo os profissionais anteriormente realocados para a Sala 2, ocasionando o problema da *histerese*. A fim de tratar o problema da *histerese* na elasticidade de recursos humanos, decidiu-se por utilizar uma estratégia baseada em um *cooldown period* (YU; FU, 2023) para evitar que ocorra o efeito indesejado. Assim, definiu-se que o *cooldown period* para e elasticidade baseia-se no fato de que sempre que um recurso é realocado de uma determinada sala A para outra sala B, se a sala A precisar desse recurso de volta no intervalo de tempo subsequente, sua necessidade só será atendida se outra sala tiver recursos livres, evitando o efeito de histerese. Dessa forma, no exemplo anterior, se a Sala 1 receber recursos humanos da Sala 2 no primeiro ciclo, no próximo ciclo, mesmo se a Sala 2 necessitar de atendentes, os mesmos serão fornecidos de outra fonte, evitando que ocorra a *histerese*. É importante salientar que esta estratégia pode aumentar o tempo de resposta, mas nunca deixará uma sala desassistida. Caso uma sala necessite com maior urgência de um recurso que já disponibilizou, o modelo designará outra pessoa para esse cargo. Portanto, o HealCity aceita um ligeiro aumento de custo para reduzir os efeitos indesejados da histerese.

Algoritmo 4: Desalocação de Recursos Humanos

Entrada: Lista ordenada com todas as salas l do ambiente de saúde e o vetor a com todas os atendentes alocados no ambiente de saúde

Saída: Lista de salas l atualizada

```
1  início
2      qd ← quantidade de atendentes disponíveis nas salas de l;
3      qf ← quantidade de atendentes faltando nas salas de l ;
4      enquanto qd < qf e size(a) > 0 faça
5          ordena a, por tempo em atendimento decrescente;
6          rh ← a.get(0);
7          libera rh;
8      fim
9      return l;
10 fim
```

Por fim, caso o modelo HealCity identifique que a demanda de atendimento de todos os quartos do ambiente de saúde é muito baixa e que a desalocação de atendentes de alguma sala não prejudicará o todo, o modelo deve identificar quais atendentes foram alocados fora do horário normal de trabalho e desalocá-los para reduzir os custos financeiros do ambiente

de saúde. O Algoritmo 4 apresenta o pseudo-código proposto para a desalocação de recursos humanos com complexidade $O(n)$. Para evitar os efeitos da histerese, da mesma forma que ocorre na realocação, a desalocação também é protegida pela estratégia de *cooldown period*. Além disso, outro ponto importante é que se um determinado recurso humano for desalocado, ele não poderá mais ser alocado no mesmo turno de trabalho.

4.4.3 Elasticidade Proativa a Nível Regional

A nível regional, o modelo HealCity precisa testar diferentes alocações para os atendentes para garantir que todos as salas de todos os ambientes de saúde identificados na etapa anterior (nível de ambiente de saúde) tenham atendentes suficientes, para minimizar a superlotação. Aqui, o modelo segue a mesma ideia do nível anterior, mas considerando a possibilidade de movimentar profissionais de saúde entre ambientes de saúde diferentes e também alocar novos profissionais se necessário. Para este processo, foi proposta uma nova definição para o custo de alocação ou realocação de pessoas. Este custo é o tempo estimado para realocar um recurso humano entre dois ambientes de saúde ou para alocar um novo atendente a um ambiente de saúde específico. Foi definido esse custo como a distância mensurada em tempo para movimentação fornecida em tempo real pela *API Distance Matrix* do Google Cloud (Google Maps Platform, 2024) somando o tempo estimado para o processo de alocação interna. A *API Distance Matrix* possui um método que retorna o tempo de viagem entre dois pontos (partida e chegada), considerando as condições atuais do trânsito, expresso em segundos. Portanto, novos formalismos matemáticos são necessários para expressar a função de custo proposta.

Seja $ATV(r)$ o vetor de tempo de alocação (*Allocation Time Vector*) da sala r durante o intervalo de tempo entre o primeiro instante de tempo do sistema (0) e o instante de tempo atual. Usando esta função e a $size(x)$ acima mencionada, podemos expressar o tempo estimado de alocação na sala r usando a Equação 10, onde $x[] = ATV(r)$ é o vetor de tempo de alocação para a sala r, e $ADT(x[j])$ representa o tempo de duração de alocação para o i-ésimo elemento do vetor x.

$$EAT(r) = \frac{1}{size(x)} \sum_{i=0}^{size(x)-1} ADT(x[j]) \tag{10}$$

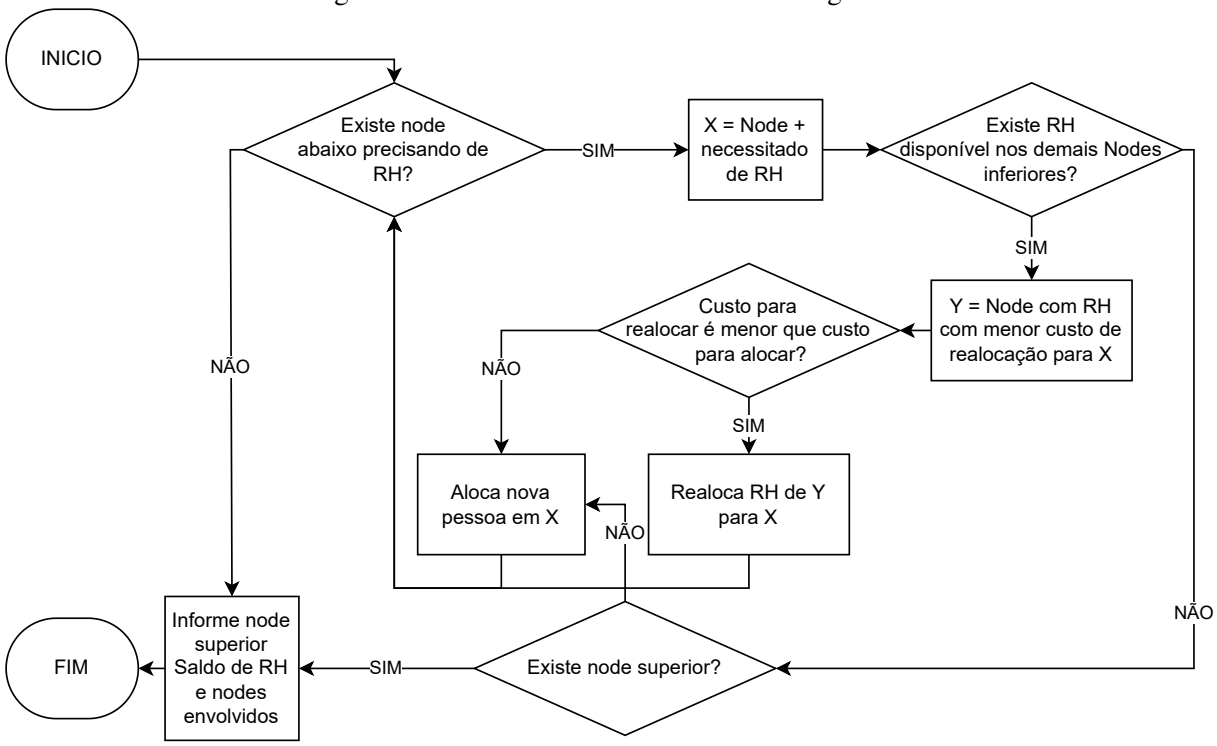
$L(x)$ é uma função que retorna a localização geográfica de um determinado ambiente de saúde ou recurso humano x. Usando esta função, a Equação 11 representa o custo de alocação ou realocação:

$$Cost(o, d) = DistanceMatrixAPI(L(o), L(d)) + EAT(d) \tag{11}$$

onde, o representa o local de origem, que pode ser um ambiente de saúde, no caso de realocação, ou a residência do recurso humano, no caso de uma alocação, e d representa o local de destino

que necessita de recurso humano.

Figura 31 – Elasticidade Proativa a Nível Regional.



Fonte: Elaborada pelo autor.

A Figura 31 apresenta o fluxograma da elasticidade proativa à nível regional. Para realizar a alocação e/ou realocação de recursos humanos a nível regional, cada nó executa as seguintes etapas: verificar se os ambientes de saúde nos nós abaixo requerem recursos humanos, identificar o nó com necessidade mais significativa e verificar se há recursos humanos disponíveis nos nós inferiores. Se disponível, o modelo seleciona o nó com menor custo de realocação e verifica se o custo de realocação é inferior ao custo de realocação. Se o custo de realocação for menor, o modelo realiza a realocação. Se o custo a alocar for menor, então o modelo aloca um novo recurso humano. Se não houver recursos humanos disponíveis nos nós inferiores, o algoritmo verifica se há um nó acima, passando a decisão para o próximo nível, se existir. Caso contrário, o modelo aloca um novo recurso humano. O gerenciador de elasticidade, em cada nível de nó, monitora constantemente a carga de trabalho e os recursos do ambiente de saúde, incluindo a disponibilidade de atendentes com as especialidades exigidas. Sempre que uma sala necessita de recursos humanos adicionais, o gerenciador de elasticidade procura atendentes disponíveis com a especialidade necessária e os realoca (ou aloca) à sala pretendida. Desta forma, o ambiente de saúde pode alocar eficientemente os seus recursos e minimizar o tempo de espera dos pacientes. Assim, quando necessário e o custo em tempo for inferior a realocar um profissional, o modelo aloca profissionais de saúde que não estão em nenhum ambiente de saúde, mas estão disponíveis para alocação.

Em todas as possibilidades, alocação, realocação ou desalocação, o HealCity deve gerar

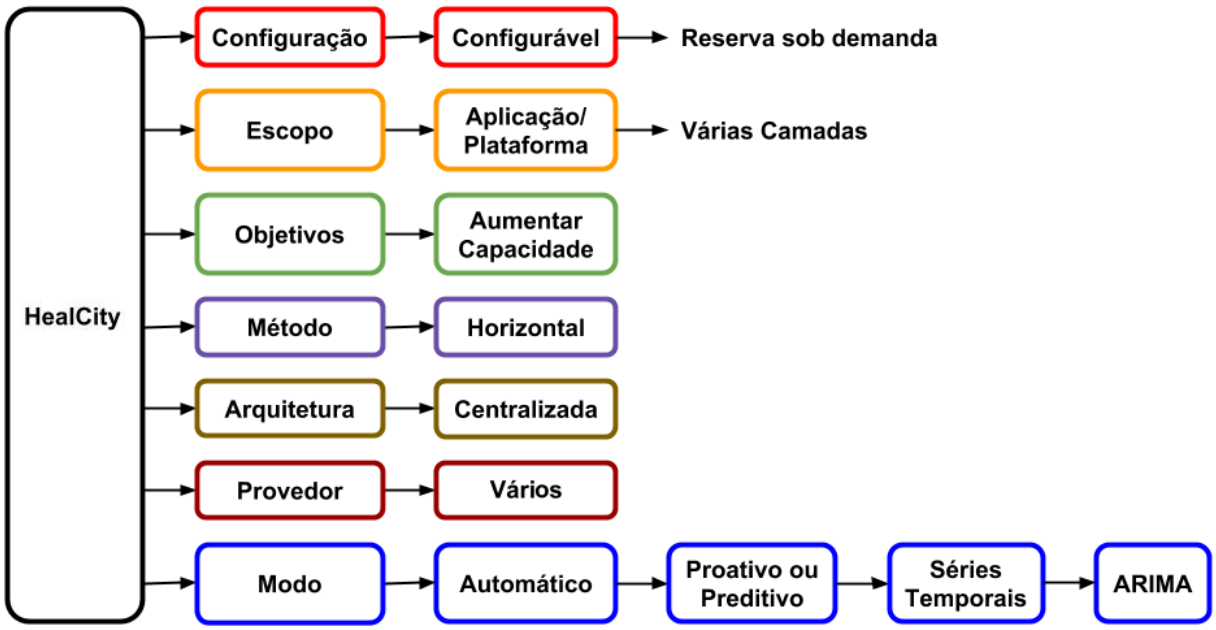
uma notificação e encaminhar ao recurso humano para que o mesmo possa realizar a ação escolhida pelo algoritmo. As notificações de elasticidade de recursos humanos devem estar todas disponíveis para os gestores de cada ambiente de saúde e da cidade inteligente visualizarem na hora em que elas ocorrem. Isso se torna necessário, pois situações pontuais podem exigir que os gestores sejam contra uma determinada escolha do algoritmo, conforme a necessidade da cidade inteligente naquele momento específico. Além disso, o algoritmo apenas envia as notificações e monitora a movimentação dos atendentes, não sendo responsável pela efetiva movimentação dos profissionais de saúde, que podem decidir acatar a movimentação ou não. Nos casos em que um determinado atendente não realizar a movimentação indicada pelo algoritmo, o HealCity deve notificar os gestores tanto dos ambientes de saúde envolvidos, quanto da cidade inteligente, para que os mesmos sejam capazes de entrar em contato com o atendente e avaliarem a sua recusa em seguir o que o algoritmo indicou. Com base em todas as descrições, algoritmos, fluxogramas e formalismos apresentados até o momento, o modelo HealCity introduz na literatura o conceito de **Elasticidade de Recursos Humanos para Saúde em Cidades Inteligentes**, que pode ser definida da seguinte forma:

> **Definição 1: Elasticidade de Recursos Humanos para Saúde em Cidades Inteligentes** é uma extensão do conceito de elasticidade de recursos da computação em nuvem (AL-DHURAIBI et al., 2017) para alocar dinamicamente profissionais de saúde com base na demanda dos pacientes em cidades inteligentes. A Elasticidade de Recursos Humanos para Saúde em Cidades Inteligentes utiliza sensores para acompanhar a demanda dos pacientes *dentro* e *fora* dos ambientes de saúde e, com base em uma abordagem de elasticidade proativa, sugere uma distribuição estratégica do corpo médico, movimentando os profissionais para as áreas com maior necessidade em toda a cidade inteligente, sempre levando em consideração a qualidade dos serviços oferecidos atualmente pelos ambientes de saúde.

Conforme proposto por Al-Dhuraibi et al. (2017), e detalhado na Seção 2.5, sistemas elásticos podem ser classificados de várias formas, dependendo de seus atributos, dentre eles: (i) Configuração, (ii) Escopo, (iii) Objetivo, (iv) Método, (v) Arquitetura, (vi) Provedor e (vii) Modo de Operação. A Figura 32 sumariza as classificações do modelo HealCity para os atributos apresentados. Com relação ao atributo Configuração, o modelo HealCity propõe um sistema elástico configurável, uma vez que o gerenciador de elasticidade deve ser capaz de solicitar atendentes que tenham as características (ou configurações) necessários ao atendimento da demanda de uma sala em específico. Se o sistema fosse rígido, só haveria um tipo de atendente disponível, mas no modelo proposto podem existir vários. Ainda nesse atributo, com relação ao método de reserva, o HealCity faz uso da reserva sob demanda, uma vez que assim que o modelo identificar a necessidade de novos atendentes, o recurso humano é "reservado" imediatamente, ou a reserva é rejeitada caso não existam profissionais disponíveis para alocação.

Já para o atributo Escopo, o modelo HealCity propõe um sistema de elasticidade embarcada

Figura 32 – Classificações do modelo HealCity para os atributos propostos por Al-Dhuraibi et al. (2017).



Fonte: Elaborada pelo autor.

para aplicação ou plataforma. Essa é a melhor classificação, pois segundo Al-Dhuraibi et al. (2017) na elasticidade embarcada, as aplicações elásticas são capazes de ajustar seus próprios recursos de acordo com os requisitos de tempo ou devido a mudanças no fluxo de execução, tal qual o modelo HealCity propõe. Além disso, o modelo faz elasticidade em várias camadas, pois toma decisões de elasticidade em três níveis, de sala, de ambiente de saúde e de cidade. Para o atributo Objetivo, o modelo HealCity apresenta um sistema elástico para aumentar capacidade, uma vez que o principal objetivo do HealCity é justamente o aumento da capacidade de atendimento da cidade inteligente sempre que a demanda de pacientes exigir. Com relação ao Método, o modelo HealCity propõe um sistema de elasticidade horizontal apenas. Isso ocorre pois o HealCity trabalha com recursos humanos que já são funcionários dos ambientes de saúde da cidade inteligente, sendo impossível para o modelo propor alterações na capacidade de atendimento individual de cada profissional de saúde. Dessa forma, a única ação de elasticidade possível ao HealCity é adicionar ou remover recursos humanos horizontalmente para atender a demanda.

No que se refere a Arquitetura, o modelo HealCity apresenta um gerenciador elástico centralizado, uma vez que o módulo Elastic é o único responsável por tomar as decisões de elasticidade. Quanto ao Provedor, o HealCity propõe vários provedores. No modelo proposto, estão sendo considerados para a alocação funcionários que trabalham para vários ambientes de saúde, dessa forma existem vários provedores de recursos. Por fim, com relação ao Modo de Operação, o modelo HealCity propõe um sistema elástico automático proativo e preditivo, através de uso de análise de séries temporais por meio do modelo ARIMA. O HealCity é considerado proativo, pois as decisões de elasticidade são tomadas antes que o modelo atinja os *thresholds*

pré-definidos, e não após como na elasticidade reativa. Ao mesmo tempo o modelo é preditivo, pois faz uso de técnicas de predição computacional para prever a demanda de pacientes e propor as ações de elasticidade baseada na demanda prevista.

4.5 Gerenciamento de futuros pacientes na cidade inteligente

O modelo tem como diferencial e missão considerar pessoas com sinais vitais alterados fora dos ambientes de saúde ao planejar a capacidade de atendimento médico da cidade inteligente. Assim, além do modelo propor um ajuste automático de profissionais com base na demanda esperada de cada sala médica, considerando o fluxo de pacientes dentro dos ambientes de saúde, o sistema também avisa as pessoas que precisam de atendimento médico para se deslocarem a esses ambientes e as leva em consideração quando propõe os ajustes automáticos na estrutura de recursos humanos. Portanto, o HealCity assume que sempre que receber um paciente com sinais vitais alterados, com indicação de intervenção médica, pelo VitalSense (RODRIGUES et al., 2023), deve considerar a pessoa como um possível paciente e incluí-lo em uma espécie de fila virtual para buscar a melhor unidade de tratamento para ele. A fila virtual deve sempre levar em consideração o risco envolvido nos sinais vitais alterados. Pacientes com maior risco devem ser avaliados e direcionados antes dos pacientes com menor risco, para garantir um atendimento mais ágil aos mesmos. Além disso, dependendo do risco, ações adicionais podem ser tomadas. Assim, sempre que houver alguém nesta fila virtual, o modelo executa o Algoritmo 5 para definir o ambiente de saúde apropriado para o paciente com complexidade $O(n)$.

Primeiro, o modelo deve identificar a localização geográfica do paciente e o tempo necessário para o paciente se deslocar até cada um dos ambientes de saúde presentes no algoritmo. Para isso, o algoritmo utiliza a localização do paciente como ponto de origem e o endereço do ambiente de saúde como ponto de destino e faz uma consulta à *API Distance Matrix* para encontrar o tempo de viagem. A API permite calcular o menor tempo ou o menor caminho. Como em determinados locais, a distância física, mesmo sendo menor, pode ter um custo de tempo muito maior, devido às regras de trânsito ou ao fluxo de pessoas, o modelo faz do uso do menor tempo, tal como é utilizado nas estratégias de alocação e realocação de recursos humanos. Assim, calculados os tempos de distância estimados, o algoritmo ordena os ambientes de saúde aptos a atender o paciente na lista do mais próximo para o mais distante. Por ambiente de saúde apto a atender, entende-se que o mesmo possua profissionais de saúde e equipamentos adequados a tratar o problema potencial identificado no paciente. Isso ocorre pois, por exemplo, se os sinais vitais indicam um possível infarto, o paciente não pode ir para um ambiente de saúde incapaz de prestar atendimento cardiológico. A princípio, a abordagem mais interessante para o paciente seria ir ao ambiente de saúde com o menor tempo de deslocamento possível, mas nem sempre essa pode ser a melhor escolha. Essa situação é ilustrada na Figura 33, onde mesmo que existam ambiente de saúde mais próximos, o mais indicado para o paciente é aquele que está mais distante. Imagine a situação em que temos um surto viral com foco em um determinado

Algoritmo 5: Seleciona o ambiente de saúde apropriado para o paciente

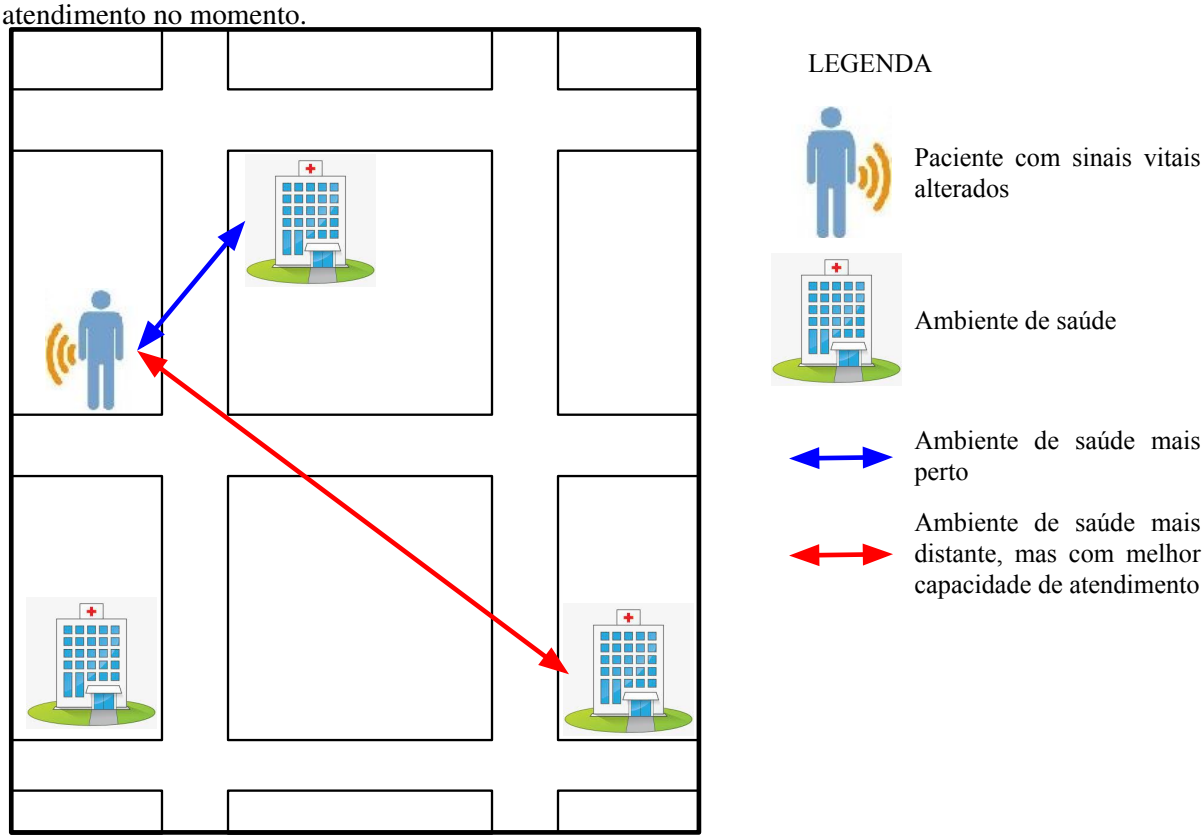
Entrada: Lista de ambientes de saúde c aptos a atender o paciente p, paciente p
Saída: Paciente p com a informação de ambiente de saúde destino

```
1  início
2      o ← localização do paciente p;
3      l ← novo vetor de ambientes de saúde e tempo estimado de deslocamento;
4      a ← novo vetor de ambientes de saúde e tempo estimado de deslocamento;
5      para cada Ambiente de saúde h na lista de ambientes de saúde c faça
6          d ← localização do ambientes de saúde h;
7          t ← tempo estimado de deslocamento usando Distance Matrix API do Google Cloud usando o,
             d como Entrada;
8          se h possui atendentes alocados então
9              a.addAll(t, h);
10         senão
11             l.addAll(t, h);
12         fim
13     fim
14     ordena l, menor tempo estimado de deslocamento;
15     para cada Ambiente de saúde h no vetor de ambientes de saúde e tempo estimado de
         deslocamento l faça
16         se tempo de triagem em h está adequado então
17             p.setAmbienteDeSaude(h);
18             retorna p;
19         fim
20     fim
21     para cada Ambiente de saúde h no vetor de ambientes de saúde e tempo estimado de
         deslocamento l faça
22         se sala de triagem de h pode alocar mais atendentes então
23             p.setAmbienteDeSaude(h);
24             retorna p;
25         fim
26     fim
27     para cada Ambiente de saúde h no vetor de ambientes de saúde e tempo estimado de
         deslocamento a faça
28         se tempo de triagem em h está adequado então
29             p.setAmbienteDeSaude(h);
30             retorna p;
31         fim
32     fim
33     para cada Ambiente de saúde h no vetor de ambientes de saúde e tempo estimado de
         deslocamento a faça
34         se sala de triagem de h pode alocar mais atendentes então
35             p.setAmbienteDeSaude(h);
36             retorna p;
37         fim
38     fim
39     Ordena l, menor número de pacientes;
40     p.setAmbienteDeSaude(l.getFirst());
41     retorna p;
42 fim
```

bairro da cidade inteligente. Nessa situação hipotética, todos os pacientes doentes estarão mais próximos do mesmo ambiente de saúde. O resultado é que se todos se deslocarem para este ponto de atendimento, o mesmo ficará sobrecarregado enquanto os demais terão menos carga.

