

UNIVERSIDADE DO VALE DO RIO DOS SINOS — UNISINOS
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA
NÍVEL MESTRADO

MATEUS ROVEDA

VISPAC: COMPRESSÃO ADAPTATIVA E PRIORIZAÇÃO INTELIGENTE NO
MONITORAMENTO DE SINAIS VITAIS DE PACIENTES EM ARQUITETURAS
EDGE-FOG-CLOUD

SÃO LEOPOLDO
2025

Mateus Roveda

VISPAC: COMPRESSÃO ADAPTATIVA E PRIORIZAÇÃO INTELIGENTE NO
MONITORAMENTO DE SINAIS VITAIS DE PACIENTES EM ARQUITETURAS
EDGE-FOG-CLOUD

Dissertação apresentada como requisito parcial
para a obtenção do título de Mestre 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
2025

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)

Roveda, Mateus

VISPAC: COMPRESSÃO ADAPTATIVA E PRIORIZAÇÃO INTELIGENTE NO MONITORAMENTO DE SINAIS VITAIS DE PACIENTES EM ARQUITETURAS Edge-Fog-Cloud / Mateus Roveda — 2025.

?? f.: il.; 30 cm.

Dissertação (mestrado) — Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Computação Aplicada, São Leopoldo, 2025.

“Orientador: Prof. Dr. Rodrigo da Rosa Righi, Unidade Acadêmica de Pesquisa e Pós-Graduação”.

1. Internet das Coisas. 2. Sinais vitais. 3. Compressão de dados adaptativa. 4. Compressão de dados híbrida. 5. Cidades inteligentes. I. Título.

CDU 004.732

Bibliotecária responsável: Fulana da Silva — CRB 12/3456

(Esta folha serve somente para guardar o lugar da verdadeira folha de aprovação, que é obtida após a defesa do trabalho. Este item é obrigatório, exceto no caso de TCCs.)

AGRADECIMENTOS

Aos meus pais (Leonel e Eliane). Mesmo sem acesso ao conhecimento quando jovens, sempre reforçaram, ao longo da minha vida, que estudar, aprender e compartilhar são fundamentais. A vocês, o meu mais sincero obrigado. Vocês fizeram escolhas e renúncias que me permitiram estudar e alcançar meus objetivos.

À minha esposa Kátia, que sempre ofereceu apoio emocional e motivacional para que essa etapa fosse cumprida. Seu apoio permitiu-me seguir em frente diante dos desafios e do cansaço. Você foi muito importante nessa jornada. Sem você, dificilmente chegaria ao final tão empolgado.

Ao meu orientador, Rodrigo da Rosa Righi, que, ao longo da jornada, sempre instigou a ir além e direcionou os caminhos a serem trilhados e os objetivos a serem alcançados. Estendo o agradecimento a todos os professores que fizeram parte desse caminho, deixando um pouco de si em prol do conhecimento e do avanço científico, além dos avaliadores, que apontaram itens importantes para ajustes e visões, para que o trabalho ficasse mais coeso e atingisse os objetivos estipulados. Também estendo os agradecimentos a todos os alunos/professores com quem, indiretamente, acabei tendo contato ao longo do processo.

Por fim, a Universidade do Vale do Rio dos Sinos, que abriu suas portas para essa pesquisa e me permitiu ter uma experiência enriquecedora no processo, alcançando meu objetivo e contribuindo para a iniciação científica.

RESUMO

A Internet das Coisas (IoT) tem impactado de forma expressiva diversas áreas, especialmente a medicina, ao viabilizar o monitoramento contínuo e remoto de pacientes por meio de arquiteturas *Edge-Fog-Cloud*, que combinam conectividade e processamento em tempo hábil. O desafio central está no grande volume de dados gerado por múltiplos sensores, que pode sobrecarregar a infraestrutura de rede, aliado à necessidade de garantir que alterações clínicas críticas sejam comunicadas com mínima latência, permitindo intervenções rápidas. Neste cenário, o tema desta dissertação é a otimização do monitoramento de sinais vitais em ambientes de saúde conectados, com foco na utilização eficiente de recursos computacionais e de rede. Embora existam avanços relevantes na área, a maioria dos trabalhos relacionados trata compressão e priorização de forma dissociada, sem integrar uma compressão multissinal adaptativa a uma priorização clínica baseada em escores padronizados. Além disso, falta um ciclo fechado onde a inteligência das camadas *fog-cloud* reconfigure dinamicamente a coleta na camada *edge*. Esta dissertação propõe o modelo *Vital Sign Prioritization and Adaptive Compression* (ViSPAC), que unifica compressão e priorização adaptativa com base no escore clínico National Early Warning Score 2 (NEWS2). O escore é utilizado para ajustar, em tempo hábil, os parâmetros de compressão e a frequência de coleta, otimizando recursos sem comprometer a fidelidade dos sinais. O modelo incorpora uma estratégia de controle que equilibra eficiência e segurança clínica por meio de três algoritmos interdependentes: configuração dinâmica por sinal vital, *back-off* exponencial condicionado ao risco e vigilância contínua (*keep-alive*). Os experimentos foram conduzidos em ambiente distribuído na AWS com 20 nós de borda operando sob restrições de hardware similares a dispositivos IoT. Os resultados demonstraram que o ViSPAC reduziu em 96,7% o número de transmissões comparado à linha de base, alcançou taxa de compressão média de 81,6% (superior aos 75,0% do cenário estático) e obteve distorção PRD média global de apenas 1,16%, substancialmente inferior aos 3,62% do modelo estático. A adoção de serialização binária *MessagePack* reduziu o volume bruto dos *payloads* em aproximadamente 30% em relação ao JSON textual. A latência média do ciclo de *feedback* foi de 1.048,9 ms, adequada para monitoramento remoto não invasivo. Em síntese, o ViSPAC apresenta-se como uma solução equilibrada para o monitoramento de saúde em larga escala no contexto de Cidades Inteligentes, otimizando recursos de rede e energia sem comprometer a capacidade de resposta a eventos clínicos adversos.

Palavras-chave: Internet das Coisas. Sinais vitais. Compressão de dados adaptativa. Compressão de dados híbrida. Cidades inteligentes.

ABSTRACT

The Internet of Things (IoT) has substantially impacted various fields, especially medicine, by enabling continuous and remote patient monitoring through Edge-Fog-Cloud architectures, which combine connectivity with real-time processing. The central challenge lies in the large volume of data generated by multiple sensors, which can overload the network infrastructure, coupled with the need to ensure that critical clinical changes are communicated with minimal latency to allow timely interventions. In this context, the central theme of this dissertation is the optimization of vital sign monitoring in connected healthcare environments, focusing on the efficient use of computational and network resources. Although notable advances have been made in the field, most related works address compression and prioritization separately, without integrating adaptive multi-signal compression with clinical prioritization based on standardized scores. Moreover, there is a lack of a closed-loop system in which intelligence in the fog-cloud layers dynamically reconfigures data collection at the edge. This dissertation proposes the Vital Sign Prioritization and Adaptive Compression (ViSPAC) model, which unifies adaptive compression and prioritization based on the clinical score National Early Warning Score 2 (NEWS2). The score is used to adjust compression parameters and sampling frequency in real time, optimizing resources without compromising signal fidelity. The model incorporates a control strategy that balances efficiency and clinical safety through three interdependent algorithms: dynamic configuration per vital sign, risk-conditioned exponential back-off, and continuous vigilance (keep-alive). Experiments were conducted in a distributed AWS environment with 20 edge nodes operating under hardware constraints similar to IoT devices. The results demonstrated that ViSPAC reduced the number of transmissions by 96.7% compared to the baseline, achieved an average compression rate of 81.6% (higher than the 75.0% of the static scenario), and obtained a global average PRD distortion of only 1.16%, substantially lower than the 3.62% of the static model. The adoption of binary MessagePack serialization reduced raw payload volume by approximately 30% compared to textual JSON. The average feedback loop latency was 1,048.9 ms, suitable for non-invasive remote monitoring. In summary, ViSPAC presents itself as a balanced solution for large-scale health monitoring in the context of Smart Cities, optimizing network and energy resources without compromising the ability to respond to adverse clinical events.