Nessa situação, poderia-se dizer que o modelo deve deslocar atendentes de outros ambientes de saúde para este, mas nem sempre isso é possível devido às próprias restrições dos ambientes de saúde. Às vezes, embora tenhamos cinco técnicos de raio-x disponíveis para alocar em uma unidade médica, se ele tiver apenas duas máquinas de raio-x, o número máximo de pessoas alocadas será de duas. Assim, o modelo tenta identificar o ambiente de saúde mais próximo que já pode atender o paciente, ou que pode ter sua capacidade de atendimento ajustada para isso.

Figura 33 – Exemplo de paciente com sinais vitais alterados, em uma cidade inteligente com três ambientes de saúde disponíveis. Mesmo que haja ambientes de saúde mais próximos, o mais indicado para o paciente é o mais distante, visto que é o ambiente de saúde que possui a melhor capacidade de atendimento no momento.



Fonte: Elaborada pelo autor.

Em um primeiro momento, o algoritmo analisa se determinado ambiente de saúde alocou funcionários, isto é, se o ambiente de saúde possui funcionários em atendimento que estão trabalhando fora de seu turno habitual de trabalho. Se o ambiente de saúde em questão tem casos assim ele é descartado. Essa decisão ocorre porque o algoritmo só aloca funcionários quando não é possível realocá-los de qualquer lugar. Assim, o ambiente de saúde já está acima da capacidade e não é o mais adequado. Além disso, o algoritmo quer desalocar esse recurso humano o mais rápido possível, pois tem um custo maior para o ambiente de saúde. Caso o ambiente de saúde não tenha ninguém alocado, é analisado se o tempo da fila de triagem está adequado ao *threshold* superior definido pelo gesto do ambiente de saúde. Se o tempo estiver adequado, é indicado este ambiente de saúde para o paciente. Caso contrário, o algoritmo passa

a analisar o próximo ambiente de saúde e assim sucessivamente. Se forem percorridos todos os ambientes de saúde que não têm ninguém alocado e todos eles têm tempo de triagem ruim, o algoritmo retorna ao início da lista (ou seja, o ambiente de saúde mais próximo em termos de tempo estimado de deslocamento) e é verificado se esse ambiente de saúde ainda pode receber mais recursos humanos para triagem. Se sim, é indicado este ambiente de saúde para o paciente. Se não, o algoritmo segue para o próximo ambiente de saúde. Se percorridos novamente todos os ambientes de saúde, são realizados os mesmos procedimentos considerando os ambientes de saúde que têm pessoas alocadas. Caso ainda seja impossível indicar ao paciente algum ambiente de saúde após percorrer toda a lista, é selecionado o ambiente de saúde com o menor número de pacientes aguardando para indicar ao paciente.

Uma vez identificado o ambiente de saúde adequado, o sistema deve alertar o paciente informando que ele deve ir para o ambiente de saúde determinado via notificação por SMS ou APP, solicitando uma resposta se ele aceitará ou não a sugestão de ir ao atendimento médico. Se o paciente responder SIM, o sistema possui a garantia que o paciente irá para o ambiente de saúde escolhido em breve. Isso garante a chegada do paciente. Caso o paciente responda não, o sistema já sabe de antemão que o paciente não virá ao ambiente de saúde, e mesmo desconsiderando o mesmo, continuará monitorando o risco do paciente. Se o paciente não responder, o modelo usa uma equação de confiança para estimar se o paciente irá ou não procurar atendimento.

A função $PNIP(r, t_i, t_f)$ define o número previsto de pacientes chegando em uma sala de ambiente de saúde r entre os tempos t_i e t_f. Ressalta-se que a função $PNIP(r, t_i, t_f)$ considera apenas os pacientes que não responderam à notificação, ou seja, aqueles que o algoritmo está antecipando a chegada, mas que não confirmaram. Assim, a confiança de chegada do paciente na sala do ambiente de saúde r entre os tempos t_i e t_f é definida pela Equação (12), onde $NIP(r, t_n)$ refere-se ao número de pacientes que não responderam a notificação e mesmo assim foram para a sala do ambiente de saúde r no instante de tempo t_n.

$$C(r, t_i, t_f) = \frac{1}{PNIP(r, t_i, t_f)} \sum_{t_n=t_i+1}^{t_f-1} NIP(r, t_n) \tag{12}$$

O número previsto de pacientes chegando, ou seja, aqueles que confirmaram que estão indo ao ambiente de saúde, em uma sala r entre os tempos t_i e t_f é definido pela função $FNIP(r, f_i, f_f)$. O possível número futuro de pacientes esperando para atendimento na sala r do ambiente de saúde entre os tempos futuros f_i e f_f é definido pela Equação 13, onde são considerados os pacientes que confirmaram que estão indo ao ambiente de saúde, através da equação $FNIP(r, f_i, f_f)$, e as estimativas para aqueles pacientes que não se tem certeza se estão indo ou não, através da função $PNIP(r, f_i, f_f)$ e da equação de confiança proposta $C(r, t_i, t_f$. Para a equação de confiança, é utilizado como instante de tempo inicial t_i o primeiro instante de tempo de sistema (0), e como instante de tempo final t_f, o instante de tempo atual (t_a). Assim, a equação proposta é usada para calcular a função $ENP(r, f_i, f_f)'$ para a sala de

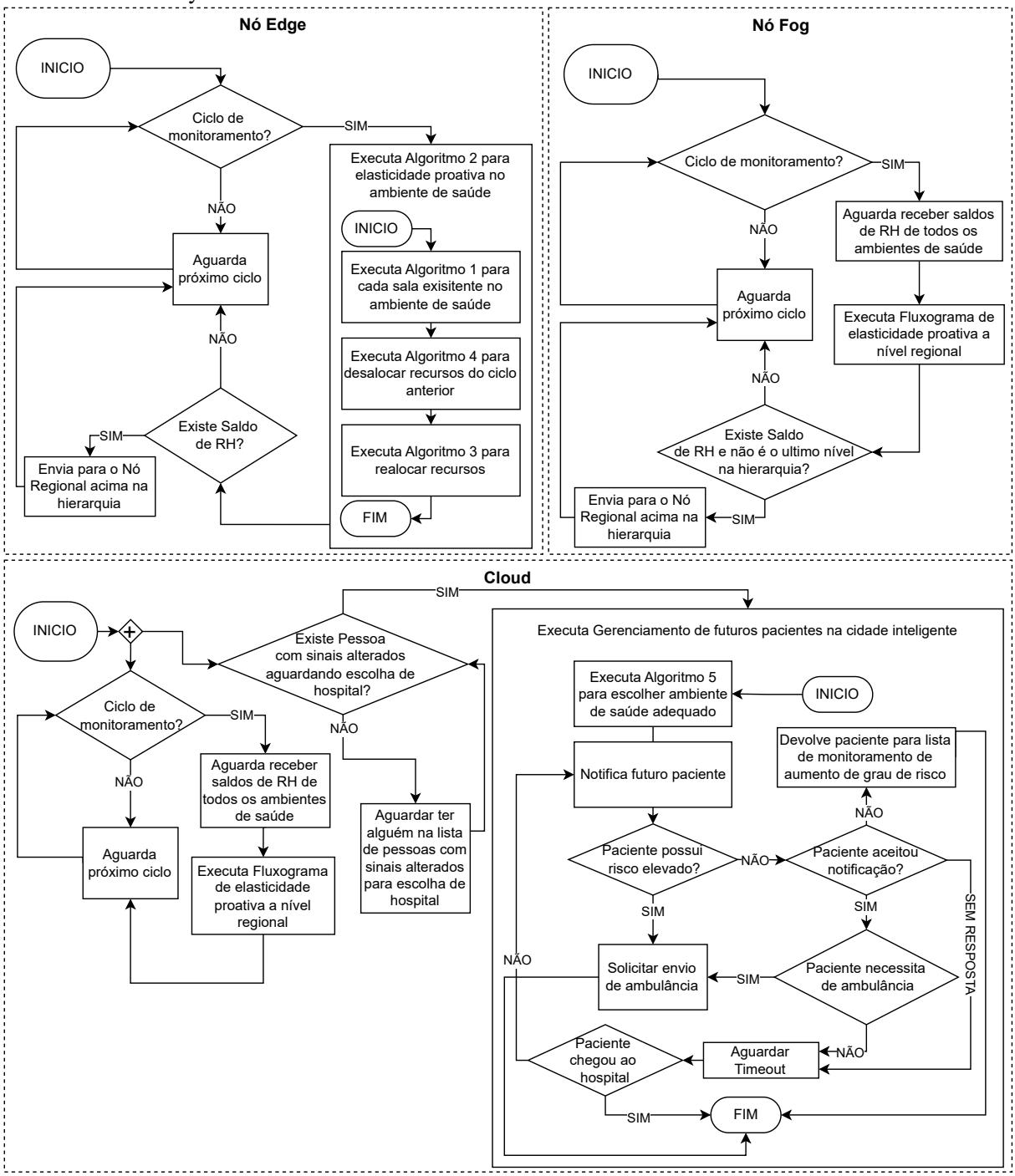
triagem, e assim é possível calcular o $ECT(r, f_i, f_f)' = ACT(r, f_i, f_f)' \cdot ENP(r, f_i, f_f)'$, onde $ACT(r, f_i, f_f)'$ são as previsões do tempo médio de atendimento e $ENP(r, f_i, f_f)'$ é calculado usando o algoritmo proposto.

$$NIP(r, f_i, f_f)' = FNIP(r, f_i, f_f) + (C(r, 0, t_a) \cdot PNIP(r, f_i, f_f)) \tag{13}$$

Além disso, além do modelo realizar as previsões e ajustes de recursos humanos visando melhoria no tempo de espera por atendimento, o HealCity precisa se preocupar também com a saúde dos pacientes. Dessa forma apenas gerar notificações pode não ser o suficiente dependendo da situação e do risco envolvido. Assim em casos que os sinais vitais tenham identificado que o risco de saúde do paciente seja elevado, o modelo envia esse alerta diretamente ao ambiente de saúde, solicitando o encaminhamento de uma ambulância diretamente ao paciente, sem esperar qualquer resposta da notificação. Um outro ponto é que em casos de riscos menos elevados o HealCity parte do princípio que o paciente pode se deslocar até o ambiente de saúde. No entanto, existem situações em o paciente não pode fazer tal deslocamento (devido a diversas razões, desde trânsito, ausência de mobilidade, entre outros). Neste caso, junto à notificação solicitando que o mesmo se dirija a um ambiente de saúde o modelo deve questionar a mobilidade do paciente, e caso o mesmo informe que aceita a notificação de atendimento, porém não possui meios de transporte, o HealCity deve enviar esse alerta diretamente ao ambiente de saúde, solicitando o encaminhamento de uma ambulância ao paciente. Por fim, além de encaminhar as notificações, o modelo deve identificar se os pacientes, aos quais que não houve solicitação de ambulância, chegaram efetivamente ao hospital. Conforme já apontado anteriormente, para os casos em que a pessoa responder negativamente a notificação, o modelo vai monitorar o risco e retornar a notificar se necessário. Para os casos em que o paciente responder positivamente, ou não responder, o HealCity aplica uma política de *timeout*. Assim, caso o paciente não chegue no tempo previsto após o envio da notificação, uma nova notificação deve ser disparada para o mesmo paciente, visando alertá-lo novamente do seu risco de saúde. É importante salientar que o modelo HealCity recebe esta indicação de sinais vitais alterados através do modelo VitalSense, uma vez que o modelo já possui mecanismos para lidar com falsos positivos, eliminando essa tarefa na abordagem do HealCity.

A Figura 34 apresenta o fluxograma de interações entre os diferentes algoritmos propostos nas diferentes camadas da hierarquia do modelo HealCity. Assim, com base nos algoritmos e equações apresentadas pelo modelo HealCity, bem como os fluxogramas propostos, é possível de fato implementar o cenário hipotético apresentado na Figura 16, onde existe uma estratégia de ajuste de recursos humanos e a indicação de ambiente de saúde adequado para um novo paciente. Assim como proposto na Seção 4.1, o HealCity criar para os cidadãos a impressão de uma *nuvem de recursos humanos*, onde a quantidade de atendentes pode ser provisionada conforme a necessidade.

Figura 34 – Fluxograma de interações entre os algoritmos propostos nas diferentes camadas da hierarquia do modelo HealCity.



Fonte: Elaborada pelo autor.

4.6 Considerações Parciais

Este capítulo apresentou o modelo HealCity, um modelo multinível de elasticidade para gerenciamento de recursos humanos de ambientes de saúde em uma cidade inteligente. Diferente dos trabalhos relacionados, o modelo propõe o uso da elasticidade para antecipar problemas futuros, fornecendo um modelo de otimização de recursos humanos para melhorar a experiên-

cia do paciente. A elasticidade proposta deve ser transparente aos pacientes e não deve afetar negativamente o atendimento já oferecido aos mesmos. O modelo trabalha com elasticidade em três níveis: (i) a nível de sala, onde é analisada a demanda de cada sala do ambiente de saúde e identificada a necessidade de alocar ou liberar atendentes; (ii) a nível de ambiente de saúde, onde é analisada a necessidade identificada por cada uma das salas de forma a serem propostas ações de realocação ou desalocação de recursos humanos em cada ambiente de saúde; e (iii) a nível regional, onde é analisada a necessidade identificada em cada ambiente de saúde de forma a serem propostas ações de alocação, realocação ou desalocação de recursos humanos entre diferentes ambientes de saúde da cidade inteligente.

Também nesse capítulo estão presentes as principais decisões de projeto, como a proposta de um RTLS com precisão a nível de sala, o uso de *thresholds* baseados em tempo de espera para a elasticidade proativa, e o uso da abordagem de predição baseada em probabilidade e estatística de séries temporais, através do modelo ARIMA. Está presente também uma descrição detalhada da arquitetura proposta, com a descrição de cada um dos componentes do modelo, bem como as suas interligações, e as novas equações e algoritmos propostos para cada um deles. Dentro dos componentes do HealCity, destacam-se os mecanismos para gerenciamento de recursos humanos, onde é proposta a (i) **Elasticidade de Recursos Humanos para Saúde em Cidades Inteligentes**, uma extensão do conceito de elasticidade de recursos em *Cloud Computing* (AL--DHURAIBI et al., 2017) para gerenciamento do uso de recursos humanos em diferente níveis de uma cidade inteligente, de forma que os profissionais são alocados, desalocados e realocados conforme a demanda prevista de pacientes na cidade. Por fim, este capítulo também apresenta as estratégias para gerenciamento de futuros pacientes na cidade inteligente, com a indicação aos mesmos do hospital melhor preparado para atendê-los.

5 METODOLOGIA DE AVALIAÇÃO

Para definir a metodologia de avaliação do modelo HealCity, primeiramente foi observado como foram avaliados os trabalhos relacionados a este. Para o contexto proposto, não foi identificado um *benchmark* único que pudesse ser utilizado para mensurar a performance do modelo e compará-lo com o estado-da-arte. Portanto, para avaliação e validação do modelo, definiu-se o uso de uma simulação computacional dos diversos ambientes de saúde que podem compor uma cidade inteligente. Com o uso de uma simulação, um grande número de dados sintéticos podem ser gerados para treinar o modelo de predição, e alimentar a aplicação desenvolvida. Além disso, é uma estratégia amplamente utilizada na literatura, sendo usada por Oueida, Aloqaily e Ionescu (2019), Apornak et al. (2021a), Apornak et al. (2021b), Hafezalkotob et al. (2022), Liao, Chu e Ho (2022), Duma e Aringhieri (2023), Lazebnik (2023) e Rasidi et al. (2024). Cargas de trabalho sintéticas foram escolhidos em função da indisponibilidade de dados reais. Além disso, de acordo com Malti, Hakem e Benmammar (2023), cargas de trabalho sintéticas são amplamente utilizadas na comunidade acadêmica para avaliar a eficácia das abordagens de elasticidade em nuvens computacionais. Nas Seções abaixo, serão detalhados os cenários simulados, os parâmetros utilizados, a infraestrutura dos ambientes de saúde inteligentes, quais as métricas de avaliação utilizadas, as cargas de trabalho geradas e os testes do sistema RTLS.

5.1 Implementação e Ambiente de Testes

Para a realização dos experimentos nos ambientes de saúde, foi desenvolvido um protótipo em Java e em Python do sistema HealCity. Esse protótipo contemplou o desenvolvimento do gerenciador de elasticidade Elastic, juntamente com o sistema de predição Predict. Além destes, foi desenvolvido um simulador para os ambientes de saúde da cidade inteligente.

5.1.1 Protótipo desenvolvido

Para realização da avaliação do modelo HealCity, foi implementado um protótipo do gerenciador de elasticidade Elastic a nível de ambiente de saúde e um protótipo do gerenciador de elasticidade Elastic a nível regional, ambos utilizando a linguagem Java. Esse protótipo consiste em um gerenciador que realiza as operações de monitoramento de recursos humanos, tempo de espera e quantidade de pacientes aguardando atendimento para realizar elasticidade de recursos humanos em um ambiente de saúde inteligente. O monitoramento do ambiente é realizado através de consultas ao banco de dados da aplicação, onde estão armazenadas informações referente a movimentação de pessoas entre as salas do ambiente de saúde, bem como os tempos de permanência em cada um dos setores. Para o controle do ambiente, o modelo envia notificações de elasticidade aos recursos humanos, informando para qual sala os mesmos devem se deslocar. Apesar do gerenciador de elasticidade armazenar o resultado da movimentação dos funcioná-

rios, no ciclo seguinte de monitoramento, o mesmo faz novamente todas as consultas ao banco de dados para verificar o estado real do sistema. Isso faz-se necessário pois em um ambiente real se um recurso humano não cumprir a notificação de elasticidade, o resultado de elasticidade do ciclo anterior não vai condizer com a realidade do sistema. Para tomar as decisões de elasticidade, o modelo consulta a cada ciclo de monitoramento os limites de tempo estabelecidos individualmente para cada sala do ambiente de saúde. Dessa forma, caso seja interessante para os gestores de um ambiente de saúde, podem ser estabelecidos limites diferentes para salas com características distintas e os limites podem ser alterados dinamicamente.

Para o gerenciador Elastic realizar o controle elástico dos recursos humanos, ele necessita de uma predição das chegadas futuras de pacientes, bem como do tempo de atendimento em cada ambiente. Para realizar esta predição, foi implementado um protótipo do módulo Predict utilizando a linguagem Java e Python. Esse protótipo consiste em uma aplicação em Java responsável por identificar as séries temporais que representam a realidade do ambiente de saúde e uma aplicação em Python responsável por aplicar o modelo ARIMA sobre estas séries temporais de forma a prever os valores futuros para as séries. Para identificar as séries temporais, a aplicação faz consultas ao banco de dados, e agrupa as informações identificadas em intervalos de tempo equivalentes aos ciclos de monitoramento. A cada novo ciclo, o sistema Predict, faz uma atualização nas séries temporais e reaplica o modelo ARIMA. Para este protótipo, foi utilizada uma implementação do Auto ARIMA, através da função $auto_arima$ da biblioteca $pmdarima$ (PYTHON SOFTWARE FOUNDATION, 2024). O Auto-ARIMA é um algoritmo automatizado utilizado para encontrar os parâmetros ideais do modelo ARIMA. O algoritmo Auto-ARIMA pesquisa automaticamente por diferentes combinações de parâmetros do modelo ARIMA para encontrar o melhor modelo que se ajuste aos dados (JAIN; PATEL, 2024). No modelo proposto, o HealCity realiza a previsão dos próximos dois ciclos de monitoramento a cada ciclo analisado.

5.1.2 Simulador do ambiente de saúde

Partindo-se do modelo proposto, foi desenvolvida uma implementação capaz de realizar a simulação da cidade inteligente proposta e armazenar algumas estatísticas após sua execução, para serem utilizadas para análise dos resultados. Para a implementação do simulador foi utilizada a linguagem Java. Para a simulação das filas do ambiente de saúde, optou-se pelo uso de um relógio com incrementos discretos de dez segundos. A cada avanço no relógio da simulação, o simulador verifica os pacientes que estão em atendimento e aqueles que devem sair do atendimento. A cada ciclo de monitoramento, deve ser verificada a chegada de novos pacientes. Para definir o tempo de atendimento em cada um dos setores, e a movimentação dos pacientes ao longo do ambiente de saúde, o simulador utiliza a classe *StdRandom* desenvolvida por Sedgewick e Wayne (2022), disponibilizada pela Universidade de Princeton. A classe *StdRandom* é responsável por gerar números aleatórios para as mais diversas distribuições de probabilidade.

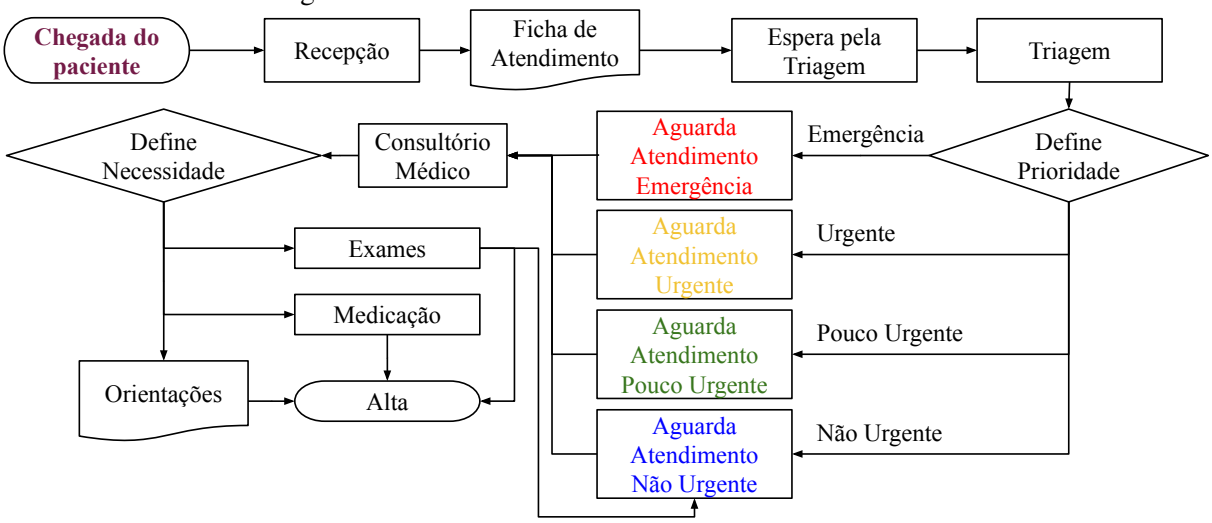
5.1.3 Ambiente de Simulação

O ambiente utilizado para a simulação de saúde e para os testes do protótipo desenvolvido é um equipamento com as seguintes características: Processador Intel Core I9 10900f com dez núcleos físicos de 5.2 GHz, Memória de 64GB com frequência de 3200 MHz, SSD NVMe de 512 GB ADATA SX8100NP, Sistema operacional Windows 10 Pro, Servidor MySQL 5.7.23, Java Virtual Machine versão Java 8 Update 333 e Python versão 3.11.