Keywords: Internet of Things. Vital signs. Adaptive data compression. Hybrid data compression. Smart cities.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
AWS	Amazon Web Services
BVP	Blood Volume Pulse
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CD	Compression Deviation
CPU	Central Processing Unit
ECG	Eletrocardiograma
FAPERGS	Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul
FC	Frequência Cardíaca
FR	Frequência Respiratória
HTTP	HyperText Transfer Protocol
IaC	Infrastructure as Code
IoT	Internet of Things
JSON	JavaScript Object Notation
MessagePack	Formato de serialização binária compacto
MQTT	Message Queuing Telemetry Transport
MSE	Mean Squared Error
NEWS2	National Early Warning Score 2
PRD	Percentage Root Mean Difference
PSNR	Peak Signal-to-Noise Ratio
QoS	Qualidade de Serviço
RMSE	Root Mean Square Error
SaaS	Software as a Service
SAMU	Serviço de Atendimento Móvel de Urgência
SCADA	Supervisory Control and Data Acquisition
SDT	Swinging Door Trending
SoC	System on a Chip
SpO2	Saturação Periférica de Oxigênio
SSIM	Structural Similarity Index
TCP	Transmission Control Protocol
TLS	Transport Layer Security
ViSPAC	Vital Sign Prioritization and Adaptive Compression
VSAC	Vital Sign Adaptive Compressor

WBAN Wireless Body Area Network

SUMÁRIO

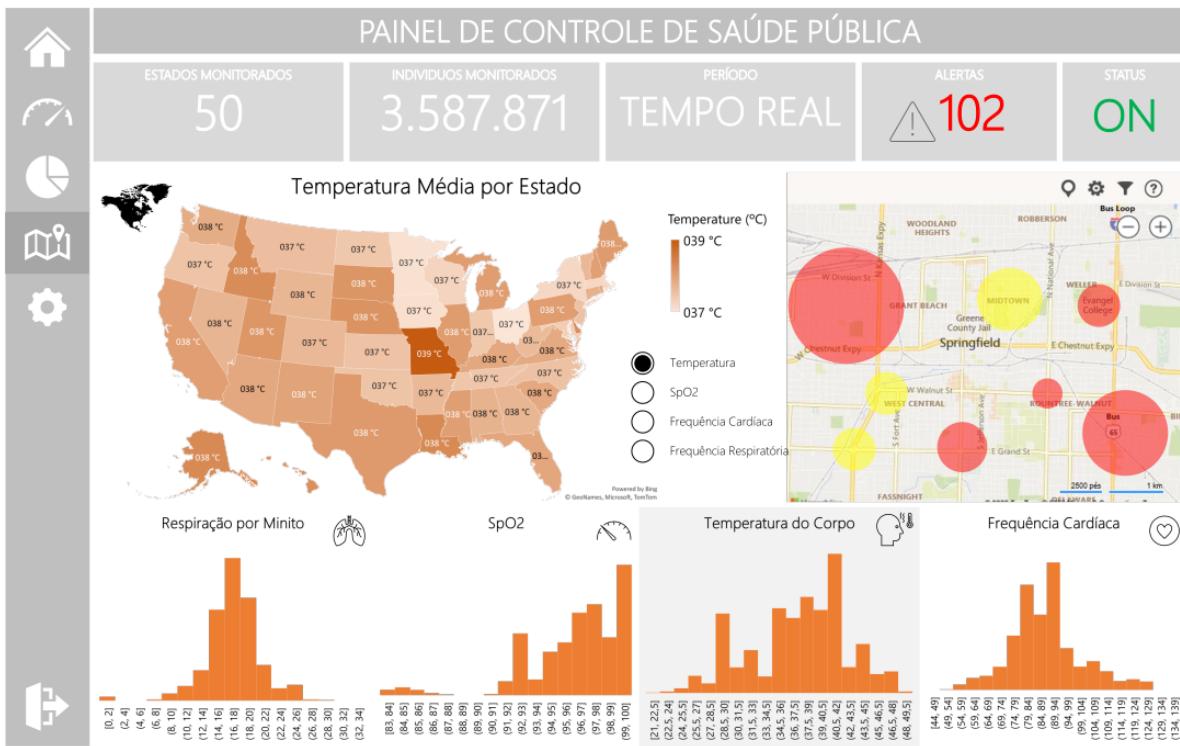
1 INTRODUÇÃO

Nos últimos anos, a Internet das Coisas (IoT) tem ganhado destaque na área da saúde, apresentando crescimento significativo impulsionado pelo desenvolvimento de dispositivos especializados para aplicações médicas. Uma revisão sistemática recente identificou avanços substanciais em tecnologias de sensores vestíveis para monitoramento de saúde, abrangendo aplicações clínicas, reabilitação e avaliação de risco de doenças (?). Em particular, o monitoramento remoto de pacientes por meio da IoT vem se mostrando de grande importância, pois permite acompanhar continuamente sinais vitais fora do ambiente hospitalar, aliviando a sobrecarga dos hospitais e oferecendo melhor suporte a pacientes com doenças crônicas (?). Sinais vitais como frequência cardíaca (FC), temperatura corporal, saturação de oxigênio periférico (SpO_2) e frequência respiratória são indicadores essenciais do estado de saúde; seu acompanhamento em tempo hábil possibilita a detecção precoce de alterações clínicas relevantes e uma intervenção médica mais rápida. Pacientes idosos ou com enfermidades crônicas frequentemente necessitam desse monitoramento contínuo de parâmetros fisiológicos, pois a falta de vigilância pode acarretar riscos graves à sua saúde. Nesse contexto, a combinação de IoT e sistemas de saúde conectados tem impulsionado a telemedicina e a criação de soluções de *e-health* inteligentes, integrando sensores biomédicos que coletam esses sinais e os encaminham para avaliação remota em tempo hábil (?). A Figura ?? demonstra um painel de exemplo para monitoramento de saúde pública em tempo hábil, onde é possível analisar sinais vitais, alertas e o status do sistema a nível macro.

Entretanto, o monitoramento unificado traz um grande desafio: gerenciar a quantidade gigantesca de dados gerados. Estudos recentes demonstram que sistemas de monitoramento contínuo de sinais vitais podem ser escalados para operar em redes hospitalares com múltiplas unidades, evidenciando tanto a viabilidade quanto os desafios de infraestrutura envolvidos (?). Relatórios de mercado mostram que a infraestrutura conectada continua crescendo rápido. Como mostra a Figura ??, o mundo chegou a 21,1 bilhões de dispositivos IoT ativos em 2025, e a previsão é que esse número suba para **23,9 bilhões já em 2026** (?). Isso faz com que o volume de dados gerados no mundo todo dispare. Estimativas apontam que vamos atingir a marca de 221 Zettabytes (ZB) de dados armazenados até 2026 (??). Nesse cenário, o problema deixa de ser apenas limitações de rede e passa a ser o alto custo financeiro e energético para transmitir e guardar tudo isso na nuvem (?).

Quando examinamos o tráfego gerado por sensores de saúde, percebemos que a maioria dos dados transmitidos não é informação clínica de fato, mas sim informações extras exigidas pelos protocolos de comunicação (*overhead*). Para entender melhor esse problema, vamos analisar uma transmissão padrão via HTTP/1.1. Os dados úteis contêm as informações vitais básicas: FC (1 byte), SpO_2 (1 byte), identificador (4 bytes) e horário (8 bytes), totalizando apenas 14 bytes. Porém, para que diferentes sistemas consigam se comunicar, esses dados precisam ser formatados em estruturas como JSON, que adiciona vários caracteres extras e aumenta o tama-

Figura 1: Painel de controle de exemplo para monitoramento da saúde pública



Fonte: Adaptado do projeto Minha História Digital, de autoria de Rodrigo da Rosa Righi e Alexandre Luis de Andrade

nho para cerca de 49 bytes:

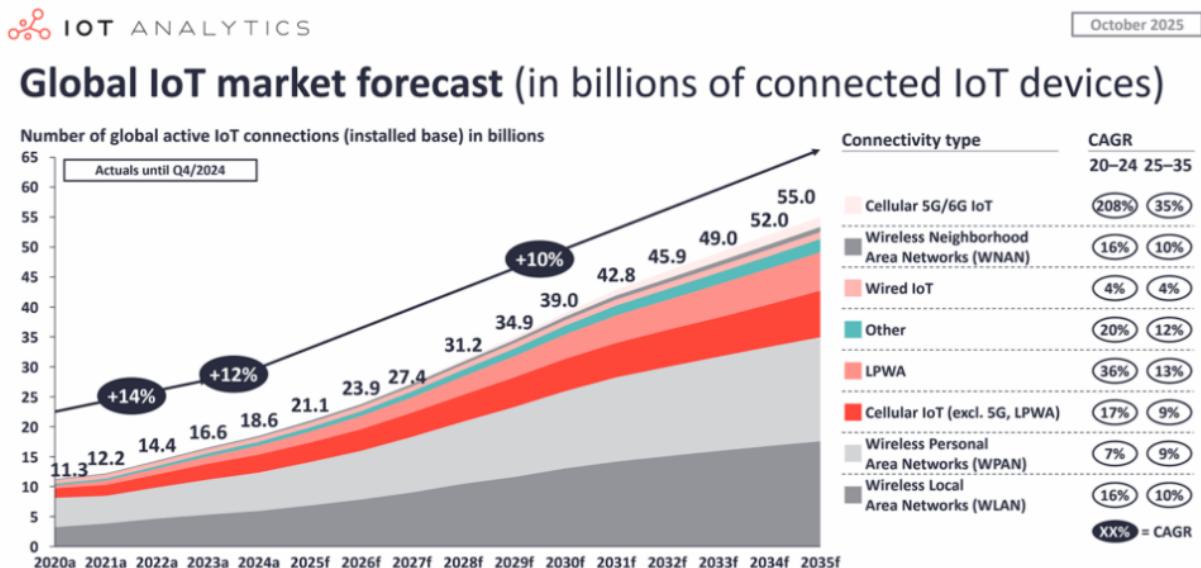
```
{"pid":1234,"ts":1726713420000,"fc":78,"spo2":98}
```

O problema fica ainda maior quando esses dados são preparados para o transporte pela rede. Cada transmissão precisa incluir os cabeçalhos do HTTP/1.1 (cerca de 124 bytes) (?), além dos cabeçalhos das camadas de infraestrutura: IPv4 (20 bytes) (?) e TCP (20 bytes) (?). Somando tudo, o custo final para transmitir aqueles 14 bytes de informação real chega a aproximadamente 213 bytes, criando uma proporção bem desequilibrada entre o conteúdo útil e toda a estrutura necessária para enviá-lo.

Esse desequilíbrio gera consequências quando consideramos o cenário de larga escala mencionado antes. Transmitir pacotes onde a maior parte é apenas estrutura de protocolo esgota rapidamente a bateria dos dispositivos portáteis e aumenta os custos operacionais na nuvem, já que os serviços de armazenamento e transferência de dados cobram pelo volume total movimentado (?). Dessa forma, a ineficiência no transporte de pequenos volumes de dados impõe barreiras técnicas e financeiras que podem inviabilizar a sustentabilidade do monitoramento contínuo.

Nesse contexto, torna-se crucial empregar técnicas eficientes de compressão de dados para reduzir o tráfego sem perder informações clínicas relevantes. A compressão adequada dos sinais vitais permite minimizar o volume transmitido da borda até a nuvem, evitando sobrecarga

Figura 2: Previsão de mercado global de IoT e crescimento da base instalada (2020-2035)



Fonte: Adaptado de (?)

da rede e atendendo aos requisitos de tempo de resposta dos sistemas de telemonitoramento (?). Diversas abordagens vêm sendo exploradas na literatura para dados biomédicos, incluindo métodos com perdas controladas (lossy) e sem perdas (lossless), com o objetivo de otimizar a utilização de banda e armazenamento sem comprometer a integridade dos sinais necessários ao diagnóstico.

Além da compressão, outro aspecto fundamental em sistemas de IoT para a medicina é a priorização dos dados coletados para suportar decisões clínicas em tempo hábil. Nem todos os sinais transmitidos possuem a mesma criticidade: leituras dentro da normalidade toleram certa latência, enquanto alterações maiores exigem entrega imediata para permitir intervenções rápidas. Abordagens recentes propõem classificar pacotes em níveis de prioridade de acordo com limiares fisiológicos, demonstrando redução da latência de alerta quando filas de prioridade são empregadas (?). Essa diferenciação visa garantir qualidade de serviço (QoS) adequada, de modo que informações críticas (e.g., detecção de arritmia cardíaca ou queda significativa de saturação de oxigênio) sejam comunicadas com o mínimo de atraso e confiabilidade máxima, pois atrasos na entrega de dados vitais anômalos podem ser potencialmente fatais. Estratégias de roteamento e gerenciamento de fila sensíveis à prioridade têm se mostrado eficazes para reduzir a latência de dados emergenciais em redes de sensores corporais e ambientes hospitalares conectados (?). Em um cenário prático, essa capacidade de priorização pode beneficiar diretamente serviços de emergência como o SAMU (Serviço de Atendimento Móvel de Urgência): ao receber, em tempo hábil, o escore de risco clínico de múltiplos pacientes monitorados em uma região, a central de regulação médica poderia otimizar o despacho de ambulâncias, priorizando os casos mais críticos e permitindo que médicos e socorristas tomem decisões de triagem fundamentadas em dados objetivos, especialmente em cenários de múltiplas vítimas ou recursos

limitados.