5.2 Parâmetros

Para realizar a simulação dos ambientes de saúde, optou-se pelo uso dos dados levantados na pesquisa de Capocci et al. (2016) realizada em um ambiente de saúde localizado na cidade de Guarulhos no estado de São Paulo com foco em reduzir o tempo de espera na unidade. O fluxo de atendimento nessa unidade de saúde pode ser visto na Figura 35. De acordo com Capocci et al. (2016), todos os pacientes ao entrarem na unidade passam primeiramente por uma recepção, onde é elaborada a ficha de atendimento. Após a elaboração da ficha, os pacientes são encaminhados para a espera pela triagem. No procedimento de triagem, os pacientes são examinados por enfermeiros e classificados em prioridades conforme a urgência do problema de saúde e são encaminhados para a espera pelo atendimento médico. O médico que atender o paciente pode encaminhá-lo para a área de medicação, para exames laboratoriais, ou dar alta. A área de medicação conta com atendimentos diferentes para aplicação intravenosa, intramuscular e por inalação. Na área de exames laboratoriais, o ambiente de saúde possui raio-x, eletrocardiograma, ou exames de coleta (exame de sangue, por exemplo). Após os exames ou medicação, o paciente pode receber alta pela equipe de enfermagem ou retornar ao atendimento médico.

Figura 35 – Fluxo de atendimento do ambiente de saúde.



Fonte: Elaborada pelo autor com base em Capocci et al. (2016).

Ainda de acordo com Capocci et al. (2016), o tempo de atendimento em cada uma das

Tabela 5 – Distribuições triangulares de probabilidade para tempos de atendimento.

Atendimento	**Tempo de Atendimento**		
	Mínimo	**Médio**	**Máximo**
Entrada no Ambiente de Saúde			
Elaboração da Ficha de Atendimento	2 minutos	3 minutos	5 minutos
Processo de Triagem	5 minutos	8 minutos	10 minutos
Consultório Médico			
Primeiro atendimento no consultório médico	5 minutos	11 minutos	16 minutos
Atendimento de retorno ao consultório médico	4 minutos	7 minutos	10 minutos
Exames Laboratoriais			
Exames de Coleta	6 minutos	8 minutos	13 minutos
Raio-x	10 minutos	15 minutos	23 minutos
Eletrocardiograma	30 minutos	45 minutos	60 minutos
Medicação			
Medicação intramuscular	3 minutos	3,5 minutos	5 minutos
Preparação para a medicação intravenosa e por inalação	0,5 minuto	1,5 minutos	2,5 minutos
Medicação intravenosa	40 minutos	70 minutos	120 minutos
Medicação por inalação	8 minutos	10 minutos	13 minutos

Fonte: Elaborada pelo autor com base em Capocci et al. (2016).

salas do ambiente de saúde sempre segue uma distribuição triangular, com tempos mínimo e máximo e um tempo médio mais frequente. A Tabela 5 apresenta as distribuições para todos os atendimentos possíveis nessa unidade de saúde, conforme identificado na pesquisa de Capocci et al. (2016). Ainda segundo Capocci et al. (2016), na unidade de saúde estudada pelos autores, os seguintes parâmetros foram identificados com relação a chegada de pacientes:

- Em média são atendidos 190 novos paciente por dia;
- Após a triagem, 5% dos pacientes são classificados como emergência, 12% como urgência, 23% como pouco urgentes e 60% como não urgentes.

Com relação ao primeiro atendimento no consultório médico:

- Em geral, 24% dos pacientes são encaminhados para Raio-x, 37% para exames de coleta, 8% para eletrocardiogramas e 31% não necessitam fazer nenhum exame;
- Apenas 1% dos pacientes não tomam medicação e recebem alta apenas com uma receita, porém 50% dos pacientes necessitam de medicação intravenosa, 30% de medicação intramuscular e 19% de medicação por inalação.

Com relação ao atendimento na área de medicação:

- Medicação intravenosa possui 10 lugares para o processo;
- Medicação por inalação disponibiliza 6 lugares para o procedimento;

Com relação ao atendimento na área de exames:

- O resultado de exames de coleta sai em 4 horas;
- O Raio-x leva meia hora para ficar pronto;
- Após os exames, em média 60% dos pacientes precisam retornar ao atendimento médico os outros 40% recebem alta;
- Os pacientes de retorno são reclassificados nas prioridades pela espera do atendimento médico com 90% como não urgentes, 9% como pouco urgentes e apenas 1% como urgente.

Com relação ao atendimento de retorno no consultório médico:

- Após o atendimento 78% recebem alta, 2% precisam de novos exames e 20% necessitam de medicação no local.

Com relação ao gerenciamento de elasticidade, de acordo com Al-Dhuraibi et al. (2017), é necessária a definição de *thresholds* ou limites para que o sistema reaja disparando as ações para adaptar o sistema à demanda. No contexto do gerenciamento elástico de recursos humanos, optou-se por *thresholds* de tempo de espera em cada um dos pontos de atendimento do ambiente de saúde. Para definir o tempo máximo que um paciente pode aguardar pelo atendimento, em um primeiro momento pensou-se em utilizar os limites definidos no Protocolo de Triagem de Manchester. Entretanto, no ambiente de saúde estudado por Capocci et al. (2016), não é utilizado a classificação *Muito Urgente*, dessa forma, o menor *threshold* para cada ponto de atendimento seria a classificação *Urgente* com 60 minutos de espera. Em função de uma hora de espera parecer ser um tempo considerável para aguardar, optou-se por realizar uma busca na legislação brasileira a respeito do tema. Não encontrou-se nenhuma lei que defina limites de tempo de espera para unidades de saúde, entretanto encontrou-se o Projeto de Lei de 14 de junho de 2018 (GOIANIA, 2018) que propõe um tempo máximo de espera por atendimento em ambientes de saúde, clínicas e laboratórios de 30 minutos em dias normais (de segunda-feira a domingo) e 45 minutos em feriados, véspera de feriados e nos dias subsequentes a eles. Dessa forma foi estabelecido ao HealCity o limite superior de 30 minutos de espera.

Para definir o limite inferior, adotou-se uma estratégia baseada em trabalhos como os propostos por Rostirolla et al. (2017); Righi et al. (2016); Rosa Righi et al. (2016), onde são adotados *thresholds* inferiores para elasticidade entre 30% e 50% da carga máxima do sistema. Esse intervalo de limites inferiores também foi baseado no trabalho de Al-Haidari, Sqalli e Salah (2013), que propõe uma análise teórica com teoria de filas para observar o desempenho de elasticidade em nuvem. Dessa forma, partindo-se do tempo máximo de espera de 30 minutos, 30% seria equivalente a um *threshold* inferior de 9 minutos e 50% seria equivalente a 15 minutos. Dessa forma optou-se por um limite inferior de 9 minutos, pois partiu-se do pressuposto de que

quanto menor a espera de um paciente por atendimento, melhor será a experiência do mesmo com o ambiente de saúde. Assim, a Tabela 6 apesenta os limites de tempo estabelecidos para as ações de elasticidade no modelo HealCity.

Tabela 6 – *Thresholds* de tempo para Gerenciamento de Elasticidade do modelo HealCity.

Thresholds	**Tempo de Atendimento**
Threshold Superior (TS)	30 minutos
Threshold Inferior (TI)	9 minutos

Fonte: Elaborada pelo autor.

Conforme já abordado na Subseção 2.2.2, no Brasil em um ambiente de saúde geralmente é utilizado um modelo de trabalho conhecido como 12x36. Conforme a Lei nº 13.467 (BRASIL, 2017) nesse regime de trabalho um funcionário pode trabalhar por doze horas consecutivas (com direito a uma hora de almoço dentro desse período) e deve descansar por trinta e seis horas seguidas antes de uma nova jornada de doze horas. Nesse regime de trabalho, com quatro pessoas atuando em um mesmo posto de trabalho é possível garantir a atuação vinte e quatro horas, sete dias por semana, de forma que todos trabalhem na mesma proporção dentro do período mensal. Ainda conforme o entendimento da lei, se por qualquer motivo um funcionário precisar trabalhar dentro do seu período de descanso, deve ser tratado como hora extra, a não ser que as horas sejam compensadas em outro momento. Assim, enquanto um recurso humano do ambiente de saúde está em efetivo trabalho, existem outros três funcionários que desempenham a mesma função dentro do seu período de descanso remunerado. Conforme o Decreto-Lei nº 5.452 (BRASIL, 1943), o descanso mínimo entre duas jornadas de trabalho deve ser de onze horas consecutivas. Dessa forma, mesmo que existam horas extras, um funcionário obrigatoriamente precisa descansar onze horas para retornar ao próximo turno de trabalho. Dessa forma, podemos dizer que esses três funcionários em período de descanso não estariam disponíveis para serem alocados a qualquer momento para atendimento, sendo necessário respeitar as seguintes regras para os mesmos estarem disponíveis:

Regra 1: O período mínimo de descanso para o recurso humano estar disponível para alocação é de onze horas;

Regra 2: Um funcionário alocado não está autorizado a trabalhar fora de seu turno regular de trabalho por mais de doze horas. Apesar da legislação trabalhista autorizar o acordo individual para definir questões como essa, o maior período possível de trabalho na legislação é o de doze horas, em função disso, decidiu-se utilizar este limite neste modelo;

Regra 3: O funcionário alocado deve ser desalocado no máximo onze horas antes do seu próximo período normal de trabalho; e

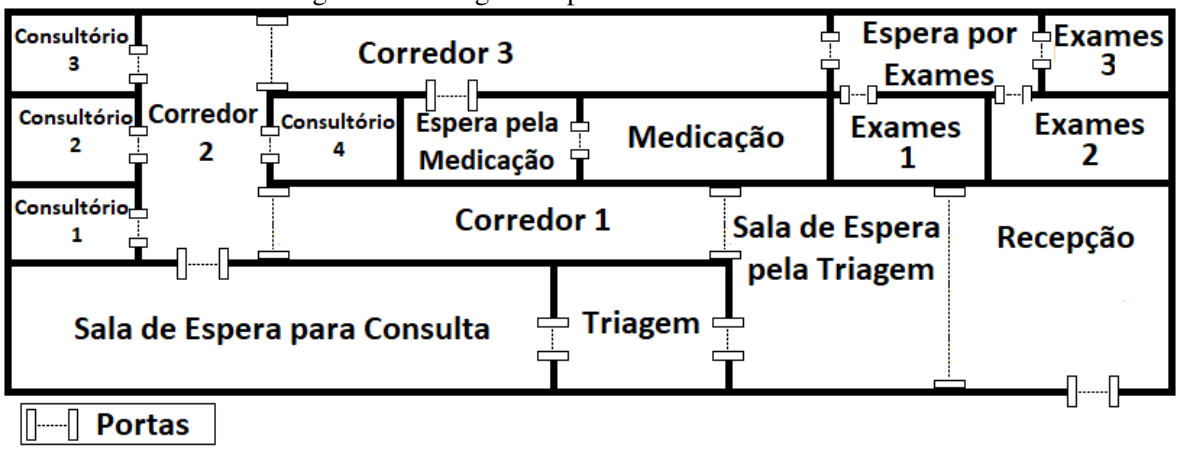
Regra 4: Cada funcionário deve obrigatoriamente cumprir um dos períodos de 36 horas de descanso dentro da mesma semana, a fim de cumprir o Artigo da lei que exige que

todos os trabalhadores possuam um período de 24 horas seguidas de descanso por semana.

5.3 Infraestrutura

De forma a se realizar as simulações da cidade inteligente proposta no modelo HealCity, foram desenvolvidas unidades de atendimento teóricas, usando como base as especificações da unidade estudada por Capocci et al. (2016). Nessas unidades de saúde propostas existem diversos ambientes distintos, onde cada um é caracterizado com um dos possíveis lugares nos quais o sistema estará preparado para identificar pacientes e profissionais de saúde. Na Figura 36 temos a representação visual da planta desse ambiente de saúde. Dessa forma, o ambiente médico proposto possui uma recepção que acomoda os atendentes que são responsáveis pelo primeiro contato com os pacientes e geração do boletim de atendimento. Além disso, existe uma sala de espera onde os pacientes já inclusos no sistema devem aguardar a chamada para a triagem. Logo após, existe uma área de triagem, onde os pacientes devem receber o primeiro atendimento clínico por enfermeiros capacitados para tanto, de forma a se identificar o risco de cada paciente. Após a triagem, existe uma nova sala de espera, onde os pacientes devem aguardar para o atendimento nos consultórios dos médicos. Também existem diversos consultórios, onde os doutores devem estar alocados para atender os pacientes já triados. Ainda, o ambiente de saúde possui uma sala de espera para pacientes aguardando para receber medicações indicadas pelos doutores e a área onde eles devem receber estas medicações e enfermeiros habilitados para aplicá-las. No ambiente estudado por Capocci et al. (2016), e nos ambientes propostos para o HealCity, existem dez vagas para pacientes estarem recebendo medicação intravenosa ao mesmo tempo, seis vagas para medicação por inalação e uma vaga para medicação intravenosa. Além da área de medicação, deve existir uma sala de espera para os exames clínicos que podem ser solicitados pelos doutores, e várias salas onde os exames de raio-x, eletrocardiograma e coleta podem ocorrer.

Figura 36 – Imagem da planta teórica desenvolvida.



Fonte: Elaborada pelo autor.

5.4 Cenários

Para avaliar o modelo HealCity, é necessário comparar uma cidade inteligente simulada com o uso do modelo proposto e sem o uso do modelo para visualizar melhorias no atendimento à população com o uso do HealCity. Portanto, conforme explicado anteriormente, o conjunto de hospitais a serem simulados foi baseado no descrito por Capocci et al. (2016) em sua pesquisa. Embora Capocci et al. (2016) não faça parte dos trabalhos relacionados, uma vez que o estudo focou apenas em realizar simulações do ambiente estudado, os autores identificaram gargalos na unidade de saúde estudada e propuseram um ajuste no serviço, onde sempre que o número de pacientes em espera fosse maior que 2 na sala de triagem, um dos enfermeiros da sala de medicação deveria ir auxiliar o atendente da triagem e após o número retornar a ser igual ou menor que 2, retornar para sua sala de origem, o equivalente a uma realocação entre as duas salas com base no número de pacientes em espera. Portanto, para melhor avaliar o modelo, o HealCity também pode ser comparado com a sugestão proposta por Capocci et al. (2016), já que usam os mesmos parâmetros de entrada como base.

Dessa forma, para avaliação do modelo HealCity, foram definidos três cenários de estudo para análise. Em todos os cenários serão utilizados os mesmos parâmetros de entrada, com os mesmos padrões de carga para chegadas de pacientes, e as mesmas distribuições de probabilidade para tempos médios e limites de atendimento e espera. s distinções entre os cenários estão associadas à aplicação do modelo HealCity e à metodologia proposta por Capocci et al. (2016) e serão delineadas a seguir:

C1: Cidade inteligente sem qualquer abordagem de gestão de recursos humanos: de forma a se ter os dados para comparação, o primeiro cenário de testes se baseia na simulação de uma cidade com um grupo não elástico de quatro ambientes de saúde sem o uso do modelo HealCity. Neste cenário a alocação de atendentes é estática, como ocorre na maioria dos ambientes de saúde.

C2: Cidade inteligente utilizando a abordagem proposta por Capocci et al. (2016): o segundo cenário de testes foca na simulação de uma cidade inteligente com um grupo de quatro ambientes de saúde utilizando o ajuste de recursos humanos proposto por Capocci et al. (2016). Nesse cenário a alocação de atendentes não é mais estática, porém serão utilizados apenas os atendentes já alocados nos ambientes de saúde, sem alocação externa de recursos humanos, ou realocações entre ambientes médicos distintos.

C3: Cidade inteligente com o modelo de elasticidade proposto pelo HealCity: o terceiro cenário de testes é baseado na simulação de uma cidade inteligente com um grupo de quatro ambientes de saúde utilizando o modelo de elasticidade proposto pelo HealCity para gestão de recursos humanos. Neste cenário o modelo faz uso da alocação, realocação e desalocação de pessoas proposta no modelo HealCity, possuindo uma alocação inicial de

atendentes idêntica aos cenários anteriores, entretanto, com base na Elasticidade Multinível de Recursos Humanos em cidades inteligentes proposta neste trabalho, a alocação se torna dinâmica ao longo do dia de atendimento, com os profissionais de saúde sendo ajustados de forma elástica conforme a demanda. Ao contrário do segundo cenário, poderá haver realocação entre os diferentes ambientes de saúde da cidade inteligente, bem como a alocação de novos recursos humanos para atendimento conforme a necessidade.

5.5 *Workload* de Entrada

Para o presente trabalho, será utilizado a alocação de recursos humanos encontrada por Capocci et al. (2016) no ambiente de saúde estudado em sua pesquisa. Nessa unidade clínica o atendimento é prestado 24 horas por dia, sete dias por semana, possuindo uma alocação de 11 recursos humanos, atuando diretamente no atendimento aos pacientes, destes sendo dois médicos, seis enfermeiros, um técnico em radiologia e dois atendentes. Em qualquer horário do dia existe a mesma quantidade de atendentes, sendo mantido o número através de mais de um turno de trabalho. A alocação dos profissionais de saúde nos ambientes é definida da seguinte forma:

- 2 atendentes sem qualificação médica atuando na recepção;
- 1 enfermeiro atuando na triagem de pacientes;
- 2 médicos atuando no atendimento em consultórios;
- 2 profissionais de enfermagem atuando com exames de coleta;
- 2 enfermeiros atuando em toda a área de medicação;
- 1 enfermeiro atuando no eletrocardiograma; e
- 1 técnico em radiologia atuando com os exames de raio-x.

Com relação a carga de pacientes, optou-se pelo uso dos mesmos parâmetros levantados por Capocci et al. (2016), dessa forma será utilizada a quantidade média de pacientes identificada pelo autor em um dia de atendimento na unidade de saúde. Dessa forma, Esses parâmetros foram utilizados para modelar diferentes cargas em onda para cada ambiente de saúde na cidade inteligente proposta. O comportamento da carga de trabalho em onda é baseado no proposto por Righi et al. (2016) em seu trabalho. Assim, a carga de trabalho de onda está mais intimamente relacionada à realidade do hospital. Dessa forma, para cada instante de tempo em que ocorre entrada de pacientes, um parâmetro $carga(x)$ é recalculado individualmente, modelando um determinado comportamento de carga. A Tabela 7 apresenta as funções para cálculo do valor para $carga(x)$ para a geração de cada comportamento de carga, onde existem 48 instantes de tempo por dia, o equivalente a um instante de tempo a cada 30 minutos.

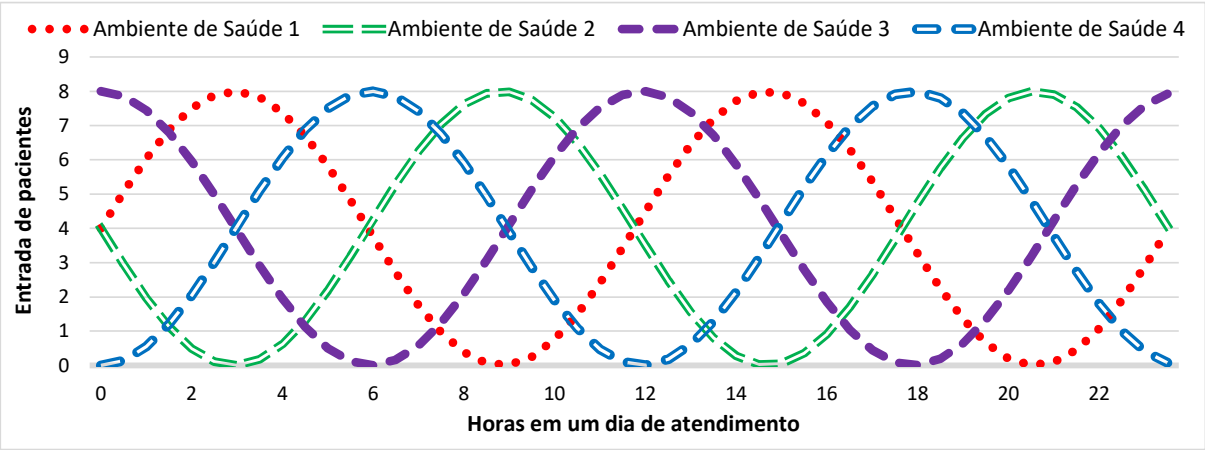
Tabela 7 – Funções para expressar os diferentes comportamentos de carga. Em $carga(x)$, x é o instante de tempo que será realizada a entrada de pacientes.

Ambiente	**Função de Carga**	Parâmetros			
		v	w	t	z
1	$carga(x) = v * z * sen(t * x) + v * z + w$	1	0	0,2674	4
2	$carga(x) = v * z * sen(t * x) + v * z + w$	1	0	-0,2674	4
3	$carga(x) = v{*}z{*}sen(t{*}(x{+}6)){+}v{*}z{+}w$	1	0	0,267	4
4	$carga(x) = v{*}z{*}sen(t{*}(x{+}6)){+}v{*}z{+}w$	1	0	-0,267	4

Fonte: Elaborada pelo autor.

A Figura 37 apresenta graficamente uma representação de cada padrão de carga do modelo. O eixo x expressa o tempo disponível em um dia de atendimento na unidade de saúde, enquanto que o eixo y representa a chegada de pacientes em cada instante de tempo.

Figura 37 – Representação gráfica dos padrões de carga utilizados na avaliação do HealCity.



Fonte: Elaborada pelo autor.

Em função das equações propostas para os padrões de carga gerarem números decimais, foi necessário estabelecer uma estratégia de *discretização* (DAI; TEMLYAKOV, 2024) para gerar números inteiros para chegada dos pacientes no ambiente de saúde. Isso ocorre pois em um ambiente real, não é possível a chegada de 0,2 paciente ou 1,7 pacientes, por exemplo. Dessa forma, adotou-se uma estratégia baseada no acúmulo de carga. Assim, se em um determinado instante houver algo ente 0,1 e 0,9 paciente, esse valor é acumulado com a carga do instante seguinte. Um exemplo dessa estratégia seria um instante qualquer com a carga de 0,6 paciente. Como não haveria um número inteiro de carga, não seria introduzido um paciente no sistema e a carga ficaria acumulada para o próximo instante de tempo. No instante seguinte, com uma nova carga de 0,6 paciente, a carga acumulada seria de 1,2 paciente, resultando na entrada de 1 paciente no ambiente de saúde. Assim, sobraria ainda 0,2 paciente, que seria acumulado para o próximo instante e assim sucessivamente.

5.6 Métricas de Avaliação

A fim de avaliar o modelo proposto, serão analisadas três principais métricas:

- **M1:** Tempo máximo de espera por atendimento,
- **M2:** Custo de Recursos Humanos e
- **M3:** Quantidade elástica de recursos humanos utilizados.

Para avaliar o tempo de espera, será utilizado como parâmetro a variação do tempo máximo de espera entre os cenários e a adequação do tempo máximo de espera aos limites estabelecidos. Dessa forma, devem ser realizadas simulações da cidade inteligente proposta, de forma a encontrar os tempos de espera por atendimento para todos os cenários propostos nessa metodologia de avaliação. A escolha do uso do maior tempo de espera como parâmetro ao invés de usar o tempo médio de espera têm como objetivo avaliar se todos os pacientes serão atendidos dentro dos tempos de atendimento estabelecidos. Um exemplo em que o uso do tempo médio de espera seria inadequado aos objetivos do HealCity poderia acontecer se houvessem cem atendimentos, dos quais 10 pacientes tivessem que esperar 40 minutos para serem atendidos enquanto que os outros 90 pacientes tivessem que esperar apenas 10 minutos cada. Apesar do tempo médio ficar bem inferior ao limite de tempo de 30 minutos, ainda haveriam 10 pacientes insatisfeitos. A escolha dessa métrica se baseia na utilização da análise do tempo de espera nos trabalhos relacionados propostos por Oueida, Aloqaily e Ionescu (2019), Fischer et al. (2020), Liao, Chu e Ho (2022) e Rasidi et al. (2024).

Para determinar o custo de recursos humanos, se torna necessário propor uma forma de medir o custo de um recurso humano em horário normal de trabalho e o custo de um recurso humano fora de seu horário de trabalho. De acordo com a Lei nº 13.467 (BRASIL, 2017) e o Decreto-Lei nº 5.452 (BRASIL, 1943), o pagamento de horas extras será no mínimo 50% (cinquenta por cento) superior à hora normal. Dessa forma, um profissional de saúde alocado fora de seu turno de trabalho custa 50% a mais que um funcionário em seu turno de trabalho. Com base nisso, foi utilizada a Equação 14 para *Custo de recursos humanos* da seguinte forma:

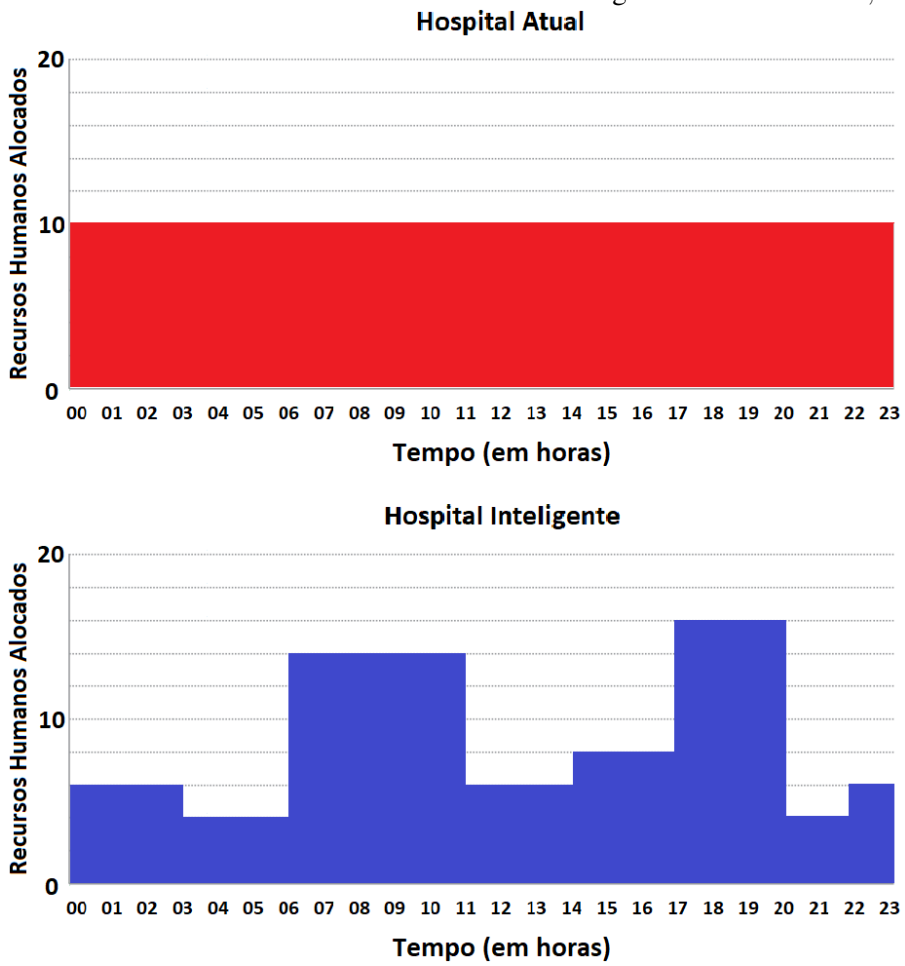
$$Cost(t_i, t_f) = \frac{1}{t_f - t_i} \sum_{t_n=t_i}^{t_f-1} HR(t_n) + (1.5 \cdot AllocatedHR(t_n)) \tag{14}$$

onde $HR(t_n)$ refere-se a todos os recursos humanos em seu turno de trabalho no instante de tempo t_n, e $AllocatedHR(t_n)$ refere-se a todos os recursos humanos alocados, ou em processo de alocação, fora do horário normal de trabalho no instante de tempo t_n. A escolha dessa métrica se baseia na utilização de análises semelhantes de custo de recursos humanos nos trabalhos relacionados propostos por Apornak et al. (2021a) e Hafezalkotob et al. (2022).

Com relação a quantidade elástica de recursos humanos, espera-se que o modelo seja capaz de utilizar os profissionais de saúde existentes na cidade inteligente de uma forma otimizada,

de forma que a alocação fixa do primeiro cenário, com quarenta e quatro funcionários atuando 24 horas por dia, possam ser comparados com a alocação proposta pelo HealCity, com a quantidade de profissionais da cidade variando ao longo do dia conforme a demanda. Nesse contexto, espera-se que o modelo utilize a mesma quantidade de recursos humanos já utilizada pelos ambientes de saúde, entretanto aceita-se o aumento da quantidade de profissionais atuando desde que o tempo de espera seja reduzido. A Figura 38 apresenta graficamente uma representação da alocação de recursos humanos ao longo de um dia de atendimento em uma unidade de saúde com e sem o uso da Elasticidade Multinível de Recursos Humanos em cidades inteligentes. O eixo x expressa o tempo, enquanto que o eixo y representa a alocação de atendentes em cada instante de tempo. A escolha dessa métrica se baseia na utilização da quantidade de recursos humanos no trabalho relacionado proposto por Fischer et al. (2020).

Figura 38 – Representação gráfica de um exemplo de alocação de recursos humanos ao longo do tempo de atendimento no ambiente de saúde com e sem o uso do modelo HealCity, onde no Ambiente de Saúde Atual têm-se $M2 = 10$ atendentes e no Ambiente de Saúde Inteligente têm-se $M2 = 8,75$ atendentes.



Fonte: Elaborada pelo autor.

A Tabela 8 apresenta todas as métricas para avaliação descritas anteriormente, relacionando os resultados esperados para o segundo e terceiro cenários, quando comparado aos ambientes de saúde atuais, sem o uso de estratégias de gerenciamento de recursos humanos.

Tabela 8 – Métricas para avaliação e resultados esperados em cada um dos cenários propostos.

	Métrica	C1	C2 (Esperado)	C3 (Esperado)
M1:	Tempo Máximo de Espera	Atual	Menor que no C1	Menor que no C2
M2:	Custo de recursos humanos	Atual	Igual a C1	Maior que no C1 e C2
M3:	Quantidade Elástica de Recursos Humanos utilizados	11 por turno	11 por turno	11 ou mais por turno

Fonte: Elaborada pelo autor.

5.7 Metodologia de Testes para o Sistema RTLS

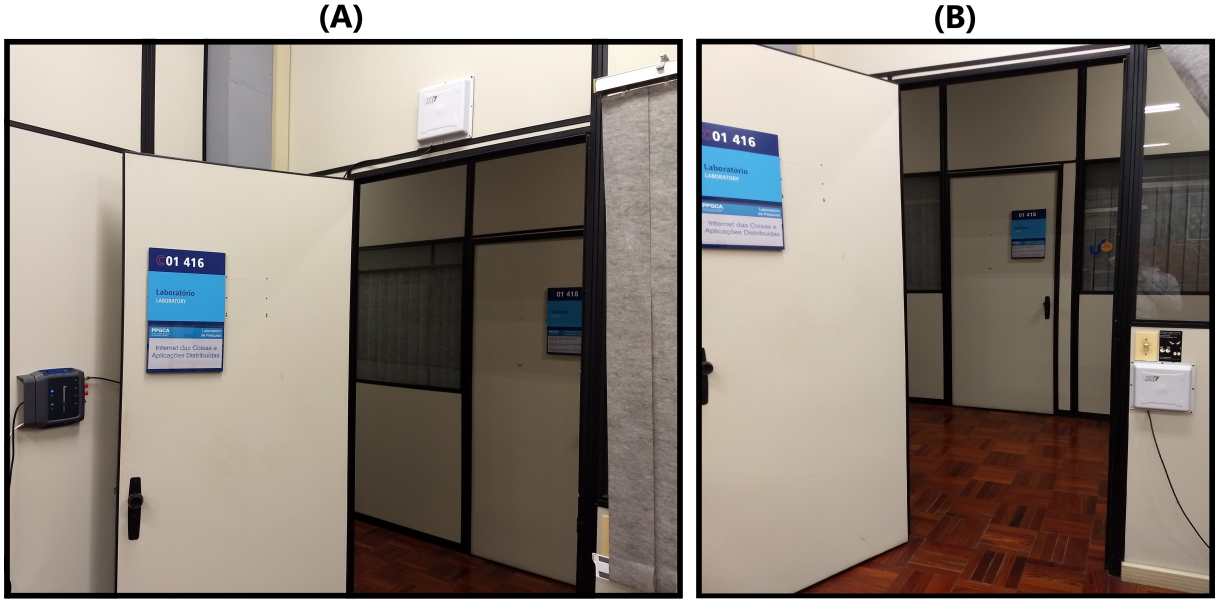
Além de testes a nível de software para verificar a eficácia do modelo HealCity para o gerenciamento elástico de recursos humanos, também optou-se por realizar testes a nível de hardware. Como o modelo HealCity baseia-se no uso da Internet das Coisas e um sistema RTLS para identificar a localização e movimentação de pacientes, bem como o tempo de permanência em cada ambiente e consequente o tempo despendido para atendimento, decidiu-se que seria interessante fazer a instalação do sistema RTLS em um ambiente real para testar sua capacidade de capturar a movimentação de pessoas. Dentre as tecnologias estudadas para Internet das Coisas, optou-se pelo uso da tecnologia RFID para os testes do modelo. A escolha do RFID se baseia na sua capacidade de permitir identificação automática, evitando a existência do erro humano nos processos de identificação. Além disso, segundo Boulos e Berry (2012) o uso de soluções RTLS que incorporem tecnologias complementares como RFID, podem fornecer níveis de precisão e flexibilidade inatingíveis por qualquer tecnologia concorrente. O ambiente selecionado para instalação do equipamento RFID para realização dos testes foi o laboratório de Internet das Coisas e Aplicações Distribuídas, localizado na sala C01 416 do Programa de Pós-Graduação em Computação Aplicada (PPGCA) da Universidade do Vale do Rio dos Sinos.

Para os testes do sistema RTLS, foi escolhido o uso do leitor RFID *Intermec IF2*[1]. A escolha desse leitor baseia-se no fato do mesmo trabalhar com até quatro antenas para leitura de *tags*, o que permite o aumento da área de captura com a instalação de diversas antenas. Além disso, ao contrário de leitores como o *Identix miniPad*[2] que possuem um raio de leitura de alguns centímetros, cada antena do leitor IF2 pode realizar leituras de *tags*-RFID localizadas até a alguns metros de distância. As *tags* utilizadas nos testes foram etiquetas adesivas EPC. Para os testes, o leitor IF2 foi instalado com apenas uma antena, em um primeiro momento localizada acima da porta de entrada do laboratório, e após foi remanejada para ao lado dessa mesma porta conforme pode ser visto na Figura 39. O objetivo desses testes é identificar qual o raio de captura de um leitor RFID em um ambiente real, de forma a propor o melhor local para instalação do mesmo de forma a capturar a entrada de pessoas em ambientes físicos, tais como o ambiente de saúde inteligente proposto nesse trabalho. Para tanto, serão medidas as distâncias máximas em que o leitor é capaz de captar o sinal da *tag* em todas as direções. Através dos valores

[1] https://www.honeywellaidc.com/products/rfid/fixed-readers/if2
[2] https://www.identix.us.com/minipad

encontrados, espera-se identificar se o uso de quatro antenas é o suficiente para cobrir a área de portas simples e também de portas duplas.

Figura 39 – Leitor RFID Intermec IF2 instalado no laboratório de Internet das Coisas e Aplicações Distribuídas do PPGCA da Universidade do Vale do Rio dos Sinos onde em (A) a antena foi instalada acima da porta e em (B) a antena foi instalada ao lado da porta.



Fonte: Elaborada pelo autor.

Outro módulo parcialmente desenvolvido foi o Capture, responsável por receber as leituras geradas pelos sensores RFID e identificar as informações relevantes a cada captura para serem utilizadas nos módulos seguintes. Dessa forma, propõe-se a realização de testes de carga no módulo a fim de se verificar a quantidade máxima de registros que a aplicação é capaz de processar em um intervalo de um segundo. E a partir disso, identificar quantos pacientes ou recursos humanos do ambiente de saúde poderiam passar ao mesmo tempo em um conjunto de leitores, sem gerar-se filas para processamento das inclusões na base de dados. O ambiente usado para estes testes foi uma máquina virtual com as seguintes características: Processador Intel Core I7 3770, Memória 16GB com 1600 MHz, Disco rígido SSD de 120 GB Kingston v300, Sistema operacional Windows 7 Professional SP1, Servidor MySQL 5.6.17 e Java Virtual Machine versão Java 8 Update 31.

Para a simulação do sistema RFID, definiu-se o uso da ferramenta *Rifidi* (TRANSCENDS, 2011). Segundo Neiva (2012), o *Rifidi Emulator* trata-se de um ambiente de emulação que permite ao usuário adicionar e remover leitores e *tags* ao sistema, permitindo a emulação dos mesmos. Já para captura dos dados através dos leitores RFID, definiu-se o uso do *middleware* Fosstrak (FOSSTRAK, 2017). O Fosstrak é um dos *middlewares* mais usados para prover a conexão das leituras de tags RFID com as aplicações de usuário (MARQUES; RIGHI, 2015). A escolha do *middleware* se baseia no fato do mesmo ser *open source*, o que facilita a criação de ferramentas e módulos para interagir com o mesmo, e por ele implementar o padrão EPC-

Global (GS1 EPCGLOBAL, 2016). Para um ambiente de produção, ou de testes, é necessária a instalação de no mínimo três componentes do Fosstrak: o ALE *Middleware*, o EPCIS e o *Capturing Application*. Os leitores RFID são conectados logicamente ao ALE que recepciona os dados capturados por eles e faz a filtragem dos mesmos (FOSSTRAK, 2017). O ALE então passa essa informação para o *Capturing Application* que é responsável de fazer a conexão entre os outros dois módulos, e definir o que vai chegar ao EPCIS, através de algumas definições de regras de negócio (MARQUES; RIGHI, 2015). Por fim, o EPCIS recepciona as informações geradas e é responsável por armazenar no banco de dados MySQL.

Dessa forma, foi configurado um ambiente composto por duas máquinas virtuais para execução e testes do sistema RFID. A primeira máquina ficou responsável pela execução dos componentes ALE e *Capturing Application* do Fosstrak e do *Rifidi Emulator*. As especificações da VM utilizada são as seguintes: Processador Intel Core i7-3770 (1 núcleo destinado para a VM), Memória 2GB com 1600MHz, Sistema Operacional Debian GNU/Linux 8 (Jessie) 32-bit, Apache Tomcat 8.5, ALE *Middleware* 1.2.0, *Capturing Application* 0.1.1 e *Rifidi Emulator* 1.6 Já a segunda máquina virtual ficou responsável pela execução dos componente EPCIS do Fosstrak e pela base de dados da aplicação. As especificações da VM foram: Processador Intel Core i7-3770 (1 núcleo destinado para a VM), Memória 2GB com 1600MHz, Sistema Operacional Windows 10 Home 64-bit, Apache Tomcat 9.0, EPCIS *Repository* 0.5.0 e MySQL 5.7. A fim de testar o ambiente configurado e verificar o funcionamento do Fosstrak e do *Rifidi Emulator*, foram realizados diferentes testes de carga no sistema. O objetivo deste teste é verificar o tempo para as requisições ao EPCIS, de forma a identificar a capacidade do *middleware* de lidar com as requisições do módulo Capture. As cargas escolhidas para o teste foram:

- **4x0**: Simular 4 leitores com nenhuma *tag* sendo lida, resultando em carga nula;
- **4x1**: Simular 4 leitores com 1 *tag* por leitor, resultando em carga de 4 *tags* por ciclo;
- **4x2**: Simular 4 leitores com 2 *tags* por leitor, resultado em carga de 8 *tags* por ciclo.
- **4x4**: Simular 4 leitores com 4 *tags* cada, resultando em carga de 16 *tags* por ciclo.
- **4x8**: Simular 4 leitores com 8 *tags* cada, resultando em carga de 32 *tags* por ciclo.

5.8 Considerações Parciais

Este capítulo apresentou a metodologia proposta para avaliar e validar o modelo HealCity. Assim como a maioria dos trabalhos relacionados, o HealCity será avaliado através do uso de técnicas de simulação. Com relação à implementação e ao ambiente de testes, foi proposto o desenvolvimento dos módulos Elastic e Predict do modelo, através do uso das linguagens de programação Java e Python. Para o ambiente de testes foi utilizado um equipamento com as seguintes características: Processador Intel Core I9 10900f com dez núcleos físicos de 5.2 GHz, Memória de 64GB com frequência de 3200 MHz, SSD NVMe de 512 GB ADATA SX8100NP,

Sistema operacional Windows 10 Pro, Servidor MySQL 5.7.23, Java Virtual Machine versão Java 8 Update 333 e Python versão 3.11. Com relação aos parâmetros utilizados para simulação dos ambientes de saúde, foi proposto o uso dos dados levantados por Capocci et al. (2016) em seus estudos. Além disso, foi proposto o uso de leis e projetos de leis brasileiras para nortear os *thresholds* e regras para definição dos recursos humanos disponíveis para alocação.

Também neste capítulo estão presentes os três cenários que serão comparados: (i) **Cidade inteligente sem qualquer abordagem de gestão de recursos humanos**, (ii) **Cidade inteligente utilizando a abordagem proposta por Capocci et al. (2016)** e (iii) **Cidade inteligente com o modelo de elasticidade proposto pelo HealCity**. Estão presentes também as três métricas de avaliação e os trabalhos relacionados que as inspiraram: (i) **Tempo máximo de espera por atendimento** (OUEIDA; ALOQAILY; IONESCU, 2019; FISCHER et al., 2020; LIAO; CHU; HO, 2022; RASIDI et al., 2024), (ii) **Custo de Recursos Humanos** (APORNAK et al., 2021a; HAFEZALKOTOB et al., 2022) e (iii) **Quantidade elástica de recursos humanos utilizados** (FISCHER et al., 2020).

Por fim, neste capítulo foi apresentado o ambiente e a metodologia de testes para o sistema RTLS, mostrando os equipamentos escolhidos, o local para instalação dos mesmos, para identificar se o uso de quatro antenas é o suficiente para cobrir a área de portas simples e também de portas duplas em ambientes de saúde. Além disso, também foram vislumbrados testes para de forma a identificar a capacidade de um *middleware* de lidar com as requisições do módulo Capture proposto pelo modelo HealCity.

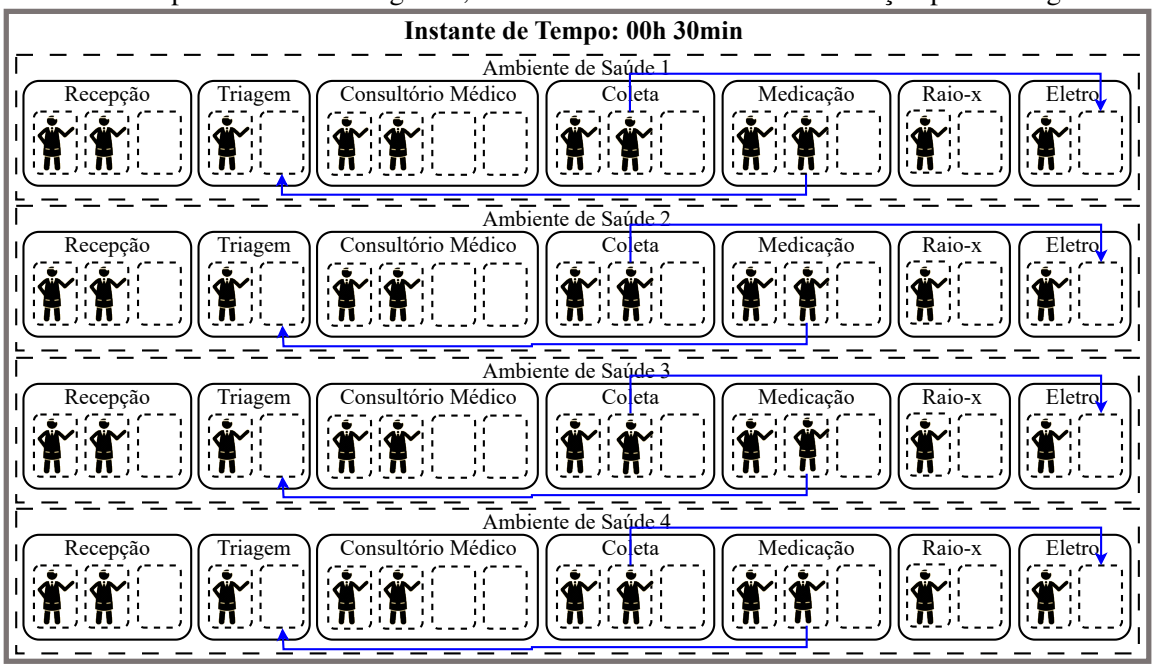
6 RESULTADOS E DISCUSSÃO

A partir da metodologia de avaliação proposta para o modelo HealCity, para cada cenário proposto, foi executada uma simulação da cidade inteligente proposta. Em cada cidade simulada, foram analisados os ambientes de saúde individualmente, dessa forma cada um dos quatro ambientes de saúde foi simulado para os três cenários, resultando em um total de doze ambientes hospitalares simulados. Nas Seções 6.2, 6.3 e 6.4 serão analisadas cada umas das métricas de avaliação propostas e os resultados obtidos. Além disso, a partir da metodologia de testes proposta para o sistema RTLS, foi instalado o equipamento para leitura RFID em um ambiente real e realizados sucessivos testes para definir a sua capacidade de captura de *tags* e identificar a capacidade do mesmo de captar a entrada de pessoas em ambientes físicos. A Seção 6.5 apresenta os resultados obtidos com os testes do sistema RTLS. Por fim, a Seção 6.6 apresenta uma discussão sobre os resultados encontrados durante a avaliação do modelo proposto.

6.1 Funcionamento do Modelo

A fim de mostrar o funcionamento do modelo proposto, resolveu-se apresentar alguns momentos em que o modelo proposto realizou a (i) realocação de recurso humanos dentro de um mesmo ambiente de saúde, (ii) realocação de atendentes entre diferentes ambientes de saúde, (iii) alocação e a (iv) desalocação de equipe médica.

Figura 40 – Instante de tempo em que foi identificada a necessidade de (i) realocar atendentes internamente nos ambientes de saúde. Para cada ambiente de saúde, foi proposta a realocação de um enfermeiro da sala de coleta para o eletrocardiograma, e um enfermeiro da sala de medicação para a triagem.

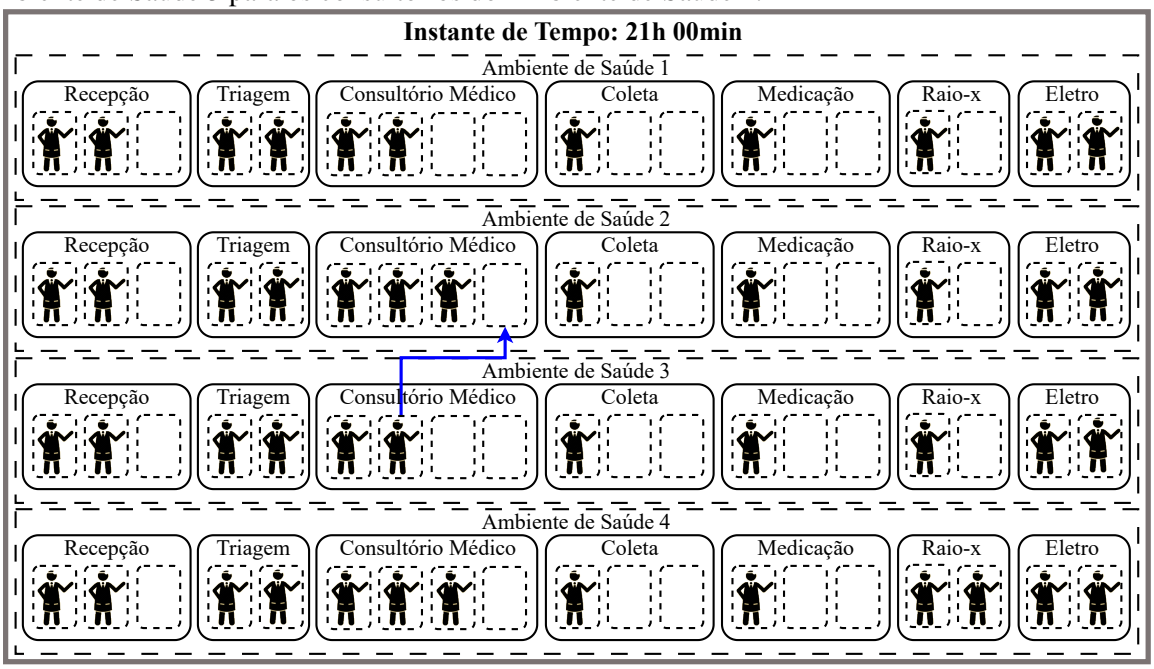


Fonte: Elaborada pelo autor.