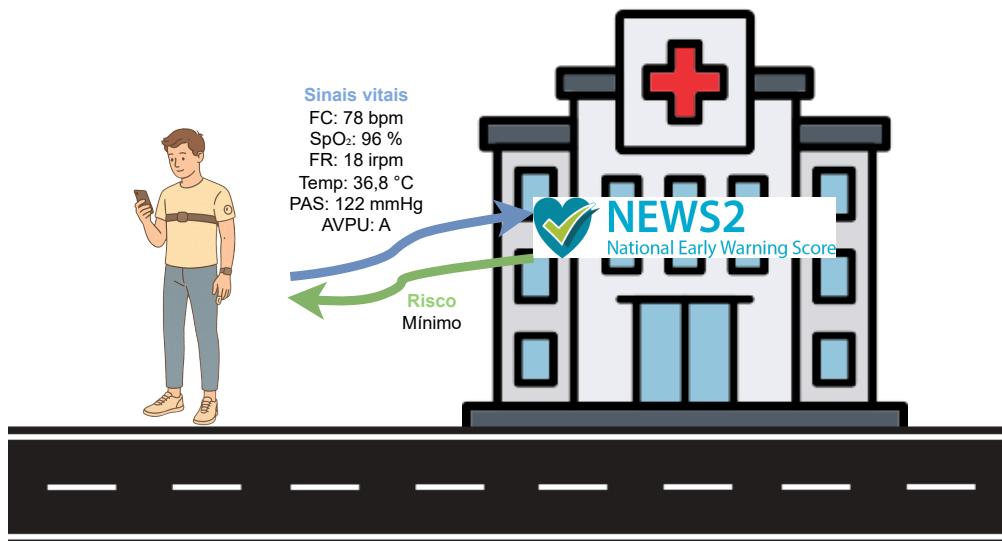
Portanto, a utilização da arquitetura *edge-fog-cloud* se torna fundamental. Considerada o estado da arte para sistemas distribuídos, sua estrutura remete à organização político-geográfica de uma cidade: a borda (*edge*) corresponde à coleta local de dados (residências, empresas); a névoa (*fog*) atua como um centro de processamento regional (postos de saúde, centros de dados municipais); e a nuvem (*cloud*) representa a infraestrutura centralizada (secretaria de saúde estadual ou nacional ou hospitais). Nesse ecossistema, mecanismos de compressão e priorização de dados como os discutidos aqui funcionam como um *middleware* essencial, que viabiliza a tomada de decisão clínica em tempo hábil e insere o monitoramento de saúde como um serviço chave em aplicações de Cidades Inteligentes.

Nesse cenário, a literatura propõe modelos arquiteturais que aliem compressão de dados e entrega seletiva de informações conforme sua relevância clínica. Dentre eles, destaca-se o modelo VSAC (*Vital Sign Adaptive Compressor*), concebido para transmissão eficiente de sinais vitais em aplicações IoT para a medicina (?). O VSAC combina algoritmos de compressão com perdas e sem perdas em uma arquitetura hierárquica de múltiplas camadas (*edge–fog–cloud*), reduzindo drasticamente o volume de dados transmitido sem comprometer a qualidade da informação. Em sua operação, inicialmente aplica-se compressão *lossy* adaptativa para filtrar e reter apenas as informações mais relevantes de cada sinal vital, seguida de compressão *loss-less* para compactar os dados antes do envio às camadas superiores. Essa abordagem híbrida demonstrou desempenho superior aos métodos convencionais de redução de dados, alcançando ganhos de compressão de até 46% na taxa de redução de bytes transmitidos, quando comparada a técnicas isoladas (?). Em outras palavras, o VSAC provê um mecanismo otimizado de coleta e agregação de sinais heterogêneos distribuídos geograficamente, com processamento em tempo hábil na borda, atendendo às demandas de aplicações de tele-saúde em cenários de cidades inteligentes.

Na presente dissertação, propõe-se uma continuidade do VSAC, com testes com múltiplos sinais vitais (incluindo, além da FC já considerada, parâmetros como o SpO_2) e a incorporação de um mecanismo de priorização inteligente dos dados baseado em escores clínicos de gravidade, utilizando o NEWS2, denominado *Vital Signs Prioritized Adaptive Compression* (ViSPAC). Em termos práticos, o sistema passa a adaptar dinamicamente os níveis de compressão e a ordem de transmissão dos sinais de acordo com indicadores clínicos de urgência, potencializados por *feedbacks* das camadas *fog* ou *cloud*. Essa estratégia permite que informações críticas recebam tratamento preferencial (menor compressão ou envio imediato), assegurando que eventos potencialmente perigosos sejam comunicados sem atrasos significativos.

Assim, o modelo proposto não apenas mantém os benefícios de redução de volume de dados e arquitetura do VSAC, como também agrega uma camada adicional na gestão dos dados médicos, priorizando-os conforme o estado do paciente. Espera-se, com isso, otimizar a transmissão dos sinais vitais mais relevantes para a decisão clínica em tempo hábil, melhorando a capacidade de resposta e a eficácia de sistemas de monitoramento contínuo de pacientes em ambientes

Figura 3: Coleta de sinais vitais através de sensores, compressão adaptativa e envio para um centro de processamento (hospital) para cálculo do risco clínico. Essa coleta acontece como um ciclo, coletando diversas vezes os sinais vitais dos pacientes e transmitindo, podendo inundar a rede com pacotes caso não haja compressão e/ou tratamentos



Fonte: Elaborado pelo autor

IoT de saúde. Em suma, o trabalho aqui apresentado busca contribuir para o desenvolvimento de soluções de telemonitoramento mais robustas, adaptativas e focadas na relevância clínica dos dados, unindo compressão adaptativa e priorização para suportar a tomada de decisão médica imediata em contextos de IoT aplicada à saúde.

A Figura ?? traz, de maneira resumida e simples, uma exemplificação do trabalho aqui proposto, onde a coleta de sinais vitais é feita através de dispositivos vestíveis, comprimida e transmitida para um hospital ou ponto de saúde próximo.

1.1 Questões de pesquisa

Este trabalho busca responder à seguinte questão de pesquisa principal:

Como aplicar priorização dinâmica com base em escores de múltiplos sinais vitais, de modo a otimizar a compressão, a coleta e a transmissão de dados clínicos críticos em um ambiente Edge–Fog–Cloud sem prejudicar a qualidade das informações?

Para fins deste trabalho, entende-se por priorização dinâmica o ajuste contínuo e automático do grau de relevância dos dados, orientado por critérios clínicos avaliados em tempo hábil. O termo escores refere-se a indicadores numéricos de gravidade, sendo adotado neste estudo o NEWS2. Os sinais vitais, por sua vez, são os principais parâmetros fisiológicos cuja monitorização sustenta a avaliação do estado do paciente. O verbo otimizar refere-se à maximizar a

eficiência do uso de recursos de comunicação, armazenamento e processamento, preservando a utilidade clínica da informação transmitida. Por fim, a expressão Edge–Fog–Cloud descreve uma arquitetura computacional em camadas na qual o processamento se distribui desde dispositivos junto ao paciente (Edge), passando por nós intermediários de proximidade (Fog), até centros de dados centralizados (Cloud).

Com o foco em responder a questão principal, foram definidos objetivos que abordam trade-offs de compressão versus fidelidade e o impacto das notificações de prioridade no desempenho do sistema. Essas questões subsidiarão o avanço dos experimentos, a escolha das métricas de avaliação e a validação da solução em cenários de monitoramento remoto de pacientes.

1. Como testar mais de um sinal vital a estrutura já existente?
2. Como receber notificações das camadas *Fog/Cloud* e adaptar a compressão na camada de *Edge*?
3. Qual o impacto das camadas de compressão na qualidade dos dados?
4. Qual o impacto, na camada de borda, do recebimento de notificações de prioridade (callbacks) geradas pela camada *Fog/Cloud* sobre o uso de recursos computacionais e a gestão da coleta de sinais vitais?