Um exemplo de realocação de recursos humanos pode ser identificado logo no instante de

tempo 00h 30min, onde em cada ambiente de saúde havia a alocação original de atendentes proposta, sem nenhuma movimentação. Como o modelo identificou que as salas de triagem e de eletrocardiogramas seriam mais utilizadas que as salas de exame de coleta e de medicação, o HealCity propôs a realocação de enfermeiros entre essas salas conforme ilustrado na Figura 40.

Figura 41 – Instante de tempo em que foi identificada a possibilidade de (ii) realocar um atendente entre diferentes ambientes de saúde. Neste caso um doutor foi realocado dos consultórios médicos do Ambiente de Saúde 3 para os consultórios do Ambiente de Saúde 2.



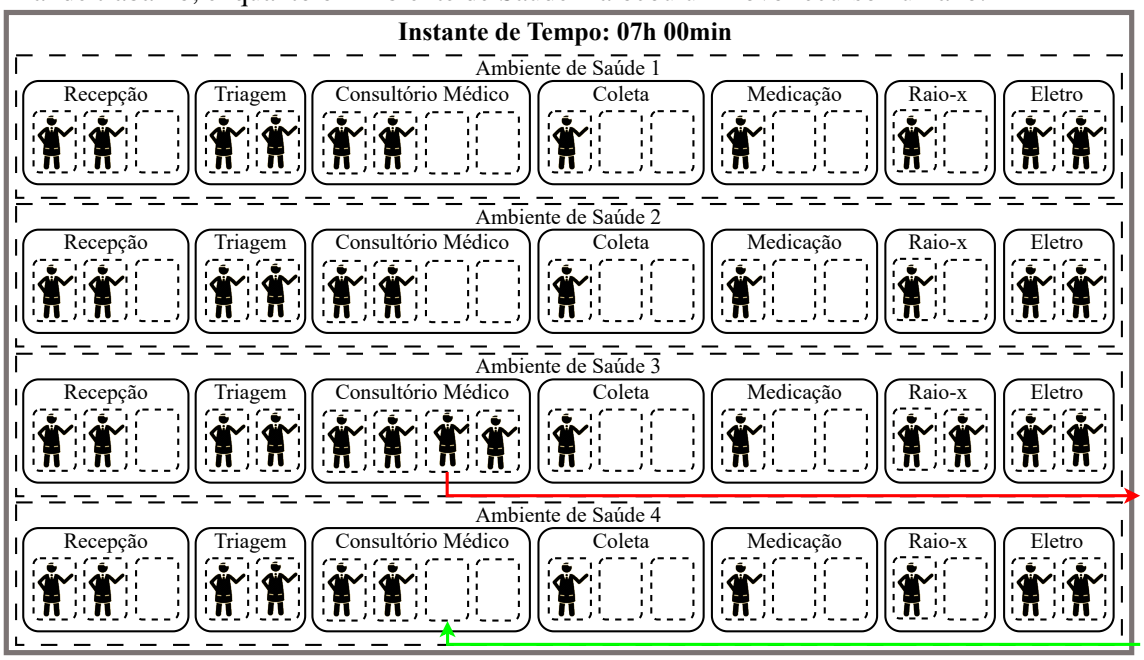
Fonte: Elaborada pelo autor.

Outro exemplo de realocação, agora entre diferentes ambientes de saúde foi identificada no instante de tempo 21h 00min. Neste momento o modelo identificou que o consultório médico do Ambiente de Saúde 2 estaria necessitando de mais atendentes do que haviam disponíveis. Ao mesmo tempo, o HealCity percebeu que o Ambiente de Saúde 3 poderia disponibilizar um médico. Assim o HealCity propôs a movimentação de um doutor entre estes dois ambientes de saúde conforme ilustrado na Figura 41. É possível observar um exemplo de alocação de um novo recurso humano no instante de tempo 07h 00min. Neste momento o modelo identificou que o consultório médico do Ambiente de Saúde 4 estaria necessitando de um novo atendente. Assim o HealCity propôs a alocação de um doutor que estava em sua residência para atender no ambiente de saúde necessitado. Por fim, um exemplo de desalocação pôde ser visto no mesmo instante de tempo, onde um doutor já alocado há quatro horas no Ambiente de Saúde 3 pôde ser liberado, conforme é possível observar na Figura 42.

6.2 Métrica M1: Tempo Máximo de Espera

O primeiro parâmetro que foi avaliado é o tempo máximo de espera por atendimento nas filas do ambiente de saúde proposto. Para a comparação e a análise dessa métrica, a cada

Figura 42 – Instante de tempo em que foi identificada a possibilidade de (iii) alocar e (iv) desalocar um recurso humano. Neste caso o Ambiente de Saúde 3 pode liberar um médico alocado fora de seu turno normal de trabalho, enquanto o Ambiente de Saúde 4 alocou um novo recurso humano.

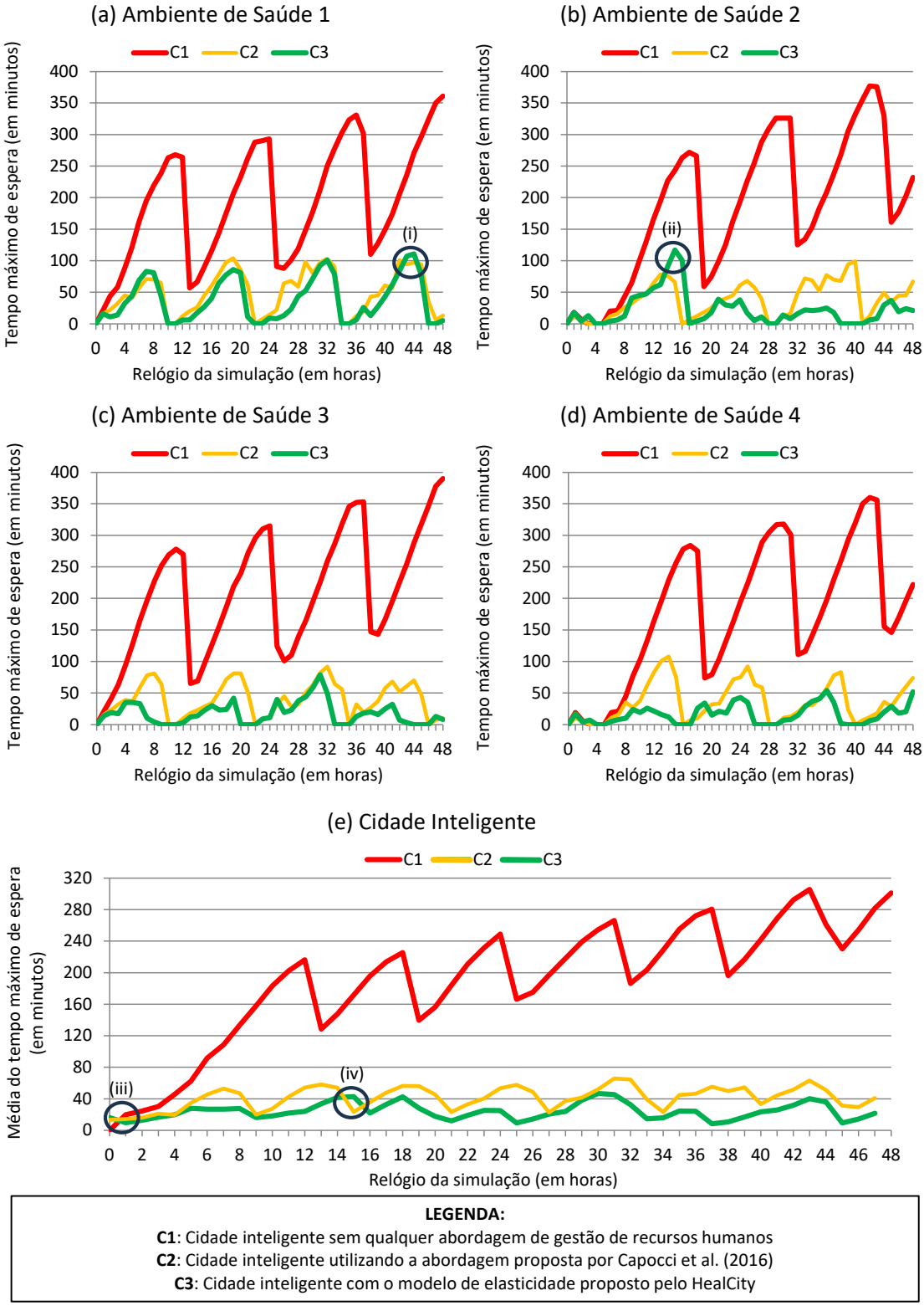


Fonte: Elaborada pelo autor.

cenário de testes, o maior tempo registrado de espera a cada hora foi medido e armazenado. O principal objetivo dessa métrica é verificar se através do uso do modelo HealCity proposto para controle de recursos humanos em ambientes de saúde de cidades inteligentes é possível diminuir o tempo de espera por atendimento de pacientes aguardando nestes ambientes de saúde, tal como proposto na hipótese de pesquisa proposta no Capítulo 1. Além disso, espera-se verificar se a maior espera identificada nas observações e a média das maiores esperas é inferior ao *threshold* de 30 minutos proposto no Projeto de Lei de 14 de junho de 2018 (GOIANIA, 2018) para dias normais ou de 45 minutos em datas especiais.

A Figura 43 apresenta o tempo máximo de espera registrado para cada um dos cenários propostos ao longo do período simulado de quarenta e oito horas. Conforme pode ser visto na Figura 43, os dados coletados mostraram uma redução significativa no tempo de espera quando as abordagens de elasticidade propostas pelo HealCity foram usadas em comparação a cidade inteligente sem elas. Devido à elasticidade de recursos humanos proposta neste trabalho, o HealCity demonstrou ser capaz de diminuir o tempo máximo médio de espera em 80,33%, 87,64%, 91,52% e 90,98% para os ambientes de saúde 1, 2, 3 e 4, respectivamente, em comparação ao cenário em que nenhuma gestão de recursos humanos é realizada, totalizando uma redução de 87,62% na média da cidade inteligente. Quando compara-se o segundo cenário (usando a estratégia proposta por Capocci et al. (2016)) e o terceiro cenário (usando a estratégia proposta neste trabalho), o HealCity demonstrou diminuir o tempo máximo de espera em 21,93%, 40,11%, 55,27% e 55,67% para os ambientes de saúde 1, 2, 3 e 4, respectivamente, totalizando uma

Figura 43 – Tempo máximo de espera registrado por ambiente de saúde para cada um dos cenários propostos, C1 (em vermelho), C2 (em laranja) e C3 (em verde), para o (a) Ambiente de Saúde 1, (b) Ambiente de Saúde 2, (c) Ambiente de Saúde 3 e (d) Ambiente de Saúde 4, e a média do tempo máximo de espera encontrado na (e) Cidade inteligente.



Fonte: Elaborada pelo autor.

redução de 41,79% na média da cidade inteligente.

Destacam-se alguns eventos que ocorreram na simulação conforme apresentado na Figura 43: (i) Neste momento, C3 teve pior desempenho devido ao fato de que a sala de Eletrocardiograma havia transferido um funcionário para Exames de Coleta no ciclo anterior. Como resultado da estratégia de *cooldown period*, não foi possível alocar o mesmo atendente de volta no ciclo seguinte. Consequentemente, o aumento da demanda só foi resolvido no ciclo subsequente, pois não existiam atendentes disponíveis para alocação; (ii) Neste momento, C3 teve um tempo de espera maior que C2 devido à sala de triagem ter atingido sua capacidade máxima de atendentes, o que impediu o modelo de alocar mais atendentes para ajudar. Em contraste, C2 havia alocado profissionais anteriormente para a sala, minimizando assim o efeito; (iii) No início da simulação, o tempo de espera em C3 foi maior que em C2, pois C2 utilizou uma estratégia de movimentação de profissionais baseada em uma condição, enquanto C3 operou por meio de ciclos de monitoramento. Consequentemente, o C2 teve um desempenho ligeiramente melhor nesta fase, pois a resposta inicial foi mais rápida que o ciclo de monitoramento do C3; e (iv) A média da cidade mudou neste ponto, conforme discutido anteriormente no item "ii". A Tabela 9 apresenta a média dos tempos máximos de esperas identificados, juntamente com o maior tempo máximo de espera identificado, em cada um dos cenários propostos. Através dos dados levantados é possível identificar uma redução significativa no tempo de espera com o uso da Elasticidade de Recursos Humanos para Saúde em Cidades Inteligentes quando comparado com a cidade sem o uso da estratégia de elasticidade.

Tabela 9 – Maior tempo de espera e média dos maiores tempos de espera registrados nos cenários propostos para cada ambiente de saúde e para a média da cidade inteligente.

Cenário	**Métrica**	**Ambiente de Saúde**				**Média da cidade inteligente**
		1	**2**	**3**	**4**	
C1	Tempo Médio	197,56	186,56	212,31	180,43	194,21
	Maior Tempo	361	377	390	360	305,7
C2	Tempo Médio	49,77	38,48	40,25	36,70	41,30
	Maior Tempo	104	99	92	108	48
C3	Tempo Médio	38,85	23,04	18,00	16,27	24,04
	Maior Tempo	111	117	79	54	46,7

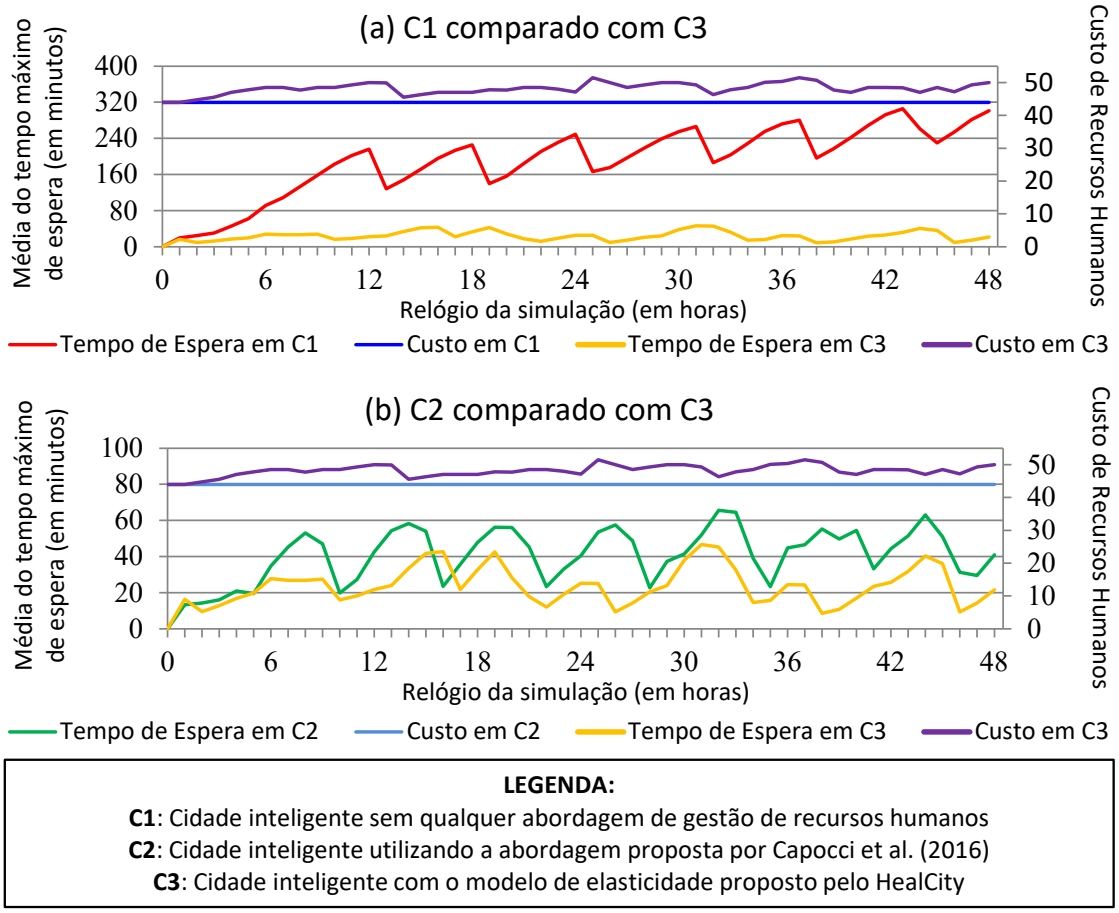
Fonte: Elaborada pelo autor.

6.3 Métrica M2: Custo de Recursos Humanos

A segunda métrica a ser avaliada é o custo estimado com recursos humanos, que leva em consideração o custo maior dos funcionários alocados fora de seus turnos normais de trabalho. Para a comparação e a análise dessa métrica, a cada cenário de testes, o custo de recursos humanos a cada hora foi calculado e armazenado. O principal objetivo dessa métrica é verificar se através do uso do modelo HealCity proposto para controle de recursos humanos em ambientes de saúde de cidades inteligentes é possível diminuir significativamente o tempo de espera sem

aumentar o custo na mesma proporção. Previmos que o custo de recursos humanos seria maior entre os cenários. Para essa métrica, não faria sentido comparar o custo para cada ambiente de saúde, já que no modelo HealCity, todos eles conseguem compartilhar profissionais. Dessa forma, foi considerado o custo total da cidade inteligente, já que em determinados momentos um ambiente médico pode ter menos que os 11 profissionais propostos em sua carga horária original, pois transferiu um profissional para outro ambiente de saúde, que nesse contexto teria 12 profissionais, sem gerar custos de hora extra. A Figura 44 mostra o custo de recursos humanos para C1 comparado a C3 e C2 comparado a C3. Foi possível observar que o custo variou de 44 a 51,52 por hora, na abordagem proposta pelo HealCity. Para o cenário C3, o custo aumentou em 9,68% comparado aos outros cenários.

Figura 44 – Custo dos recursos humanos comparado com a média do tempo máximo de espera na cidade inteligente em (a) C1 e C3 e (b) C2 e C3.



LEGENDA:
C1: Cidade inteligente sem qualquer abordagem de gestão de recursos humanos
C2: Cidade inteligente utilizando a abordagem proposta por Capocci et al. (2016)
C3: Cidade inteligente com o modelo de elasticidade proposto pelo HealCity

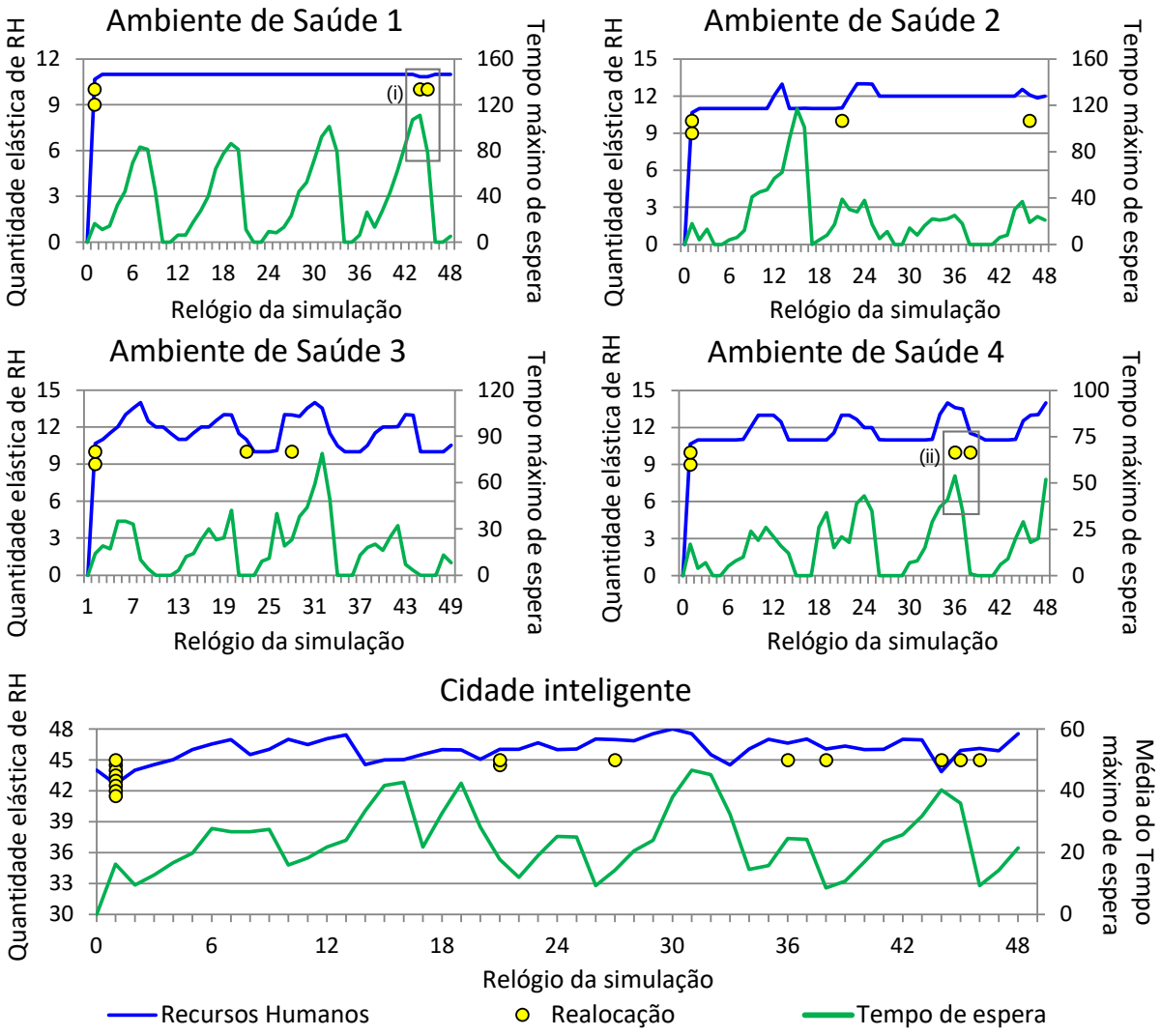
Fonte: Elaborada pelo autor.

6.4 Métrica M3: Quantidade Elástica de Recursos Humanos utilizados

A terceira métrica a ser avaliada é a quantidade elástica de recursos humanos utilizados para prover o atendimento aos pacientes na cidade inteligente proposta. Para comparação e

análise dessa métrica, foi calculada a quantidade média de recursos humanos alocados a cada hora por ambiente de saúde, e no total da cidade inteligente. O principal objetivo dessa métrica é verificar qual foi o aumento do uso de recursos humanos com o uso do modelo HealCity proposto e comparar essa métrica com o tempo de espera, de forma a verificar se o aumento da quantidade de recursos humanos acaba por impactar positivamente o atendimento aos pacientes. O aumento da quantidade de pessoas em atendimento não representa um problema ao ambiente hospitalar, uma vez que a métrica M1, já analisada, apresentou uma redução significativa.

Figura 45 – Quantidade elástica de recursos humanos utilizados para cada ambiente de saúde, e para o total da cidade inteligente, no Cenário C3, com o uso da Elasticidade de Recursos Humanos para Saúde em Cidades Inteligentes. Neste gráfico, realocação refere-se as movimentações de profissionais que ocorrem nos ambientes de saúde, em que não há acréscimo ou redução de RH no sistema.



Fonte: Elaborada pelo autor.

Conforme já mostrado anteriormente, era previsto um aumento no número de recursos humanos utilizados na cidade inteligente. A Figura 45 apresenta a quantidade elástica de recursos humanos utilizados para atendimento na cidade inteligente para o cenário C3, o único cenário

em que a quantidade de funcionários acaba por apresentar alguma variação. O número elástico de recursos humanos variou de 44 a 52 por hora. Destacam-se dois eventos que ocorreram na simulação conforme apresentado na Figura 45: (i) Neste momento, é possível ver que o tempo de espera no Ambiente de Saúde 1 estava em um ritmo crescente. Como resultado, no instante de tempo 43h e 30min, o HealCity realocou um enfermeiro entre a sala de eletrocardiograma do Ambiente de Saúde 2 e a sala de exames de coleta do Ambiente de Saúde 1 e logo após mais um enfermeiro foi realocado da sala de eletrocardiograma do Hospital 2 para a sala de exames de coleta do Hospital 1 no instante de tempo 44h e 30min, o que ocasionou uma queda acentuada no tempo de espera para o ambiente de saúde; (ii) Neste ponto, é possível ver que o tempo de espera no Ambiente de Saúde 4 também estava em um ritmo crescente. Como resultado o HealCity propôs a realocação de um enfermeiro entre as salas de eletrocardiograma e exames de coleta no instante de tempo 35h e 30min deste ambiente de saúde, o que também ocasionou uma queda perceptível no tempo de espera. Conforme pode ser visto, quando o modelo tem um aumento no númeo de recursos humanos, ou ocorre alguma realocação entre as salas de atendimento, o tempo de espera acaba por diminuir. Além disso, como pode ser visto na Figura 45, é evidente que o tempo de espera do paciente diminui 87,62% e 41,79% devido aos procedimentos de realocação e alocação do HealCity entre os cenários S1/S3 e S2/S3, respectivamente. A Tabela 10 apresenta a quantidade média elástica de recursos humanos utilizados em cada ambiente de saúde, juntamente com a média da cidade inteligente.

Tabela 10 – Quantidade média elástica de Recursos Humanos utilizados em cada ambiente de saúde, juntamente com a média da cidade inteligente para cada cenário proposto.

Cenário	**Ambiente de Saúde**				**Média da cidade inteligente**
	1	**2**	**3**	**4**	
C1	11	11	11	11	44
C2	11	11	11	11	44
C3	10,99	11,69	11,64	11,78	46,15

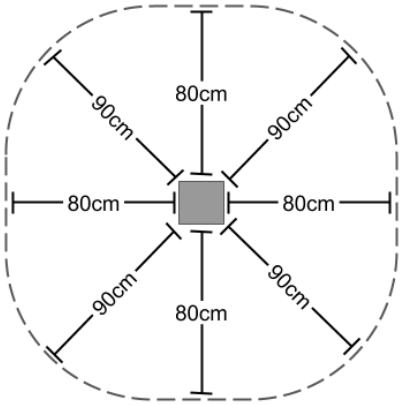
Fonte: Elaborada pelo autor.

6.5 Testes do Sistema RTLS

Baseado na metodologia de testes proposta para o sistema RTLS, o primeiro teste que foi executado foi verificar se com uma antena localizada acima de uma porta, seria possível identificar a movimentação de pessoas. Notou-se para esse cenário que o leitor foi capaz de captar uma *tag* entrando na sala em toda a largura da porta, entretanto não em toda a sua altura. Em função disso, o mesmo teste foi executado com a antena remanejada para a lateral da porta, onde os resultados foram mais satisfatórios, uma vez que a *tag* pode ser capturada em quase toda a altura da porta. Com a antena do leitor localizada dessa forma, pode-se notar que a tag só é lida ao adentrar na sala. Se uma pessoa ficasse parada na entrada, mas não passasse pelo arco, não seria capturada pelo leitor. Dessa forma, apenas ao adentrar na sala que o leitor identificaria

a pessoa. Isto ocorreu pois na parte traseira da antena existe uma chapa de material que impede a transmissão do sinal.

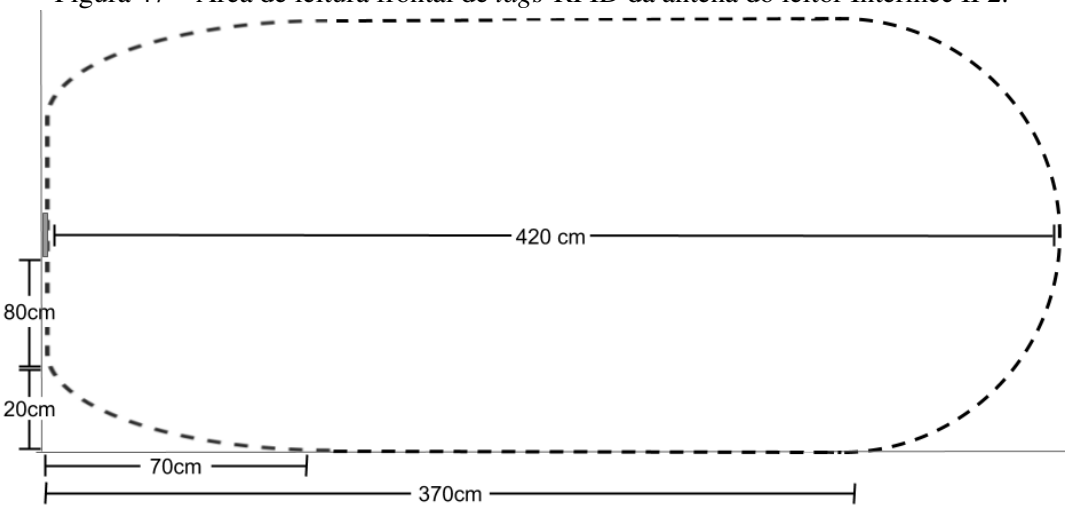
Figura 46 – Área de leitura de *tags*-RFID em torno da antena do leitor Intermec IF2.



Fonte: Elaborada pelo autor.

Para então identificar a localização ideal para cobrir todo e qualquer ponto de qualquer porta onde fosse instalada, foram feitos novos testes para identificar qual a área em torno da antena capaz de captar *tags*-RFID. Para tanto, a antena foi instalada em um local a aproximadamente 1,1 metro de altura. Com base nos testes, pode-se notar que lateralmente o limite de captura da antena é em torno de 80 centímetros. Já diagonalmente, esse limite fica em torno de 90 centímetros, o que indica que o raio de sinal não é circular. A Figura 46 apresenta o raio estimado de sinal para a antena. Já a área de captura frontal da antena provou-se ser extremamente superior quando comparado à área lateral. Nos testes executados, pode-se notar que em linha reta é possível identificar com precisão uma *tag* localizada até 4,2 metros de distância da antena. A partir de 70 centímetros da antena, o raio de captura chega ao chão e se estende até 3,7 metros de distância. A Figura 47 apresenta o raio frontal estimado de sinal para a antena.

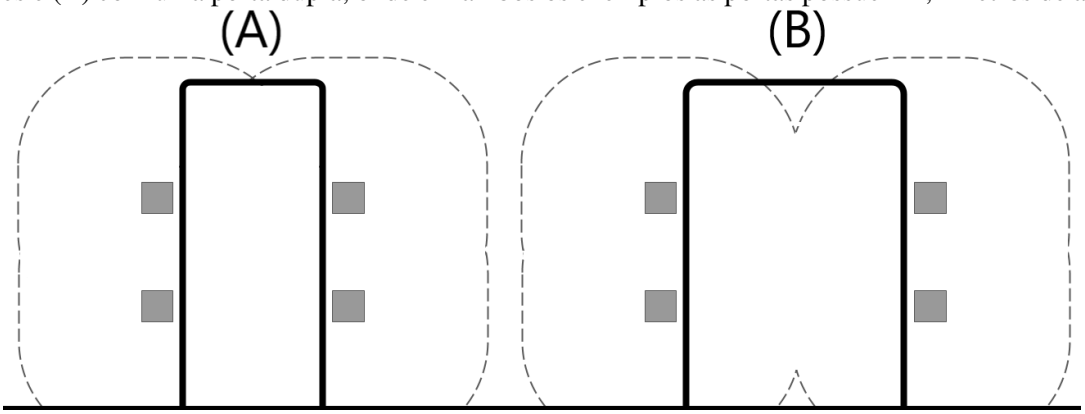
Figura 47 – Área de leitura frontal de *tags*-RFID da antena do leitor Intermec IF2.



Fonte: Elaborada pelo autor.

Assim sendo, propõe-se para a melhor precisão de captura a instalação de quatro antenas, o limite máximo disponível para o leitor IF2, em torno da porta ao qual se deseja registrar a passagem de pessoas. As antenas devem estar localizadas o mais próximo possível do marco da porta, e igualmente espaçadas com relação a altura da mesma. Utilizando-se o raio de captura medido anteriormente, pode-se estimar o raio de cobertura para uma porta simples e em uma porta dupla, utilizando as quatro antenas em conjunto, conforme pode ser visto na Figura 48. Nesse cenário, com o uso das quatro antenas em torno de uma porta simples, não existem pontos em que as mesmas não são capazes de captar uma *tag*, tornando a alocação das mesmas ideal a esse cenário. Já para o uso em porta duplas, existem alguns pontos em que as mesmas não são capazes de ler *tags*, entretanto, como mostrado na Figura 47, conforme o paciente for adentrando o ambiente, o ponto cego de leitura vai diminuindo gradativamente, fazendo com que o mesmo seja identificado.

Figura 48 – Instalação proposta das antenas do leitor Intermec IF2 em dois cenários: (A) com uma porta simples e (B) com uma porta dupla, onde em ambos os exemplos as portas possuem 2,1 metros de altura.

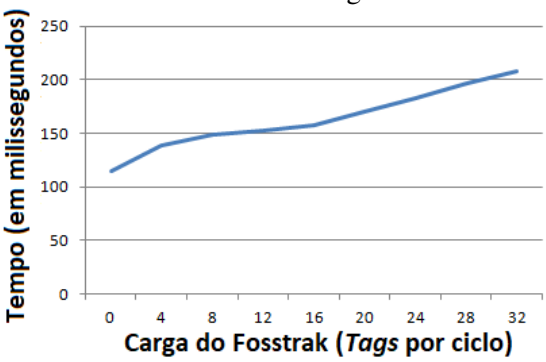


Fonte: Elaborada pelo autor.

Para avaliação do módulo Capture, conforme proposto na metodologia de avaliação, foram realizados sucessivos testes de carga no módulo a fim de se verificar a quantidade máxima de registros que a aplicação é capaz de processar em um intervalo de um segundo. Após os testes concluiu-se que, em média, o sistema é capaz de processar 230 marcações no intervalo de um segundo. Partindo-se da premissa que entre um leitor e outro exista um espaço físico que cada pessoa leva em torno de dez segundos para atravessar, estima-se que o sistema se torna capaz de processar as passagens de até 1150 pessoas em um mesmo intervalo de tempo.

Já para os testes do *middleware*, o *EventCycle* do ALE foi definido para a cada dois segundos. Isso significa que a *tag* assim que for lida pelo leitor gera um registro no EPCIS e a cada dois segundos que a mesma permanecer no raio do leitor ela gera um novo registro. O *Capturing Application* foi configurado para armazenar todos os *EventCycles* gerados no EPCIS. Nos testes, cada requisição ao EPCIS pedia a recuperação dos últimos cem *EventCycles* registrados no banco de dados. Cada teste foi executado 10 vezes e o resultado considerado é a média do tempo das execuções. A Figura 49 apresenta os resultados apurados.

Figura 49 – Resultados dos testes de carga no *middleware Fosstrak*.



Fonte: Elaborada pelo autor.

Com base nos resultados, pode-se notar que o aumento do número de *tags* não apresenta um aumento significativo no tempo de resposta do EPCIS. O gráfico traz um tempo de resposta, em função da quantidade de *tags*, próximo ao linear, sugerindo que o *middleware* possui uma excelente performance, sendo capaz de lidar com requisições com muito mais *tags* sendo capturadas do que as testadas nesse trabalho. Em função do escopo do modelo HealCity, dificilmente terão mais do que 32 *tags* sendo lidas em fluxo constante ao mesmo tempo em um mesmo ambiente de saúde, de tal forma que os testes realizados são promissores para o desenvolvimento e execução do módulo Capture.

6.6 Discussão

Com base nas métricas estabelecidas, pôde-se notar que o modelo foi capaz de melhorar o desempenho dos ambientes de saúdes simulados na cidade inteligente proposta. A Tabela 11 apresenta todos os resultados encontrados em cada uma das duas métricas de avaliação utilizadas nestes testes preliminares. Conforme proposto na metodologia de avaliação, esperava-se que o tempo máximo de espera apresentassem uma diminuição entre os cenários C1, C2 e C3, e isso de fato ocorreu, cumprindo o objetivos para essa métrica. Além disso, o custo de recursos humanos e o número elástico de recursos humanos usados aumentaram entre os cenários C1/C2 e C3, conforme esperado. Dessa forma, pode-se afirmar que os resultados esperados na metodologia de avaliação foram alcançados através do uso do modelo HealCity na cidade inteligente proposta.

Para a métrica de tempo máximo de espera, haviam dois objetivos: a diminuição do tempo de espera nos ambientes de saúde e o enquadramento do mesmo dentro do limite superior de 30 minutos estabelecido na metodologia de avaliação. Conforme já mostrado, o modelo HealCity foi capaz de diminuir significativamente o tempo de espera para a cidade inteligente proposta, com um tempo médio máximo de espera de 24,04 minutos para C3 (com desvio-padrão de 10,3 minutos). No entanto, quando observamos os tempos de espera mais longos encontrados durante o período de simulação, o limite superior foi excedido (111, 117, 79, 54 e 46,75 minutos

Tabela 11 – Métricas de avaliação e resultados encontrados, onde os melhores resultados no M1 são destacados em verde e os piores em vermelho. Para as métricas Custo de RH e Quantidade Elástica de RH, foi utilizada a média dos resultados hora-a-hora do período simulado.

Cenário		Tempo máximo de espera		Custo de RH	Quantidade elástica de RH
		Média	Maior		
C1	Ambiente de Saúde 1	197.56 (±92.7)	361	-	11
	Ambiente de Saúde 2	186.56 (±112.9)	377	-	11
	Ambiente de Saúde 3	212.31 (±97.2)	390	-	11
	Ambiente de Saúde 4	180.43 (±107.4)	360	-	11
	Média da cidade inteligente	194.21 (±73.9)	305.7	44	44
C2	Ambiente de Saúde 1	49.77 (±34.4)	104	-	11
	Ambiente de Saúde 2	38.48 (±28.2)	99	-	11
	Ambiente de Saúde 3	40.25 (±26.4)	92	-	11
	Ambiente de Saúde 4	36.70 (±30.9)	108	-	11
	Média da cidade inteligente	41.30 (±14.3)	48	44	44
C3	Ambiente de Saúde 1	38.85 (±35.1)	111	-	10.99
	Ambiente de Saúde 2	23.04 (±26.2)	117	-	11.69
	Ambiente de Saúde 3	18.00 (±17.8)	79	-	11.64
	Ambiente de Saúde 4	16.27 (±15.2)	54	-	11.78
	Média da cidade inteligente	24.04 (±10.3)	46.7	48.26	46,15

Fonte: Elaborada pelo autor.

nos Ambientes de Saúde 1, 2, 3 e 4, respectivamente). Isso provavelmente ocorreu devido ao número limitado de estações de atendimento disponíveis para a alocação de novos recursos humanos. Entretanto, pode-se dizer que os objetivos para a métrica foram alcançados, visto que o modelo HealCity foi capaz de diminui consideravelmente o tempo de espera (87,62% na média da cidade inteligente) e na média da cidade inteligente o tempo máximo de espera ficou abaixo dos 30 minutos estabelecidos como objetivo.

Para a métrica de custo de recursos humanos, esperava-se que entre os cenários propostos houvesse um aumento deste custo, e foi exatamente isso que ocorreu. O objetivo da métrica é justamente mostrar qual o aumento de custo necessário para atingir determinada redução no tempo de espera por atendimento de saúde em uma cidade inteligente. Dessa forma, os resultados mostram que o custo de recursos humanos aumentou de forma *inversamente proporcional* à diminuição do tempo de espera, conforme esperado. Além disso, era esperado que a proporção de redução do tempo de espera fosse superior a proporção do aumento de custo, e isso de fato ocorreu. Assim, ficou evidente que um mero aumento de 9,68% no custo resultou em uma diminuição de até 87,62% na média do tempo máximo de espera. Por fim, para a métrica de quantidade elástica de recursos humanos era previsto que número médio de profissionais de saúde utilizados na cidade inteligente aumentaria entre os cenários C1/C2 e C3, e esse foi de

fato o caso.

Com relação ao uso de sensores RFID para o modelo proposto, também existem evidências de que com o uso de um leitor *Intermec* IF2, equipado de quatro antenas, é possível identificar a entrada de pessoas em um ambiente se as antenas forem instaladas ao lado de portas simples ou duplas, com a mesma distância entre elas. Com relação ao processamento de dados recebidos do sistema RTLS, também existem evidências de que tanto o *middleware* escolhido, quanto o módulo Capture são capazes de atender a demanda de leituras geradas por um grupo significativo de pacientes.

Assim, o modelo teve resultados animadores, adicionando uma nova visão para a gestão de recursos humanos no contexto da saúde em cidades inteligentes. Ressalta-se que para o contexto proposto para o HealCity não é possível realizar uma análise quantitativa em comparação com os trabalhos relacionados, uma vez que os ambientes e contextos de cada trabalho são diferentes. Portanto, foi feita uma análise qualitativa comparando o modelo HealCity com os trabalhos relacionados, a fim de apresentar as características que o modelo propõe que o diferenciam de outros trabalhos na literatura, a fim de destacar a contribuição do HealCity para a área. A Tabela 12 apresenta a análise qualitativa proposta.

Tabela 12 – Avaliação qualitativa comparativa do modelo HealCity versus Trabalhos Relacionados.

Trabalho Relacionado	**Análise de sinais vitais**	**Alocação *on-the-fly***	**Análise do estado de saúde dos pacientes fora dos hospitais**	**Ajustes em vários ambientes de saúde**	**Ajustes proativos de recursos humanos**	**Utiliza sensores IoT**	**Escalabilidade é analisada**
Graham et al. (2018)	Não	Não	Não	Não	Não	Não	Não
Franke et al. (2018)	Não	Não	Não	**Sim**	Não	Não	Não
Oueida, Aloqaily e Ionescu (2019)	Não	Não	Não	Não	Não	Não	Não
Fischer et al. (2020)	Não	**Sim**	Não	Não	**Sim**	Não	Não
Apornak et al. (2021a)	Não	Não	Não	Não	Não	Não	Não
Apornak et al. (2021b)	Não	Não	Não	Não	Não	Não	Não
Hafezalkotob et al. (2022)	Não	Não	Não	**Sim**	Não	Não	Não
Liao, Chu e Ho (2022)	Não	Não	Não	Não	Não	Não	Não
Duma e Aringhieri (2023)	Não	**Sim**	Não	Não	Não	Não	Não
Lazebnik (2023)	Não	**Sim**	Não	Não	Não	Não	Não
Rasidi et al. (2024)	Não	Não	Não	Não	Não	Não	Não
HealCity	**Sim**	**Sim**	**Sim**	**Sim**	**Sim**	**Sim**	**Sim**

Fonte: Elaborada pelo autor.

Dessa forma, o modelo HealCity propôs uma abordagem para preencher todas as lacunas

encontradas em trabalhos relacionados, conforme segue:

- O HealCity integra dados capturados por sensores de IoT em cidades inteligentes e analisa o estado de saúde de pacientes fora dos hospitais para prever o *uso futuro de recursos* e a *sobrecarga de pacientes* em ambientes hospitalares em cidades inteligentes;
- O HealCity é capaz de identificar a demanda em um ambiente de saúde e propor *soluções concretas* para abordar proativamente a questão da escassez de recursos humanos *antes que a mesa aconteça*;
- O HealCity tem a capacidade de *otimizar processos* de forma automatizada, propondo ajustes *on-the- fly* para múltiplos ambientes de saúde; e
- O HealCity relaciona a admissão de pacientes em hospitais com os *sinais vitais* da população das cidades, integrando essas abordagens em um modelo de gestão de recursos humanos.

Com base nos dados levantados através das simulações e testes, pode-se notar que existem evidências do funcionamento do modelo HealCity e da sua capacidade de ajustar a força de trabalho de uma cidade inteligente à demanda de pacientes, diminuindo o tempo de espera por atendimento. Entretanto, levando em conta que os testes foram baseados em um ambiente simulado, é importante salientar que os resultados obtidos não podem ser generalizados. Assim, podemos dizer que para uma validação absoluta do sistema seria necessária a sua efetiva implantação em um ambiente de saúde, de forma a se capturar dados reais da demanda de pacientes para serem analisados. Então, do ponto de vista técnico, existem evidências do protótipo desenvolvido ser funcional, mas não podemos provar apenas com simulação a sua efetividade no mundo real.

7 CONCLUSÃO

Todos os dias, novas tecnologias estão sendo desenvolvidas para auxiliar nossas atividades diárias. O conceito de Internet das Coisas demonstra que até nossos objetos estão se integrando ao mundo virtual. Além disso, o progresso nas pesquisas possibilita que dados coletados por sensores físicos sejam transmitidos para os sistemas de computação de maneira cada vez mais rápida, precisa e transparente para o usuário. Também, a predição computacional é um conceito notável que está ganhando destaque com o avanço das tecnologias e possui um potencial imenso. Com sistemas computacionais cada vez mais poderosos e robustos, os sistemas conseguem analisar dados históricos e prever com alta confiança informações futuras para diversos tipos de aplicações. Durante o levantamento bibliográfico, foram estudadas e levantadas características das tecnologias utilizadas para Internet das Coisas, tais como a tecnologia RFID, *Real-Time Location Systems* e sensores de sinais vitais, bem como conceitos de cidades inteligentes e sua legislação aplicável, de forma a verificar se seria possível utilizar essas tecnologias para melhorar o fluxo de pacientes em ambientes de saúde. Também foram estudados os processos hospitalares mais conhecidos e técnicas utilizadas para predição de dados. A partir da análise de algumas abordagens existentes, percebe-se que a tecnologia não apenas possui o potencial necessário para melhorar o fluxo de pessoas em ambientes de saúde, como também já existem várias metodologias para gerenciamento nesses ambientes. No entanto, nenhuma das abordagens trabalha com ajustes dinâmicos de recursos humanos para múltiplos ambientes de saúde.

Nesse contexto esse trabalho propôs o modelo HealCity para gerenciamento elástico de recursos humanos em cidades inteligentes. O modelo HealCity vai além dos trabalhos relacionados ao propor o uso da elasticidade para antecipar problemas futuros, fornecendo um modelo para otimização de recursos humanos visando melhorar a experiência dos pacientes. Na Seção 1.2 foi definida a seguinte questão de pesquisa: *Quais estratégias um modelo de gerenciamento de recursos humanos para saúde precisa empregar para identificar a demanda por serviços de saúde e propor uma melhor alocação tanto de recursos humanos quanto de atendimento, independente da quantidade de pacientes?* Após o desenvolvimento do modelo HealCity, ela pode ser respondida da seguinte forma: Um modelo que possa prever a demanda por serviços de saúde e sugerir uma melhor distribuição de recursos humanos pode ser desenvolvido empregando tecnologias capazes de monitorar a movimentação e os sinais vitais de profissionais de saúde e pacientes em diferentes ambientes de saúde. Ao registrar a movimentação dos pacientes e os tempos de permanência em cada ponto de atendimento, pode-se determinar a demanda por tratamento médico, além de identificar possíveis gargalos no sistema de atendimento. Com esses dados, um modelo preditivo pode estimar a futura demanda de pacientes, permitindo a identificação de novos gargalos no atendimento clínico. Baseado nesses dados, esse modelo pode prever a capacidade futura de atendimento dos profissionais de saúde e propor soluções para a realocação de pessoal que atendam à demanda futura em uma cidade

inteligente. Isso amplia a aplicação do paradigma da elasticidade da computação em nuvem para gestão de recursos humanos, permitindo uma distribuição flexível e dinâmica da equipe médica em ambientes hospitalares localizados em cidades inteligentes.