1.2 Objetivos

Este estudo apresenta o ViSPAC, um modelo concebido para consolidar, comprimir e transmitir múltiplos sinais vitais em arquiteturas *Edge–Fog–Cloud* enquanto ajusta, em tempo hábil, o grau de compressão e a frequência de coleta segundo a criticidade clínica de cada paciente. O ViSPAC integra o escore NEWS2 nas camadas *fog-cloud* para classificar dinamicamente o risco e emitir notificações que parametrizam a camada edge, permitindo que dados de pacientes críticos sejam amostrados com maior resolução e compressão mais branda, ao passo que leituras estáveis recebam tratamento econômico.

Diante disso, o objetivo principal desta pesquisa é:

O objetivo geral deste trabalho é validar um mecanismo dinâmico de priorização, fundamentado no NEWS2, que reduza substancialmente o volume de dados transmitidos e o consumo de recursos, mantendo a utilidade clínica dos sinais vitais em um ambiente Edge–Fog–Cloud.

A fim de alcançar o objetivo principal desta pesquisa, os seguintes objetivos específicos foram definidos:

1. **Agilidade na decisão clínica:** permitir que o sistema envie alertas rapidamente sempre que o escore NEWS2 indicar risco, assegurando tempo hábil para que a equipe de saúde intervenha;

2. **Redução de Sobrecarga de Rede:** empregar compressão adaptativa para priorizar a eficiência na transmissão de sinais vitais, resultando em menor ocupação da largura de banda e redução da latência total de comunicação;
3. **Fidelidade e integridade dos dados:** preservar a qualidade dos sinais comprimidos de modo que possam ser reconstruídos sem perda de informação clínica relevante, garantindo confiança no diagnóstico;

1.3 Estrutura do texto

O restante desta dissertação está organizado da seguinte forma: o **Capítulo 2** apresenta a *Fundamentação Teórica*, descrevendo os princípios de monitoramento remoto de sinais vitais, a explicação e o entendimento do modelo VSAC e sua compressão de dados e os de priorização de dados que embasam o ViSPAC. Em seguida, o **Capítulo 3** discute os *Trabalhos Relacionados*, oferecendo uma visão crítica da literatura recente e identificando lacunas que motivam a extensão do modelo VSAC, trazendo o estado da arte atual. O **Capítulo 4** é dedicado ao *Modelo ViSPAC*, detalhando a arquitetura, os algoritmos de compressão *lossy/lossless* empregados e o mecanismo dinâmico de priorização baseado no escore NEWS2. O **Capítulo 5** descreve a *Metodologia e Experimento*, apresentando o design experimental, os materiais e dados utilizados, o ambiente de avaliação na AWS e as métricas adotadas. O **Capítulo 6** apresenta os *Resultados e Discussão*, com a análise comparativa entre os cenários testados e as respostas às questões de pesquisa. Por fim, o **Capítulo 7** encerra com a *Conclusão*, sintetizando as contribuições e indicando direções para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

O objetivo deste capítulo é apresentar os fundamentos teóricos que embasam o desenvolvimento deste trabalho, sendo organizado em quatro seções principais. Inicialmente, na Seção ??, serão discutidos os aspectos relacionados ao monitoramento remoto de sinais vitais, enfatizando os processos de aquisição dos dados, a correlação com a IoT e as principais características dos diferentes tipos de sinais que serão considerados. Em seguida, a Seção ?? dedica-se à análise da arquitetura e do funcionamento do VSAC, modelo base que será utilizado nesse trabalho. A Seção ?? tratará de priorização, explorando abordagens para determinar a importância relativa dos dados, visando otimizar a utilização de recursos e a tomada de decisão clínica. Por fim, a Seção ?? contextualizará sobre o projeto Minha História Digital, ao qual esse projeto faz parte.

2.1 Monitoramento remoto de sinais vitais

A evolução da IoT tem impulsionado consideravelmente o desenvolvimento de sistemas de monitoramento de saúde remota, possibilitando a aquisição contínua de sinais vitais sem comprometer a rotina dos pacientes. Essa tecnologia tem promovido uma transição do modelo tradicional, centrado em diagnósticos reativos, para uma abordagem preditiva e proativa no cuidado com a saúde (?). Dispositivos eletrônicos compactos e vestíveis, conectados ao corpo, são capazes de monitorar parâmetros fisiológicos essenciais, como temperatura corporal, FC, pressão arterial, frequência respiratória, saturação de oxigênio e pulso (?). A base desses sistemas costuma ser composta por arquiteturas do tipo Wireless Body Area Network (WBAN), que viabilizam a captação e o envio de dados biométricos por meio de protocolos específicos, como o IEEE 802.15.6, desenvolvido para oferecer comunicação segura e de baixo consumo energético em contextos voltados à área médica (?). Os dados coletados são geralmente encaminhados por dispositivos como smartphones ou unidades de processamento embarcadas, que atuam como gateways inteligentes, responsáveis por agregar, processar e enviar informações em tempo hábil para camadas de *fog* ou *cloud*, promovendo um acompanhamento contínuo e a geração de alertas preventivos com base em modelos preditivos de saúde (?).

Com essa infraestrutura conectada, o monitoramento dos sinais vitais torna-se o alicerce da vigilância clínica, pois descreve em tempo hábil o estado do paciente e permite reconhecer precocemente urgências ou algum problema de saúde. Quatro parâmetros são cruciais: FC, temperatura corporal, saturação periférica de oxigênio (SpO_2) e frequência respiratória (FR). Para cada um, é importante conhecer os intervalos de referência, entender os fatores que afetam suas variações, como idade e comorbidades, e escolher o método de medição mais adequado, seja em dispositivos vestíveis ou sensores dedicados. Cada métrica possui limites inferiores e superiores que devem ser ajustados ao perfil do indivíduo, motivando algoritmos adaptativos de alerta e priorização de dados.

- 1. Frequência cardíaca:** Em adultos em repouso, a FC oscila entre 60 e 100 batimentos

por minuto (bpm) (?). Atletas bem condicionados podem exibir bradicardias fisiológicas próximas de 40 bpm, ao passo que taquicardias sustentadas acima de 100 bpm suscitam investigação clínica. A medição contínua é viabilizada sobretudo por fotopletismografia (PPG) em relógios ou pulseiras, cuja análise reporta erro absoluto médio abaixo de $\pm 3\%$ em condições controladas e de $\pm 10\%$ em ambientes livres (?). Para análise de arritmias, o eletrocardiograma (ECG) permanece como sendo o padrão.