Na Seção 1.2 foi definida a seguinte Hipótese para esta pesquisa: *Um modelo para controle de recursos humanos em ambientes de saúde de cidades inteligentes é capaz de diminuir o tempo de espera por atendimento nos ambientes de saúde, sem aumentar proporcionalmente o custo, mantendo este tempo adequado à legislação pertinente, de acordo com o número de pacientes nas cidades*. Com base nos resultados alcançados, existem evidências da validação da hipótese visto que através da simulação de uma cidade inteligente, destaca-se que o sistema foi capaz de atingir uma melhoria significativa em tempos de redução do tempo de espera (87,62% e 41,79% na média da cidade inteligente, com relação a cidade sem nenhuma estratégia de gerenciamento de recursos humanos e com o uso de uma abordagem existente na literatura, respectivamente), com um aumento mínimo no custo (9,68%). Além disso, foi possível enquadrar o tempo médio de espera (24,04 minutos) aos limites propostos em projeto de legislação sobre o tema (30 minutos).

7.1 Lições Aprendidas e Transferência de Tecnologia

O modelo HealCity mostrou que usar uma abordagem baseada em elasticidade permite reduzir os tempos de espera para atendimento médico em cidades inteligentes. Para fazer isso, os pacientes não precisariam realizar nenhuma atividade extra nos ambientes de saúde, mas simplesmente usar um sensor vestível para coletar sinais vitais e sua localização aproximada. O HealCity pode ser usado para identificar gargalos no fluxo de atendimento ao paciente e otimizar processos em ambientes de assistência médica. Apesar dos resultados promissores, há algumas *necessidades* que devem ser levadas em consideração ao implementar o modelo HealCity em uma cidade inteligente real: (i) tanto a equipe médica quanto os pacientes devem usar seus sensores vestíveis enquanto estiverem na cidade inteligente; (ii) uma vez que o HealCity envia apenas notificações aos recursos humanos, a equipe médica na cidade inteligente deve seguir estas recomendações; e (iii) é necessária a instalação da infraestrutura de sensores na cidade inteligente para que o sistema funcione corretamente.

O Art. 469 da legislação trabalhista brasileira, permite que que empregadores possam compartilhar seus empregados entre suas empresas. Porém, para implementar um modelo nesse nível de compartilhamento de recursos humanos, além de ajustes na legislação, deve haver uma política pública de compartilhamento de profissionais entre ambientes de saúde de diferentes redes hospitalares, ou todos os ajustes em escala regional ficam limitados apenas entre profissionais de uma rede comum de funcionários. Salienta-se que a lei acima aplica-se apenas ao Brasil, visto que outros países possuem suas próprias leis sobre o tema. Nos Estados Unidos, por exemplo, inexiste uma lei que impeça este compartilhamento de profissionais. Além disso, conforme a Lei nº 13.709 (BRASIL, 2018), conhecida como Lei Geral de Proteção de Dados

Pessoais (LGPD), em seu Art. 11, somente seria permitido o compartilhamento de dados dos cidadãos através do consentimento expresso desses cidadãos. A legislação até permite o compartilhamento em casos relacionados a saúde (Conforme a Lei nº 13.853 (BRASIL, 2019) que alterou alguns pontos na LGPD), porém isso permitiria apenas o compartilhados dos sinais vitais, excluindo-se localização e outros dados necessários ao modelo VitalSense. A mesma lei também permite o compartilhamento de dados necessários à execução de políticas públicas previstas em leis ou regulamentos. Dessa forma, respeitando-se a LGPD, também devem haver políticas públicas que permitam a captura e o compartilhamento de dados entre os atores envolvidos. Novamente salienta-se que a lei específica acima aplica-se apenas ao Brasil. Entretanto, em países da União Europeia, por exemplo, seriam necessárias as mesmas políticas visto similaridade da *General Data Protection Regulation* (GPDR) (KNUT BLIND; RAMMER, 2024) com a lei brasileira nestes pontos. Também é necessário, para o bom funcionamento do modelo proposto, recompensar os funcionários que estão dispostos a participar do modelo de alocação com uma remuneração maior, o que seria benéfico de duas maneiras: o público receberia melhores serviços e os profissionais de saúde seriam melhor remunerados. Além disso, a equipe médica precisa ser treinada para entender a importância de seguir as recomendações propostas pelo modelo. Além disso, os gestores municipais precisam investir na infraestrutura de monitoramento proposta pelo modelo VitalSense, e os gestores hospitalares na infraestrutura proposta pelos ambientes de saúdes propostos pelo HealCity. O grande incentivo para políticas públicas e investimento em infraestrutura é a possibilidade de se salvar vidas.

Inicialmente, vislumbra-se que a implementação do modelo HealCity deveria ocorrer em uma rede pública municipal de saúde, dessa forma todos os profissionais de saúde são servidores públicos municipais. Um exemplo seria a implementação do modelo proposto na cidade de Porto Alegre, por meio da rede municipal de saúde, que atualmente é composta por 132 centros de saúde, 4 clínicas da família (PORTO ALEGRE, 2024a), 4 unidades de pronto atendimento, dois hospitais de gestão pública, dois hospitais com gestão compartilhada com a rede privada (PORTO ALEGRE, 2024b) e 4000 funcionários (PORTO ALEGRE, 2024c). Neste cenário, o modelo apresenta um grande potencial, podendo impactar políticas públicas para facilitar a mobilidade de profissionais e fomentar investimentos em infraestrutura tecnológica, melhorando assim a qualidade do atendimento ao paciente por meio da simplificação da gestão de recursos humanos em hospitais.

A cidade de Porto Alegre não possui uma infraestrutura de cidade inteligente conforme proposta neste modelo, dessa forma para implantação do modelo no estágio atual de desenvolvimento tecnológico da cidade são necessários alguns ajustes. Primeiramente, sem infraestrutura de sensores e o modelo VitalSense, o HealCity se torna incapaz de sugerir o ambiente de saúde adequado para um novo paciente, porém ainda é capaz de realizar a elasticidade de recursos humanos proposta. Para tanto, ao invés de serem utilizados sensores em todos os pacientes com monitoramento ativo de sinais vitais, os mesmos podem ser substituído por *tags* RFID, muito mais baratas para uso em larga escala. Dessa forma com a instalação de leitores portais nas salas

dos ambientes de saúde é possível monitorar o fluxo de pacientes e com isso alimentar o modelo de elasticidade de pessoas. Além disso, em um cenário em que nem a instalação desses leitores e *tags* RFID seria possível, o modelo de elasticidade poderia ser alimentado manualmente pelos funcionários do hospital, conforme a movimentação de pacientes. Assim sempre que um paciente passar pelo primeiro atendente, o mesmo já estaria inserido na primeira fila e a cada atendimento ele seria movido manualmente no sistema da fila para a sala de atendimento e após para a próxima fila. Dessa forma, o modelo de gerenciamento de recursos humanos poderia ser posto em prática com praticamente nenhum custo de infraestrutura física.

7.2 Contribuições Esperadas

O modelo HealCity visa preencher as lacunas encontradas no estado da arte por meio da investigação de trabalhos relacionados. Dessa forma, este estudo procura oferecer tanto contribuições científicas quanto benefícios para a sociedade, por meio de ambientes de saúde em cidades inteligentes que adotem o modelo proposto. A principal inovação do modelo HealCity é a proposta de um modelo para gestão de pessoas *on-the-fly* no contexto de múltiplos ambientes de saúde. Assim, o modelo propõe três contribuições científicas, conforme listado abaixo:

(i) Com a finalidade de criar ambientes de saúde em cidades inteligentes que se adaptam às variações na demanda de pacientes, este estudo apresenta o conceito **Elasticidade de Recursos Humanos para Saúde em Cidades Inteligentes**, que inclui um algoritmo para gerenciamento dinâmico da distribuição de recursos humanos em ambientes de saúde, fazendo uso de sensores de sinais vitais para monitorar a demanda de problemas de saúde dos pacientes, propondo algumas métricas de avaliação de ambientes de saúde em cidades inteligentes;

(ii) Com o propósito de permitir o crescimento sustentável da solução proposta, de forma a ser possível adicionar ou remover quantos ambiente de saúde forem necessários no sistema, foi proposta a definição de uma **arquitetura hierárquica escalável para sistemas de gerenciamento de recursos humanos de hospitais em cidades inteligentes**; e

(iii) Este trabalho introduz um novo algoritmo para gerenciamento de futuros pacientes na cidade inteligente permitindo a identificação do ambiente de saúde apropriado para um possível novo paciente, propondo alguns novos formalismos matemáticos.

Por fim, com relação as contribuições para a sociedade, podem ser destacadas:

(i) Redução do tempo de espera por atendimento nos ambientes de saúde;

(ii) Otimização do tempo dos profissionais de saúde, balanceando a carga de atendimento;

(iii) Identificação de antemão de qual ambiente de saúde está preparado para atender determinado paciente;

(iv) Possibilidade de melhoria do planejamento hospitalar com base nos dados fornecidos pelo modelo;

(v) Melhor tratamento das necessidades de atendimento médico;

7.3 Limitações

Nesta seção serão elencadas algumas das limitações que podem ser encontrados para implementação do modelo HealCity em um ambiente de saúde real.

- Antes de colocar o modelo em produção, é essencial discutir o que a elasticidade proposta significa em termos de legislação trabalhista;
- Tanto os funcionários quanto os pacientes devem obrigatoriamente portar suas respectivas *tags* de identificação durante toda sua permanência nas instalações dos ambientes de saúde, assim como os sensores de sinais vitais enquanto permanecerem na cidade inteligente;
- Funcionários podem considerar o modelo proposto invasivo, já que ele monitora constantemente as atividades dos profissionais de saúde;
- O HealCity apenas envia notificações para que os pacientes se dirijam aos ambientes de saúde e para que os recursos humanos sejam realocados, mas a movimentação real das pessoas nos ambientes da cidade depende de cada indivíduo escolher seguir ou não a orientação;
- A implementação de um sistema RTLS com a precisão necessária em todas as áreas do ambiente de saúde, bem como da infraestrutura de monitoramento na cidade inteligente, pode ser altamente custosa, tanto em termos de tempo quanto financeiramente.

7.4 Trabalhos Futuros

A título de trabalhos futuros vislumbra-se utilizar os dados gerados neste trabalho, especialmente aqueles relacionados às taxas de ocupação de profissionais de saúde, para propor novos ajustes na alocação de profissionais. A ideia é levar em conta não apenas o tempo de espera, mas também a quantidade de pessoas esperando, permitindo que as salas fiquem sem equipe médica se não houver pacientes esperando, melhorando o balanceamento de carga entre os profissionais. Além disso, pretende-se ajustar a estratégia de *cooldown-period* para levar em conta o risco das salas, permitindo que salas com maior urgência pulem esse período para melhorar o tempo de resposta. Além disso, como o modelo inclui uma interface de usuário no módulo HealCity App, espera-se realizar um estudo de *feedback* do usuário para demonstrar a usabilidade e aceitabilidade do sistema proposto. Vislumbra=se também o uso de uma abordagem

adaptativa para definição dos intervalos de processamento do módulo *Predict*. Outra possibilidade de trabalhos futuros seria o uso de *smartphones* como controladores *Edge*, dessa forma simplificando processos de localização de pessoas, utilizando o GPS do próprio aparelho e simplificando a instalação em larga escala, visto redução da necessidade de um hardware dedicado para processamentos dos dados mais próximos aos usuários. Por fim, pretende-se utilizar a elasticidade de recursos humanos em cidades inteligentes para domínios diferentes da área da saúde. Assim, espera-se que o algoritmo seja adaptado e expandido de forma que o mesmo possa ser aplicado em outros ambientes existentes em cidades inteligentes que prestem atendimento à pessoas, uma vez que o modelo proposto poderia ser usado em outras situações com poucas adaptações.

7.5 Publicações

Ao longo de todo o período de pesquisa do doutorado, foram produzidos alguns artigos científicos para publicação em periódicos e anais de eventos. A seguir são listados os artigos publicados ou aceitos para publicação.

- Artigos publicados em **revistas**:

  1. **Título:** Multi-Hospital Management: Combining Vital Signs IoT Data and the Elasticity Technique to Support Healthcare 4.0
     Revista: IoT
     Métricas: *Google H5-Index:* 21 - *Scopus CiteScore:* 8.5 - *Qualis calculado:* A1
     DOI: `https://doi.org/10.3390/iot5020019`
     Citação: FISCHER, G. S.; RAMOS, G. O.; COSTA, C. A.; ALBERTI, A.; GRIEBLER, D.; SINGH, D.; RIGHI, R. R. *Multi-Hospital Management: Combining Vital Signs IoT Data and the Elasticity Technique to Support Healthcare 4.0*. IoT. v.5, p. 381-408, 2024.

  2. **Título:** CityHealth: otimizando a gestão hospitalar em cidades inteligentes através da técnica de elasticidade e análise de sinais vitais
     Revista: Brazilian Journal of Health Review (BJHR)
     Métricas: *Google H5-Index:* 30 - *IJIFactor:* 2.91 - *Qualis 2017-2020:* B3
     DOI: `https://doi.org/10.34119/bjhrv7n2-300`
     Citação: FISCHER, G. S.; COSTA, C. A.; RIGHI, R. R. *CityHealth: otimizando a gestão hospitalar em cidades inteligentes através da técnica de elasticidade e análise de sinais vitais*. Brazilian Journal of Health Review (BJHR), v. 7, p. e68709, 2024.

3. **Título:** Looking at Smart Cities Through the Lens of a Pandemic Era: A systematic Literature Review
   Revista: International Journal of Technology Management (IJTM)
   Métricas: *Google H5-Index:* 19 - *Scopus CiteScore:* 2.7 - *Qualis calculado:* A2
   DOI: `https://doi.org/10.1504/IJTM.2024.136418`
   Citação: FISCHER, G. S.; RODRIGUES, V. F.; RIGHI, R. R.; COSTA, C. A. ; POLICARPO, L. M.; SILVA, R. G. G. *Looking at Smart Cities Through the Lens of a Pandemic Era: A systematic Literature Review*. International Journal of Technology Management (IJTM), v. 94, p. 342-384, 2024.

- Artigos publicados em **anais de eventos**:

4. **Título:** Combinando Elasticidade Proativa e Reativa para Gestão Multi-Hospitalar de Recursos Humanos no Contexto das Cidades Inteligentes
   Revista: Anais do XXIV Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2024)
   Métricas: *Google H5-Index:* 7 - *Qualis eventos:* A4
   DOI: `https://doi.org/10.5753/sbcas.2024.2179`
   Citação: FISCHER, G. S.; RIGHI, R. R.; COSTA, C. A.; ROEHRS, A. *Combinando Elasticidade Proativa e Reativa para Gestão Multi-Hospitalar de Recursos Humanos no Contexto das Cidades Inteligentes*. In: XXIV Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2024), 2024, Goiânia. Anais do XXIV Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2024), 2024. p. 268-279.

5. **Título:** Usando Dados de Sinais Vitais dos Cidadãos e a Técnica de Elasticidade para gerência de Hospitais no Contexto de Cidades Inteligentes
   Revista: Anais do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2023)
   Métricas: *Google H5-Index:* 7 - *Qualis eventos:* A4
   DOI: `https://doi.org/10.5753/sbcas.2023.229666`
   Citação: FISCHER, G. S.; COSTA, C. A. ; RIGHI, R. R. *Usando Dados de Sinais Vitais dos Cidadãos e a Técnica de Elasticidade para gerência de Hospitais no Contexto de Cidades Inteligentes*. In: XXIII Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2023), 2023, São Paulo. Anais do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2023), 2023. p. 268-279.

6. **Título:** Utilizando IoT para Gerenciamento Elástico e Eficiente de Recursos em Hospitais Inteligentes
   Revista: Anais da XXII Escola Regional de Alto Desempenho da Região Sul (ERAD-

RS 2022)
Métricas: *Google H5-Index:* 3
DOI: `https://doi.org/10.5753/eradrs.2022.19161`
Citação: FISCHER, G. S.; RIGHI, R. R. *Utilizando IoT para Gerenciamento Elástico e Eficiente de Recursos em Hospitais Inteligentes*. In: Escola Regional de Alto Desempenho da Região Sul, 2022, Brasil. Anais da XXII Escola Regional de Alto Desempenho da Região Sul (ERAD-RS 2022), 2022. p. 59.

7. **Título:** Usando IoT e Conceitos de Elasticidade em Cloud para Análise de Recursos Humanos em Hospitais Inteligentes
   Revista: Anais Principais do XX Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2020)
   Métricas: *Google H5-Index:* 7 - *Qualis eventos:* A4
   DOI: `https://doi.org/10.5753/sbcas.2020.11516`
   Citação: FISCHER, G. S.; COSTA, C. A. ; RIGHI, R. R. *Usando IoT e Conceitos de Elasticidade em Cloud para Análise de Recursos Humanos em Hospitais Inteligentes*. In: XX Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2020), 2020, Salvador. Anais Principais do XX Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2020). Porto Alegre: Sociedade Brasileira de Computação, 2020. p. 226-237.

- Capítulos de livros publicados:

8. **Título:** Use of Internet of Things with Data Prediction on Healthcare Environments: a Survey
   Livro: Research Anthology on Convergence of Blockchain, Internet of Things, and Security
   Métricas: -
   DOI: `https://doi.org/10.4018/978-1-6684-7132-6.ch043`
   Citação: FISCHER, G. S.; RIGHI, R. R.; RODRIGUES, V. F.; COSTA, C. A. *Use of Internet of Things with Data Prediction on Healthcare Environments: a Survey*. Research Anthology on Convergence of Blockchain, Internet of Things, and Security. 1ed.: Management Association, Information Resources, 2023, v. 1, p. 778-798.

REFERÊNCIAS

AHAMED, Z.; KHEMAKHEM, M.; EASSA, F.; ALSOLAMI, F.; AL-GHAMDI, A. S. A.-M. Technical Study of Deep Learning in Cloud Computing for Accurate Workload Prediction. **Electronics**, Basel, Switzerland, v. 12, n. 3, 2023.

AKEJU, O.; BUTAKOV, S.; AGHILI, S. Main factors and good practices for managing BYOD and IoT risks in a K-12 environment. **International Journal of Internet of Things and Cyber-Assurance**, Olney, England, v. 1, n. 1, p. 22–39, Jan. 2018.

AKRAM, A. S.; ABBAS, S.; KHAN, M. A.; ATHAR, A.; GHAZAL, T. M.; HAMADI, H. A. Smart Energy Management System Using Machine Learning. **Computers, Materials & Continua**, Henderson, USA, v. 78, n. 1, p. 959–973, 2024.

AL-DHURAIBI, Y.; PARAISO, F.; DJARALLAH, N.; MERLE, P. Elasticity in Cloud Computing: state of the art and research challenges. **IEEE Transactions on Services Computing**, USA, v. PP, n. 99, p. 1–1, 2017.

AL-HAIDARI, F.; SQALLI, M.; SALAH, K. Impact of CPU Utilization Thresholds and Scaling Size on Autoscaling Cloud Resources. In: IEEE 5TH INTERNATIONAL CONFERENCE ON CLOUD COMPUTING TECHNOLOGY AND SCIENCE, 2013., 2013, Bristol, UK. **Anais...** IEEE, 2013. v. 2, p. 256–261.

ALI, Z.; MAHMOOD, A.; KHATOON, S.; ALHAKAMI, W.; ULLAH, S. S.; IQBAL, J.; HUSSAIN, S. A Generic Internet of Things (IoT) Middleware for Smart City Applications. **Sustainability**, Basel, Switzerland, v. 15, n. 1, 2023.

ALSHAMAILA, Y.; PAPAGIANNIDIS, S.; ALSAWALQAH, H.; ALJARAH, I. Effective use of smart cities in crisis cases: a systematic review of the literature. **International Journal of Disaster Risk Reduction**, Netherlands, v. 85, p. 103521, 2023.

APORNAK, A.; RAISSI, S.; KERAMATI, A.; KHALILI-DAMGHANI, K. Optimizing human resource cost of an emergency hospital using multi-objective Bat algorithm. **International Journal of Healthcare Management**, United Kingdom, v. 14, n. 3, p. 873–879, 2021.

APORNAK, A.; RAISSI, S.; KERAMATI, A.; KHALILI-DAMGHANI, K. Human resources optimization in hospital emergency using the genetic algorithm approach. **International Journal of Healthcare Management**, United Kingdom, v. 14, n. 4, p. 1441–1448, 2021.

ATTARAN, H.; KHEIBARI, N.; BAHREPOUR, D. Toward integrated smart city: a new model for implementation and design challenges. **GeoJournal**, Netherlands, Jan. 2022.

AZEVEDO, A. L.; PEREIRA, A.; LEMOS, C.; COELHO, M.; CHAVES, L. Organização de serviços de emergência hospitalar: uma revisão integrativa de pesquisas. **Revista Eletrônica de Enfermagem**, Goiânia, GO, Brasil, v. 12, n. 4, p. 736–45, 2010.

BAUCAS, M. J.; SPACHOS, P.; PLATANIOTIS, K. N. Federated Learning and Blockchain-Enabled Fog-IoT Platform for Wearables in Predictive Healthcare. **IEEE Transactions on Computational Social Systems**, United States, v. 10, n. 4, p. 1732–1741, 2023.

BENDAVID, Y.; ROSTAMPOUR, S.; BERRABAH, Y.; BAGHERI, N.; SAFKHANI, M. The Rise of Passive RFID RTLS Solutions in Industry 5.0. **Sensors**, Basel, Switzerland, v. 24, n. 5, 2024.

BERTONI, A. P. S.; RODRIGUES, V. F.; ZEISER, F. A.; MELLO, B.; COSTA, C. A. da; DONIDA, B.; RIGO, S. J.; ROSA RIGHI, R. da. Internet das Coisas de Saúde: aplicando iot, interoperabilidade e aprendizado de máquina com foco no paciente. In: OLIVEIRA, L. F. de; ARAúJO, F. H. D. de (Ed.). **MINICURSOS DO XXII SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO APLICADA À SAÚDE**. Porto Alegre: SBC, 2022. p. 1–476.

BITTAR, O. J. N. V. Gestão de processos e certificação para qualidade em saúde. **Revista da Associação Médica Brasileira**, São Paulo, SP, Brazil, v. 46, p. 70 – 76, mar. 2000.

BIYA, M.; GEZAHAGN, M.; BIRHANU, B.; YITBAREK, K.; GETACHEW, N.; BEYENE, W. Waiting time and its associated factors in patients presenting to outpatient departments at Public Hospitals of Jimma Zone, Southwest Ethiopia. **BMC Health Services Research**, United Kingdom, v. 22, n. 1, p. 107, Jan. 2022.

BOULOS, M. N. K.; BERRY, G. Real-time locating systems (RTLS) in healthcare: a condensed primer. **International Journal of Health Geographics**, England, v. 11, n. 25, p. 1–8, 2012.

BRASIL. **Decreto-Lei nº 5.452, de 1º de maio de 1943.** Aprova a Consolidação das Leis do Trabalho. Disponível em: <http://www.planalto.gov.br/ccivil_03/Decreto-Lei/Del5452.htm>. Acesso em: 25 jul. 2022.

BRASIL. **Lei nº 13.467, de 13 de julho de 2017.** Altera a Consolidação das Leis do Trabalho (CLT), aprovada pelo Decreto-Lei no 5.452, de 1º de maio de 1943, e as Leis nº 6.019, de 3 de janeiro de 1974, 8.036, de 11 de maio de 1990, e 8.212, de 24 de julho de 1991, a fim de adequar a legislação às novas relações de trabalho. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2017/lei/L13467.htm>. Acesso em: 25 jul. 2022.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018.** Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 10 ago. 2024.

BRASIL. **Lei nº 13.853, de 8 de julho de 2019.** Altera a Lei nº 13.709, de 14 de agosto de 2018, para dispor sobre a proteção de dados pessoais e para criar a Autoridade Nacional de Proteção de Dados; e dá outras providências. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/lei/L13853.htm>. Acesso em: 10 ago. 2024.

BUTEAN, A.; DAVID, A.; BUDULECI, C.; DAIAN, A. Auxilum Medicine: a cloud based platform for real-time monitoring medical devices. In: INTERNATIONAL CONFERENCE ON CONTROL SYSTEMS AND COMPUTER SCIENCE, 2015., 2015, Bucharest, Romania. **Anais...** IEEE, 2015. p. 874–879.

CAPOCCI, N. R.; NASCIMENTO, B. S.; LOPES, F. B.; BUENO, M. J. C. Simulação de Eventos Discretos Aplicado em uma Policlínica. **INOVAE - Journal of Engineering, Architecture and Technology Innovation**, São Paulo, SP, Brasil, v. 4, n. 1, p. 11–22, 2016.