2. **Temperatura corporal:** Análise com 7.636 participantes redefiniu a faixa oral normal para 35,7–37,4 °, registrando queda média de 0,23 °em indivíduos com 60 anos ou mais (?). Em geral, valores superiores a 37,3 °já são interpretados como febre, sugerindo que o organismo está respondendo a um processo infeccioso ou inflamatório (?). Normalmente, essa medida é obtida por meio de termômetros convencionais aplicados nos de maneira axilar, oral ou retal (?). Recentemente surgiram dispositivos que também monitoram a temperatura da pele, um parâmetro que varia para manter o equilíbrio térmico interno, ampliando as possibilidades de acompanhamento contínuo.
3. **Saturação periférica de oxigênio (SpO_2):** A saturação periférica de oxigênio (SpO_2) expressa a proporção de hemoglobina periférica que se encontra ligada ao oxigênio. Em quadros infecciosos, a liberação de citocinas inflamatórias pode prejudicar a difusão gassosa nos alvéolos, fazendo esse percentual cair (?). Leituras inferiores a 95 % constituem um alerta precoce de possível dispneia. Em pessoas saudáveis, o valor costuma permanecer quase inalterado durante as atividades diárias. A medição rotineira da SpO_2 baseia-se na fotopletismografia: pares de emissores de luz vermelha e infravermelha posicionados em extremidades (dedos ou lóbulos da orelha) detectam variações de absorção ótica que permitem calcular o índice (?).
4. **Frequência respiratória (FR):** Esse parâmetro vital corresponde ao número de incursões respiratórias realizadas a cada minuto, conhecido em inglês como *breaths per minute* (BRPM). Estudos descrevem como faixa de normalidade valores entre 12 e 20 BRPM (?). Taxas inferiores a esse intervalo podem sinalizar comprometimento pulmonar, pois quando uma doença ataca os pulmões é notória a dificuldade de respirar bem e regular. A FR ganhou destaque na pandemia de COVID-19 devido à alta incidência de sequelas pulmonares. Métodos contemporâneos incluem cintas torácicas, microfones de contato e acelerômetros combinados a PPG. (?).

Em síntese, FC, temperatura corporal, SpO_2 e FR formam o núcleo mínimo de vigilância fisiológica. A confiabilidade de sistemas de priorização de dados em saúde conectada depende tanto da robustez dos sensores quanto da adaptação das faixas de referência às características individuais do paciente.

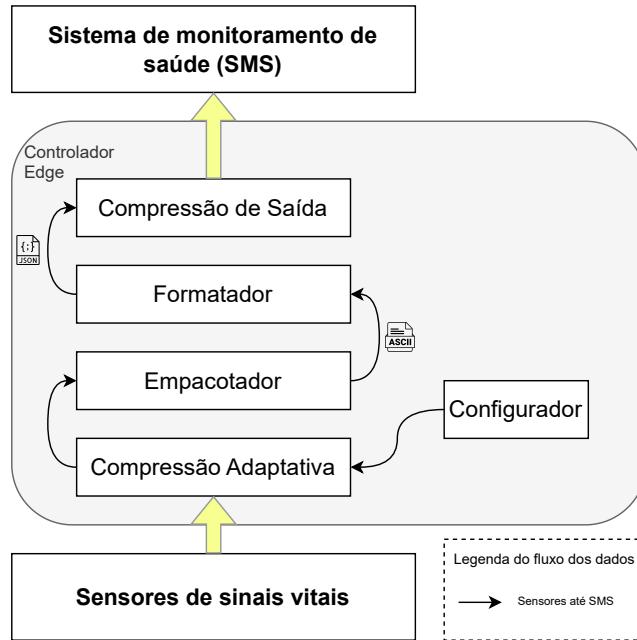
2.2 Vital Sign Adaptive Compressor

O atual trabalho terá em sua base e continuará o desenvolvimento do VSAC (Vital Sign Adaptive Compressor), que é um modelo de compressão de dados desenvolvido para ambientes de IoT aplicados ao monitoramento de sinais vitais em cidades inteligentes. O VSAC opera em uma arquitetura em camadas (*Edge–Fog–Cloud*) e combina métodos de compressão lossy e lossless para maximizar a retenção de informação clínica relevante ao mesmo tempo em que minimiza o volume de dados transmitidos.

2.2.1 Arquitetura do VSAC

O VSAC adota uma topologia em três níveis (*Edge–Fog–Cloud*), mas mantém toda a lógica de processamento no dispositivo de borda. As camadas superiores atuam como encaminhadoras e tomadoras de decisão baseadas no NEWS2, além de armazenarem o histórico clínico dos pacientes e proverem dashboards de visualização para a equipe médica (?). A Figura ?? ilustra a arquitetura completa e o fluxo de dados do modelo.

Figura 4: Arquitetura e funcionamento do VSAC



Fonte: Adaptado da arquitetura do VSAC de Alexandre Luis de Andrade

No Estágio 1 – Compressão adaptativa, o módulo Compressão com perdas realiza, em tempo hábil, a compressão com perdas dos sinais vitais brutos. Em seguida, o Integrador agrupa esses dados comprimidos ao identificador de origem correspondente, garantindo que cada fluxo de sinais permaneça rastreável. Já o módulo Empacotador, por sua vez, recebe os pacotes

vindos da etapa anterior e decide o que será feito: para pacotes de alta prioridade, o envio ao sistema de monitoramento ocorre imediatamente; para os demais, o módulo concatena múltiplas amostras e acrescenta metadados. Depois, o Formattedor de arquivo serializa os dados em JSON, produzindo um formato padronizado para a compressão sem perdas subsequente (?).

O Estágio 2 – Compressão de saída começa com o Decisão de Algoritmo, que seleciona Huffman ou LZW de acordo com o tamanho do arquivo. O Compressão de Saída aplica o algoritmo escolhido, gerando o pacote final, que é então encaminhado pelo Enviador rumo às camadas *fog* e *cloud*. No *fog* ocorre o primeiro processamento clínico, enquanto a *cloud* armazena o histórico completo e disponibiliza visualizações em dashboards (?).

Em resumo, o VSAC estabelece um fluxo em que compressão adaptativa, agregação inteligente e seleção dinâmica de algoritmos trabalham de forma coesa para reduzir drasticamente o volume de dados transmitido, mantendo, contudo, a integridade das informações essenciais à decisão clínica. Distribuir as funções entre processamento intensivo na borda, priorização na *fog* ou *cloud* e armazenamento histórico na nuvem melhora a eficiência da rede e a segurança do paciente. Isso garante respostas rápidas a eventos críticos sem sobrecarregar a infraestrutura de comunicação.

2.2.2 Compressão de dados no VSAC

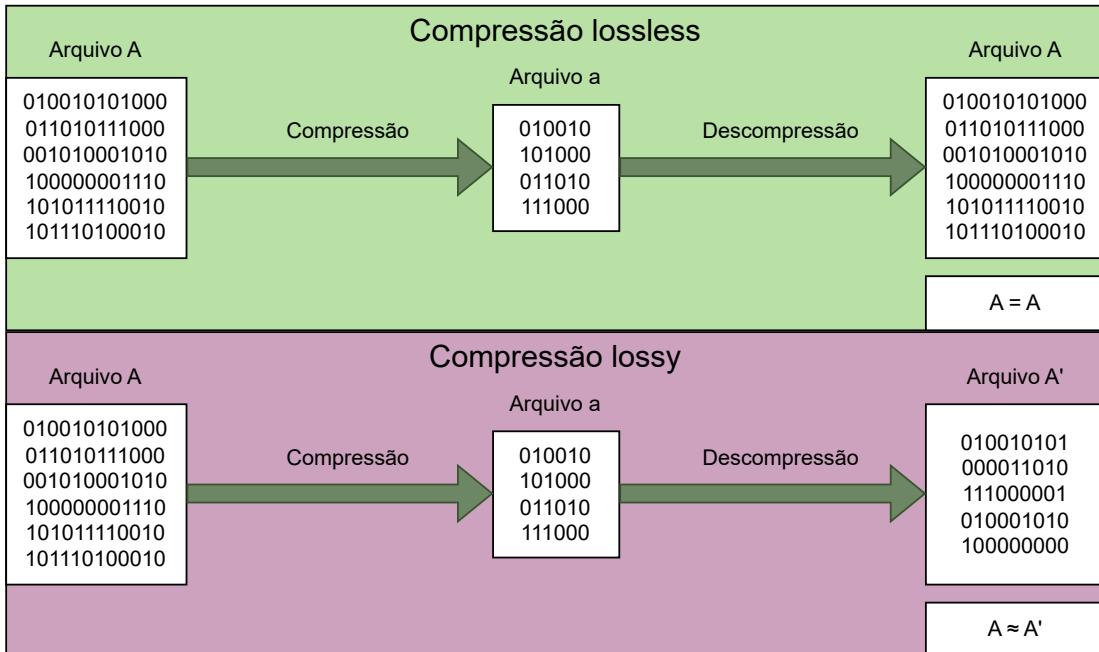
A compressão implementada pelo modelo VSAC baseia-se em duas fases principais encaixadas, explorando inicialmente uma abordagem *lossy*, seguida por uma abordagem *lossless*. A combinação dessas duas técnicas visa reduzir substancialmente o volume de dados transmitidos, assegurando que informações essenciais sejam preservadas para análise clínica, auditoria e armazenamento (?).

Os algoritmos que aplicam compressões do tipo *lossy* descartam componentes do sinal consideradas redundantes ou perceptualmente menos relevantes, obtendo taxas de redução elevadas ao custo de uma distorção limitada. O conteúdo reconstruído difere do original, mas permanece dentro de um erro máximo aceitável para a aplicação (?). Já algoritmos que aplicam técnicas de compressões *lossless* codificam o fluxo de dados sem perda de informação. Após a descompressão obtém-se exatamente os mesmos bits de entrada. Exemplos clássicos incluem Codificação de Huffman e Lempel–Ziv–Welch (LZW) (?). A Figura ?? demonstra a diferença entre ambos algoritmos de compressão.

Na primeira fase (*lossy*), o modelo VSAC utiliza o algoritmo denominado *Swinging Door Trending* (SDT), originalmente proposto para compressão de dados em sistemas SCADA industriais e adaptado para sinais biomédicos (?). O SDT caracteriza-se por descartar informações redundantes ou com variações mínimas, consideradas pouco relevantes do ponto de vista clínico, enquanto preserva picos e vales que representam eventos clinicamente significativos.

O funcionamento do SDT baseia-se no conceito de uma porta giratória imaginária, ilustrado na Figura ???. O algoritmo opera da seguinte forma:

Figura 5: Existem dois métodos principais de compressão: Lossless e Lossy. O método Lossy geralmente alcança uma taxa de compressão mais elevada, porém os dados recuperados após a descompressão não são idênticos aos do arquivo original.

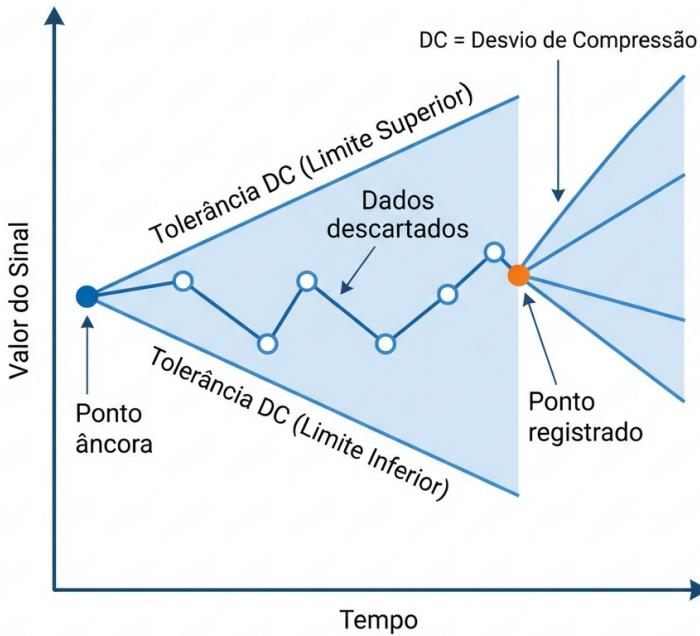


Fonte: Elaborado pelo autor

1. **Ponto Âncora:** O primeiro ponto da série é sempre registrado e serve como referência inicial.
2. **Tolerância (DC):** A partir do ponto âncora, duas retas de tolerância são projetadas, uma superior e uma inferior, definindo um corredor dentro do qual novos pontos são considerados redundantes. O parâmetro DC (*Compression Deviation*) define a largura dessa tolerância.
3. **Descarte de Pontos:** Enquanto os novos pontos permanecerem dentro do corredor de tolerância, eles são descartados (representados como círculos vazios na figura).
4. **Registro e Novo Âncora:** Quando um ponto viola o limite superior ou inferior, ele é registrado como ponto significativo e torna-se o novo ponto âncora, reiniciando o processo.
5. **Tempo Máximo (T_{SDT}):** Adicionalmente, se nenhum ponto for registrado após um intervalo máximo de tempo, o último ponto disponível é forçadamente registrado para evitar gaps prolongados na série.

Para exemplificar, considere uma série de FC com valores [72, 73, 71, 74, 72, 85, 86] bpm e $DC = 5$ bpm. O ponto 72 é o âncora inicial. Os pontos 73, 71, 74 e 72 permanecem dentro da tolerância (72 ± 5) e são descartados. O ponto 85 viola o limite superior e é registrado,

Figura 6: Funcionamento do algoritmo SDT. O ponto âncora projeta limites de tolerância (DC). Pontos dentro da tolerância são descartados; pontos que excedem o limite são registrados e se tornam o novo âncora.



Fonte: Elaborado pelo autor

tornando-se o novo âncora. O resultado comprimido seria [72, 85, 86], reduzindo 7 pontos para 3 (compressão de 57%).

Especificamente, os dados vitais brutos recebidos de sensores são processados pelo módulo Compressão Adaptativa. Inicialmente, um componente extrai e separa sinais vitais individuais, que então são comprimidos pelo SDT de acordo com parâmetros estabelecidos dinamicamente pelo módulo Configurador. Posteriormente, esses dados comprimidos são reintegrados em pacotes estruturados, contendo identificadores e metadados que descrevem as condições de coleta e prioridade.

Na segunda fase (*lossless*), o modelo VSAC aplica técnicas que garantem a reversibilidade completa dos dados após a descompressão. Os pacotes comprimidos na etapa anterior são agrupados pelo módulo Empacotador, formatados como arquivos JSON pelo Formatador, e então submetidos ao módulo Compressão de Saída. Neste módulo, o algoritmo aplicado depende do tamanho dos arquivos. Pacotes menores ou prioritários (até 33 kB) são processados usando a codificação de Huffman, mais eficiente para arquivos pequenos e com menor overhead computacional. Arquivos maiores seguem para a compressão Lempel–Ziv–Welch (LZW), que demonstra melhor desempenho para grandes volumes de dados, oferecendo taxas de compressão

superiores (?).

Essas etapas sequenciais garantem uma redução significativa do volume total de dados transmitidos pela rede, contribuindo diretamente para menor latência de comunicação, menor uso de largura de banda e uma redução dos requisitos de armazenamento na infraestrutura *fog/cloud*. Além disso, devido ao processamento ocorrer inteiramente na camada *edge*, os dados já chegam otimizados às camadas superiores, reduzindo a sobrecarga computacional e a necessidade de grandes capacidades de processamento em níveis mais altos da arquitetura.