COUTINHO, A. A. P.; CECÍLIO, L. C. d. O.; MOTA, J. A. C. Classificação de risco em serviços de emergência: uma discussão da literatura sobre o Sistema de Triagem de Manchester. **Revista Médica de Minas Gerais**, Belo Horizonte, MG, Brasil, v. 22, n. 2, p. 188–198, Mar. 2012.

DAI, F.; TEMLYAKOV, V. Random points are good for universal discretization. **Journal of Mathematical Analysis and Applications**, Netherlands, v. 529, n. 1, p. 127570, 2024.

DENG, L.; CHENG, F.; GAO, X.; YU, W.; SHI, J.; ZHOU, L.; ZHANG, L.; LI, M.; WANG, Z.; ZHANG, Y.-D.; LV, Y. Hospital crowdedness evaluation and in-hospital resource allocation based on image recognition technology. **Scientific Reports**, London, United Kingdom, v. 13, n. 1, p. 299, Jan 2023.

DUMA, D.; ARINGHIERI, R. Real-time resource allocation in the emergency department: a case study. **Omega**, United Kingdom, v. 117, p. 102844, 2023.

EGHMAZI, A.; ATAEI, M.; LANDRY, R. J.; CHEVRETTE, G. Enhancing IoT Data Security: using the blockchain to boost data integrity and privacy. **IoT**, Basel, Switzerland, v. 5, n. 1, p. 20–34, 2024.

ELASSY, M.; AL-HATTAB, M.; TAKRURI, M.; BADAWI, S. Intelligent transportation systems for sustainable smart cities. **Transportation Engineering**, USA, v. 16, p. 100252, 2024.

ENGLISH, B. **Hospital overcrowding continues despite increased ED staffing**. **Limerick Post**, Limerick, Ireland, 20 july 2024. Disponível em: <https://www.limerickpost.ie/2024/07/20/hospital-overcrowding-continues-despite-increased-ed-staffing/>. Acesso em: 12 ago. 2024.

FISCHER, G. S.; RAMOS, G. d. O.; COSTA, C. A. d.; ALBERTI, A. M.; GRIEBLER, D.; SINGH, D.; RIGHI, R. d. R. Multi-Hospital Management: combining vital signs iot data and the elasticity technique to support healthcare 4.0. **IoT**, Basel, Switzerland, v. 5, n. 2, p. 381–408, 2024.

FISCHER, G. S.; ROSA RIGHI, R. da; OLIVEIRA RAMOS, G. de; COSTA, C. A. da; RODRIGUES, J. J. ElHealth: using internet of things and data prediction for elastic management of human resources in smart hospitals. **Engineering Applications of Artificial Intelligence**, United Kingdom, v. 87, p. 103285:1–14, 2020.

FOSSTRAK. **Fosstrak**: open source rfid software platform. 2017. Disponível em: <https://fosstrak.github.io/>. Acesso em: 23 abr. 2018.

FRANKE, H.; HASSELER, M.; DICK, D.; KREBS, S. A Sustainable Approach for a Smart Human Resource Management in Healthcare. **European Journal of Sustainable Development**, Italy, v. 7, n. 3, p. 403, Oct. 2018.

GOIANIA. **Projeto de Lei 2018/000159, de 14 de junho de 2018**. Estipula o tempo máximo de espera para atendimento em hospitais, clínicas e laboratórios da iniciativa privada estabelecidos em Goiânia. Autoria: Vereador Zander Fabio, Goiania, GO: Câmara Municipal de Goiania, [2018]. Disponível em: <http://goiania.go.leg.br/sala-de-imprensa/noticias/projeto-estipula-tempo-maximo-de-espera-para-atendimento-em-hospitais>. Acesso em: 25 jul. 2022.

Google Maps Platform. **Distance Matrix API overview**. 2024. Disponível em: <https://developers.google.com/maps/documentation/distance-matrix/overview>. Acesso em: 21 jul. 2024.

GRAHAM, B.; BOND, R.; QUINN, M.; MULVENNA, M. Using Data Mining to Predict Hospital Admissions From the Emergency Department. **IEEE Access**, New York, NY, USA, v. 6, p. 10458–10469, Feb. 2018.

GRUPO PORTUGUÊS DE TRIAGEM. **Triagem no Serviço de Urgência**: Protocolo de Triagem de Manchester. 2. ed. Lisboa, Portugal: Blackwell Publishing Ltd., 2010.

GS1 EPCGLOBAL. **EPC™ Radio-Frequency Identity Protocols Generation-2 UHF RFID**: protocol for communications at 860 MHz – 960 MHz. Brussels, 2016.

HAFEZALKOTOB, A.; FARDI, K.; AICKELIN, U.; CHAHARBAGHI, S.; Akbarzadeh Khorshidi, H. A cooperative robust human resource allocation problem for healthcare systems for disaster management. **Computers & Industrial Engineering**, United Kingdom, v. 170, p. 108283, 2022.

INTERNATIONAL ORGANIZATION FOR STARDARTIZATION (ISO); INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC). **ISO/IEC 24730**: information technology: real-time locating systems (RTLS). 2th ed. Geneva, 2014.

INTERNATIONAL TELECOMMUNICATION UNION (ITU-T). **Smart sustainable citites**: an analysis of definitions. ITU-T focus group on smart sustainable cities. 1st ed. Geneva: ITU-T, 2014 (Technical report).

JAIN, H.; PATEL, R. Analysis & forecasting of juvenile crime using variance threshold and time series algorithm. **Multimedia Tools and Applications**, Netherlands, Jul 2024.

JANANI, S.; R., D.; TALUKDAR, S. B.; TALUKDAR, V. B.; MISHRA, K. N.; DHABLIYA, D. IoT and Machine Learning in Smart City Healthcare Systems. In: PRAMANIK, S.; SAGAYAM, K. M. (Ed.). **Handbook of Research on Data-Driven Mathematical Modeling in Smart Cities**. Hershey, PA, USA: IGI Global, 2023. p. 262–279.

KALASIN, S.; SURAREUNGCHAI, W. Challenges of Emerging Wearable Sensors for Remote Monitoring toward Telemedicine Healthcare. **Analytical Chemistry**, Washington, DC,USA, v. 95, n. 3, p. 1773–1784, Jan 2023.

KASSIM, M.; SALLEH, A. S.; SHAHBUDIN, S.; YUSOFF, M.; KAMALUDDIN, N. A. IoT Bus Tracking System Localization via GPS-RFID. In: IEEE INTERNATIONAL CONFERENCE IN POWER ENGINEERING APPLICATION (ICPEA), 2022., 2022, Shah Alam, Malaysia. **Anais...** IEEE, 2022. p. 1–6.

KNUT BLIND, C. N.; RAMMER, C. The impact of the EU General data protection regulation on product innovation. **Industry and Innovation**, United Kingdom, v. 31, n. 3, p. 311–351, 2024.

KOMALASARI, R. Telemedicine in Pandemic Times in Indonesia: healthcare professional's perspective. In: VAJJHALA, N. R.; EAPPEN, P. (Ed.). **Health Informatics and Patient Safety in Times of Crisis**. Hershey, PA, USA: IGI Global, 2023. p. 138–153.

KUMAR, R. Application of Smart Connected-Home Sensors Towards IoT Service. In: AL-TURJMAN, F. (Ed.). **AIoT and Smart Sensing Technologies for Smart Devices**. Hershey, PA, USA: IGI Global, 2024. p. 103–126.

LAZEBNIK, T. Data-driven hospitals staff and resources allocation using agent-based simulation and deep reinforcement learning. **Engineering Applications of Artificial Intelligence**, USA, v. 126, p. 106783, 2023.

LIAO, P.-H.; CHU, W.; HO, C.-S. An Analysis of Waiting Time for Emergency Treatment and Optimal Allocation of Nursing Manpower. **Healthcare**, Basel, Switzerland, v. 10, n. 5, 2022.

LOMBARDI, M.; PASCALE, F.; SANTANIELLO, D. Internet of Things: a general overview between architectures, protocols and applications. **Information**, Basel, Switzerland, v. 12, n. 2, 2021.

LUCIC, M. C.; BOUHAMED, O.; GHAZZAI, H.; KHANFOR, A.; MASSOUD, Y. Leveraging UAVs to Enable Dynamic and Smart Aerial Infrastructure for ITS and Smart Cities: an overview. **Drones**, Basel, Switzerland, v. 7, n. 2, 2023.

MALTI, A. N.; HAKEM, M.; BENMAMMAR, B. A new hybrid multi-objective optimization algorithm for task scheduling in cloud systems. **Cluster Computing**, Netherlands, Jul 2023.

MARQUES, B. S.; RIGHI, R. R. Análise da Escalabilidade de Middleware Para Internet das Coisas em Ambientes Centralizados e Distribuídos. In: XV ESCOLA REGIONAL DE ALTO DESEMPENHO DO ESTADO DO RIO GRANDE DO SUL (ERAD 2015), 2015, Gramado. **Anais...** Biblioteca Digital Brasileira de Computação (BDBComp), 2015.

MCCARTHY, M. L. Overcrowding in emergency departments and adverse outcomes. **BMJ**, London, United Kingdon, v. 342, 2011.

MOREIRA, H. **Drama da superlotação**: o que relatam pacientes que aguardam atendimento nos hospitais públicos de campinas. **g1**, Campinas, 13 mar. 2024. Disponível em: <https://g1.globo.com/sp/campinas-regiao/noticia/2024/03/13/drama-da-superlotacao-o-que-relatam-pacientes-que-aguardam-atendimento-nos-hospitais-publicos-de-campinas.ghtml>. Acesso em: 12 ago. 2024.

MOREIRA, P. A. T. **Gestão de Controlo de Acessos**. 2008. 241 p. Dissertação (Mestrado Integrado em Engenharia Electrotecnica e de Computadores) — Faculdade de Engenharia da Universidade de Porto, Portugal, 2008.

NASSEREDDINE, M.; KHANG, A. Applications of Internet of Things (IoT) in Smart Citie. In: KHANG, A.; ABDULLAYEV, V.; HAHANOV, V.; SHAH, V. (Ed.). **Advanced IoT Technologies and Applications in the Industry 4.0 Digital Economy**. London, England: CRC Press, 2024.

NEIVA, J. P. B. **Localização e Orientação Indoor com recurso à Tecnologia RFID**. 2012. 78 p. Dissertação (Mestrado Integrado em Engenharia Electrotecnica e de Computadores) — Faculdade de Engenharia da Universidade de Porto, Portugal, 2012.

NETO, J. P. **Uma Comparação Empirica do Desempenho Preditivo das Análises Fundamentalista, Gráfica e Econométrica no Mercado Acionário Brasileiro**. 2011. 593 p. Tese (Doutorado em Administração - Linha de pesquisa: Finanças) — Programa de

Pós-Graduação em Administração, Faculdade de Economia, Administração, Contabilidade e Ciência da Informação e Documentação, Universidade Federal de Brasília, Brasília, 2011.

NEVISTIC, Z.; BACIC, Z. The Concept, Realizations and Role of Geosciences in the Development of Smart Cities. **Technical gazette**, Croatia, v. 29, n. 1, p. 330–336, 2022.

NISHA, K. G.; SREEKUMAR, K. A Review and Analysis of machine Learning and Statistical Approaches for Prediction. In: INTERNATIONAL CONFERENCE ON INVENTIVE COMMUNICATION AND COMPUTATIONAL TECHNOLOGIES (ICICCT), 2017., 2017, Thondamuthur coimbatore, India. **Anais...** IEEE, 2017. p. 135–139.

OLSZEWSKI, D.; IWANOWSKI, M.; GRANISZEWSKI, W. Dimensionality reduction for detection of anomalies in the IoT traffic data. **Future Generation Computer Systems**, Amsterdam, North-Holland, v. 151, p. 137–151, 2024.

OUEIDA, S.; ALOQAILY, M.; IONESCU, S. A smart healthcare reward model for resource allocation in smart city. **Multimedia Tools and Applications**, Netherlands, v. 78, n. 17, p. 24573–24594, Sept. 2019.

PEREIRA, C. L. A. **Dispositivos de Identificação**. 2012. 127 p. Dissertação (Mestrado em Sistemas e Tecnologias de Informação para as Organizações). — Instituto Politecnido de Viseu, Escola Superior de Tecnologia e Gestão de Viseu, Portugal, 2012.

PORTO ALEGRE. **Atenção Primária à Saúde - Unidades de Saúde**. Porto Alegre: Prefeitura Municipal, 2024. Disponível em: <https://prefeitura.poa.br/carta-de-servicos/atencao-primaria-saude-unidades-de-saude/>. Acesso em: 23 mar. 2024.

PORTO ALEGRE. **Urgência e Emergência**. Porto Alegre: Prefeitura Municipal, 2024. Disponível em: <https://prefeitura.poa.br/sms/urgencia-e-emergencia/>. Acesso em: 23 mar. 2024.

PORTO ALEGRE. **Quadro Sintese - Servidores Municipais**. Porto Alegre: Prefeitura Municipal, 2024. Disponível em: <https://transparencia.portoalegre.rs.gov.br/sites/default/files/usudoc/uploads/smap/2023/09/quadro-sintese-agosto/2023/quadrosintese-202308.pdf/>. Acesso em: 23 mar. 2024.

PYTHON SOFTWARE FOUNDATION. **pmdarima 2.0.4**. 2024. Disponível em: <https://pypi.org/project/pmdarima/>. Acesso em: 04 ago. 2024.

QIN, Y.; SHENG, Q. Z.; CURRY, E. Dissemination of Internet of Things Streams in a Real-time Linked Dataspace. In: REAL-TIME LINKED DATASPACES: ENABLING DATA ECOSYSTEMS FOR INTELLIGENT SYSTEMS, 2020, Cham. **Anais...** Springer International Publishing, 2020. p. 191–208.

RAI, S.; GURUNG, A.; Bhakta Sharma, H.; Prakash Ranjan, V.; Ravi Sankar Cheela, V. Sustainable solid waste management challenges in hill cities of developing Countries: insights from eastern himalayan smart cities of sikkim, india. **Waste Management Bulletin**, USA, v. 2, n. 2, p. 1–18, 2024.

RASIDI, N. F.; YUSOFF, N. S. M.; SHUIB, A.; NOH, A. Y. M.; ALIAS, S.; AMINUDDIN, W. M. W. M. Estimation of New Resource Allocation in Hospital's (or Medical Care) Inpatient

Department using Discrete Event Simulation. **Applied Mathematics and Computational Intelligence (AMCI)**, [S.l.], v. 13, n. 2, p. 11–24, June 2024.

REJEB, A.; REJEB, K.; TREIBLMAIER, H.; APPOLLONI, A.; ALGHAMDI, S.; ALHASAWI, Y.; IRANMANESH, M. The Internet of Things (IoT) in healthcare: taking stock and moving forward. **Internet of Things**, Netherlands, v. 22, p. 100721, 2023.

RIGHI, R. D. R.; RODRIGUES, V. F.; NARDIN, I. F. D.; COSTA, C. A. D.; ALVES, M. A. Z.; PILLON, M. A. Towards providing middleware-level proactive resource reorganisation for elastic HPC applications in the cloud. **International Journal of Grid and Utility Computing**, Olney, England, v. 10, n. 1, p. 76–92, 2019.

RIGHI, R. R.; RODRIGUES, V. F.; COSTA, C. A. da; GALANTE, G.; BONA, L. C. E. de; FERRETO, T. AutoElastic: automatic resource elasticity for high performance applications in the cloud. **IEEE Transactions on Cloud Computing**, New York, NY, USA, v. 4, n. 1, p. 6–19, Jan. 2016.

RIO GRANDE DO SUL. **Covid-19 Dashboard**. 2024. Disponível em: <https://covid.saude.rs.gov.br/>. Acesso em: 15 ago. 2024.

RODRIGUES, V. F.; RIGHI, R. d. R.; CESCHINI, L. M.; BELLINI, B. C. L.; DONIDA, B.; COSTA, C. A. da. On revisiting vital signs IoT sensors for COVID-19 and long COVID-19 monitoring: a condensed updated review and future directions. **Journal of Ideas in Health**, Turkey, v. 4, n. 4, p. 604–614, Dec. 2021.

RODRIGUES, V. F.; ROSA RIGHI, R. da; COSTA, C. A. da; ZEISER, F. A.; ESKOFIER, B.; MAIER, A.; KIM, D. Digital health in smart cities: rethinking the remote health monitoring architecture on combining edge, fog, and cloud. **Health and Technology**, Germany, v. 13, n. 3, p. 449–472, Jun 2023.

RODRIGUES, V.; RIGHI, R. MinhaHistoriaDigital: An Scalable Fog-Based Architecture for Efficient Vital Signs Monitoring over Smart Cities. In: XXII ESCOLA REGIONAL DE ALTO DESEMPENHO DA REGIÃO SUL, 2022, Curitiba. **Anais...** SBC, 2022. p. 117–118.

ROEHRS, A.; COSTA, C. A. da; ROSA RIGHI, R. da; OLIVEIRA, K. S. F. de. Personal Health Records: A Systematic Literature Review. **Journal of medical Internet research**, Bethesda, MD, USA, v. 19, n. 1, 2017.

ROSA RIGHI, R.; COSTA, C. A.; RODRIGUES, V. F.; ROSTIROLLA, G. Joint-analysis of Performance and Energy Consumption when Enabling Cloud Elasticity for Synchronous HPC Applications. **Concurrency and Computation : Practice and Experience**, Chichester, UK, v. 28, n. 5, p. 1548–1571, Apr. 2016.

ROSA RIGHI, R. da; BELLINI, B. C. L.; FRITSCH, F.; RODRIGUES, V. F.; SINGH, M.; PASIN, M. On Defining and Deploying Health Services in Fog-Cloud Architectures. In: ROSA RIGHI, R. da (Ed.). **Ubiquitous and Pervasive Computing**. Rijeka: IntechOpen, 2023.

ROSSI, D. J. **Previsão da Velocidade dos Ventos por Redes Neurais Artificiais e ARIMA de Box & Jenkins**. 2013. 76 p. Dissertação (Mestrado em Engenharia Elétrica - Área de conhecimento: Automação) — Programa de Pós-Graduação em Engenharia Elétrica, Faculdade de Engenharia de Ilha Solteira, Universidade Estadual Paulista, Ilha Solteira, 2013.

ROSTIROLLA, G.; R. RIGHI, R. d.; BARBOSA, J. L. V.; COSTA, C. A. da. ElCity: an elastic multilevel energy saving model for smart cities. **IEEE Transactions on Sustainable Computing**, Piscataway, NJ, USA, v. 3, n. 1, p. 30–43, Jan. 2018.

ROSTIROLLA, G.; ROSA RIGHI, R. da; BARBOSA, J. L. V.; COSTA, C. A. da. ElCity: an elastic multilevel energy saving model for smart cities. **IEEE Transactions on Sustainable Computing**, Piscataway, NJ, USA, v. PP, n. 99, p. 1–1, 2017.

SEDGEWICK, R.; WAYNE, K. **Class StdRandom**. July 2022. Disponível em: <https://introcs.cs.princeton.edu/java/stdlib/javadoc/StdRandom.html>. Acesso em: 25 July 2022.

SELVI, M. S.; DEEPA, K.; BALAMURUGAN, S.; RANI, S. J.; UVAZEAHAMED, A. M. Chapter 7 RTLS: an introduction. IN: GOYAL, D.; BALAMURUGAN, S.; KARTHIKRAJAN SENTHILNATHAN, I. A.; MOHAMMAD (Ed.). **Cyber-Physical Systems and Industry 4.0**. 1st. ed. Cham: Apple Academic Press, 2022.

SEN, S.; YADEO, D.; KUMAR, P.; KUMAR, M. Chapter Thirteen - Machine learning and predictive control-based energy management system for smart buildings. In: BASETTI, V.; SHIVA, C. K.; UNGARALA, M. R.; RANGARAJAN, S. S. (Ed.). **Artificial Intelligence and Machine Learning in Smart City Planning**. Amsterdam, Netherlands: Elsevier, 2023. p. 199–220.

SIDOROV, D. Preface to Model Predictive Control and Optimization for Cyber-Physical Systems. **Mathematics**, Basel, Switzerland, v. 11, n. 4, 2023.

SILBERSCHATZ, A.; GALVIN, P. B.; GAGNE, G. **Operating System Concepts**. 9. ed. Hoboken, NJ, USA: Wiley Publishing, 2012.

SINGH, J.; WAZID, M.; DAS, A. K.; CHAMOLA, V.; GUIZANI, M. Machine learning security attacks and defense approaches for emerging cyber physical applications: a comprehensive survey. **Computer Communications**, United Kingdom, v. 192, p. 316–331, 2022.

SURESH, S.; CHAKARAVARTHI, G. RFID technology and its diverse applications: a brief exposition with a proposed machine learning approach. **Measurement**, Netherlands, v. 195, p. 111197, 2022.

SYED, A. S.; SIERRA-SOSA, D.; KUMAR, A.; ELMAGHRABY, A. IoT in Smart Cities: a survey of technologies, practices and challenges. **Smart Cities**, Basel, Switzerland, v. 4, n. 2, p. 429–475, 2021.

TOOMAN, H.; PIIRMAN, M.; KÄRP, M. Using the Principles of the Holistic Wellness Concept in Designing and Developing Hospital Services: case of pärnu hospital, rehabilitation and well-being centre. In: PFANNSTIEL, M. A.; BREHMER, N.; RASCHE, C. (Ed.). **Service Design Practices for Healthcare Innovation**: paradigms, principles, prospects. Cham: Springer International Publishing, 2022. p. 201–223.

TRANSCENDS. **Rifidi Software Defined RFID**. 2011. Disponível em: <http://www.transcends.co/www/>. Acesso em: 20 jun. 2018.

TRIPATHY, S. S.; RATH, M.; TRIPATHY, N.; ROY, D. S.; FRANCIS, J. S. A.; BEBORTTA, S. An Intelligent Health Care System in Fog Platform with Optimized Performance. **Sustainability**, Basel, Switzerland, v. 15, n. 3, 2023.

WANG, H.; SINNEN, O. List-Scheduling versus Cluster-Scheduling. **IEEE Transactions on Parallel and Distributed Systems**, USA, v. 29, n. 8, p. 1736–1749, Aug. 2018.

WANG, M.; ZHOU, T. Does smart city implementation improve the subjective quality of life? Evidence from China. **Technology in Society**, Netherlands, v. 72, p. 102161, 2023.

WANT, R. An introduction to RFID technology. **IEEE Pervasive Computing**, New York, NY, USA, v. 5, n. 1, p. 25–33, Jan. 2006.

XIANG, W.; CHEN, L.; YAN, X.; WANG, B.; LIU, X. The impact of traffic control measures on the spread of COVID-19 within urban agglomerations based on a modified epidemic model. **Cities**, United Kingdom, v. 135, p. 104238, 2023.

XIUFAN, Z.; DECHENG, F. Collaborative emission reduction research on dual-pilot policies of the low-carbon city and smart city from the perspective of multiple innovations. **Urban Climate**, Netherlands, v. 47, p. 101364, 2023.

YU, Z.; FU, M. Cloud API Oriented Analysis on Rolling Upgrade Dependability. In: IEEE 2ND INTERNATIONAL CONFERENCE ON ELECTRICAL ENGINEERING, BIG DATA AND ALGORITHMS (EEBDA), 2023., 2023, Changchun, China. **Anais...** IEEE, 2023. p. 1560–1564.

ZANCHETTIN, L. **Emergências e Unidades de Pronto Atendimento de Porto Alegre operam com superlotação**. **Zero Hora**, Porto Alegre, 10 jul. 2024. Disponível em: <https://gauchazh.clicrbs.com.br/saude/noticia/2024/07/emergencias-e-unidades-de-pronto-atendimento-de-porto-alegre-operam-com-superlotacao-clyghap5300v6015e59vucrps.html>. Acesso em: 12 ago. 2024.

ZHANG, Q. S. Environment pollution analysis on smart cities using wireless sensor networks. **Strategic Planning for Energy and the Environment**, Denmark, v. 42, n. 1, p. 239–262, 2023.

ZIOSI, M.; HEWITT, B.; JUNEJA, P.; TADDEO, M.; FLORIDI, L. Smart Cities: mapping their ethical implications. **SSRN Electronic Journal**, USA, Jan. 2022.