Em resumo, os resultados experimentais apontam que essa abordagem híbrida (SDT + Huffman/LZW) apresenta desempenho superior às estratégias isoladas. O modelo VSAC obteve taxas de compressão até 42% superiores às obtidas com algoritmos exclusivamente **lossless** e até 46% superiores às alcançadas por algoritmos **lossy** utilizados isoladamente. Dessa forma, evidencia-se a eficácia do modelo VSAC em otimizar a transmissão de sinais vitais mantendo a qualidade necessária às análises clínicas (?).

2.3 Priorização

Em ambientes de IoT para saúde, nem todos os dados coletados possuem a mesma criticidade clínica. O grande volume de informações fisiológicas transmitidas continuamente por múltiplos sensores impõe a necessidade de priorização de dados, de modo a assegurar que informações potencialmente críticas (indicativas de deterioração do paciente) sejam transmitidas e processadas com rapidez, enquanto dados menos relevantes ou redundantes possam ser comprimidos, filtrados ou enviados com menor frequência. Essa priorização busca equilibrar dois objetivos: resposta clínica oportuna (detecção imediata de sinais de alerta) e uso eficiente de recursos de comunicação e hardware (evitando congestionamento de rede e sobrecarga de dados irrelevantes).

Uma abordagem comum para direcionar a priorização e a coleta e transmissão dos sinais vitais é o uso de escores clínicos de alerta precoce. Um escore amplamente conhecido para avaliação rápida do status do paciente é o National Early Warning Score 2 (NEWS2). O NEWS2 foi desenvolvido pelo Royal College of Physicians (RCP) no Reino Unido como uma forma padronizada de registrar e pontuar variáveis fisiológicas de pacientes, facilitando a identificação precoce do agravamento do estado de saúde e a resposta oportuna das equipes médicas (?). Essa relevância do NEWS2 pode ser compreendida a partir de alguns pontos detalhados abaixo:

1. **Implementação ampla e melhora na saúde pública.** O NEWS2 conta com a aprovação do *NHS England* e encontra-se implementado em 100% dos serviços de ambulância e 76% dos serviços de emergência britânicos, com estimativa oficial de prevenção de cerca de 1.800 óbitos anuais graças à padronização do reconhecimento precoce de agravamento do estado clínico (?).
2. **Desempenho prognóstico robusto.** Análise de 30 estudos publicada em 2023, que en-

volveu 185.835 pacientes, mostrou que o escore NEWS2 teve ampla capacidade de prever mortalidade nas próximas 48 horas. Os resultados apontaram uma área sob a curva (AUC) de 0,88, com sensibilidade e especificidade de 0,81, tornando o NEWS2 um dos escores mais confiáveis em ambientes de emergência (?).

3. **Viabilidade de automação no ecossistema IoT.** Estudo piloto realizado em 2024 mostrou que um dispositivo vestível, que usa sensores ópticos (fotopletismografia), conseguiu calcular automaticamente o escore NEWS2 com resultados muito semelhantes aos obtidos por profissionais de saúde ($K = 0,76$). O estudo não registrou eventos adversos e teve boa aceitação entre os pacientes, o que reforça o potencial de utilizar essa tecnologia em tempo hábil em camadas como *edge* e *fog* (?).
4. **Evolução e melhoria contínua.** Scoping review publicada em 2025 identificou 11 variantes do NEWS2 que aumentam a acurácia prognóstica ao incorporar idade, tendência temporal dos sinais vitais ou fração inspirada de oxigênio, evidenciando uma comunidade ativa de pesquisa e garantindo que futuras melhorias possam ser adotadas sem substituir a base já consolidada (?).

O NEWS2 baseia-se na avaliação de sinais vitais fundamentais: frequência respiratória, saturação periférica de oxigênio, temperatura corporal, pressão arterial sistólica, frequência cardíaca e nível de consciência. Além desses sinais, o sistema também considera a necessidade de oxigênio suplementar como um parâmetro adicional, influenciando a pontuação final (?).

Cada parâmetro monitorado pelo NEWS2 é associado a uma escala de pontuação de 0 a 3, refletindo o grau de desvio dos valores normais. A Figura ?? apresenta o sistema de pontuação do NEWS2. A soma desses escores resulta em um valor total que expressa o nível de risco do paciente, orientando intervenções clínicas conforme a gravidade observada. Além disso, o NEWS2 contempla a administração de oxigênio suplementar, que acrescenta dois pontos à pontuação total, reconhecendo o impacto significativo da necessidade de suporte respiratório (?).

O sistema também introduz duas escalas distintas para a interpretação da saturação periférica de oxigênio: a Escala 1, aplicada à maioria dos pacientes, e a Escala 2, destinada a indivíduos com condições como Doença Pulmonar Obstrutiva Crônica (DPOC), que apresentam parâmetros fisiológicos específicos. Essa adaptação visa garantir uma avaliação mais precisa e adequada de populações clínicas distintas, ajustando a resposta do sistema conforme a condição de base do paciente (?).

A Tabela ?? apresenta a classificação de risco clínico baseada no escore NEWS2 agregado, juntamente com a frequência mínima de monitoramento e a resposta clínica recomendada conforme as diretrizes do (?).

Para fins de nomenclatura neste trabalho, adotam-se três estados clínicos derivados das faixas de risco definidas pelo NEWS2 (?): *paciente estável* corresponde aos níveis de risco BAIXO ou MÍNIMO ($\text{NEWS2} \leq 4$); *paciente em alerta* corresponde ao risco MODERADO (NEWS2

Figura 7: O sistema de pontuação do NEWS

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

Fonte: (?)

Tabela 1: Classificação de risco clínico baseada no escore NEWS2 agregado

Escore NEWS2	Nível de Risco	Frequência de Monitoramento	Resposta Clínica
0	Mínimo	Mínimo a cada 12 horas	Continuar monitoramento de rotina
1–4	Baixo	A cada 4–6 horas	Informar enfermeiro responsável; avaliar necessidade de aumento de frequência
5–6	Moderado	No mínimo a cada hora	Revisão urgente por clínico competente; considerar transferência para área de maior cuidado
≥ 7	Alto	Monitoramento contínuo	Alerta emergencial imediato; avaliação por equipe de resposta rápida; considerar transferência para UTI

Fonte: Adaptado de (?).

entre 5–6); e *paciente crítico* corresponde ao risco ALTO ($\text{NEWS2} \geq 7$). Essa nomenclatura simplificada orienta as decisões de compressão e frequência de coleta ao longo de todo o modelo proposto.

No contexto deste trabalho, o modelo proposto utilizará a classificação baseada no escore NEWS2 para orientar a priorização dos dados transmitidos. O cálculo será realizado nas cama-

das *fog-cloud* que enviarão as classificações de risco para a *edge*, adaptando dinamicamente a frequência de coleta e o nível de compressão dos sinais vitais. Pacientes com maior risco clínico terão seus dados coletados com maior frequência e sofrerão compressão reduzida, preservando a integridade das informações críticas. Por outro lado, pacientes em condições estáveis terão dados coletados em intervalos mais espaçados e poderão ser submetidos a compressão mais intensiva, otimizando o uso dos recursos de rede e processamento, sem comprometer a segurança do monitoramento.

2.4 Projeto MHD

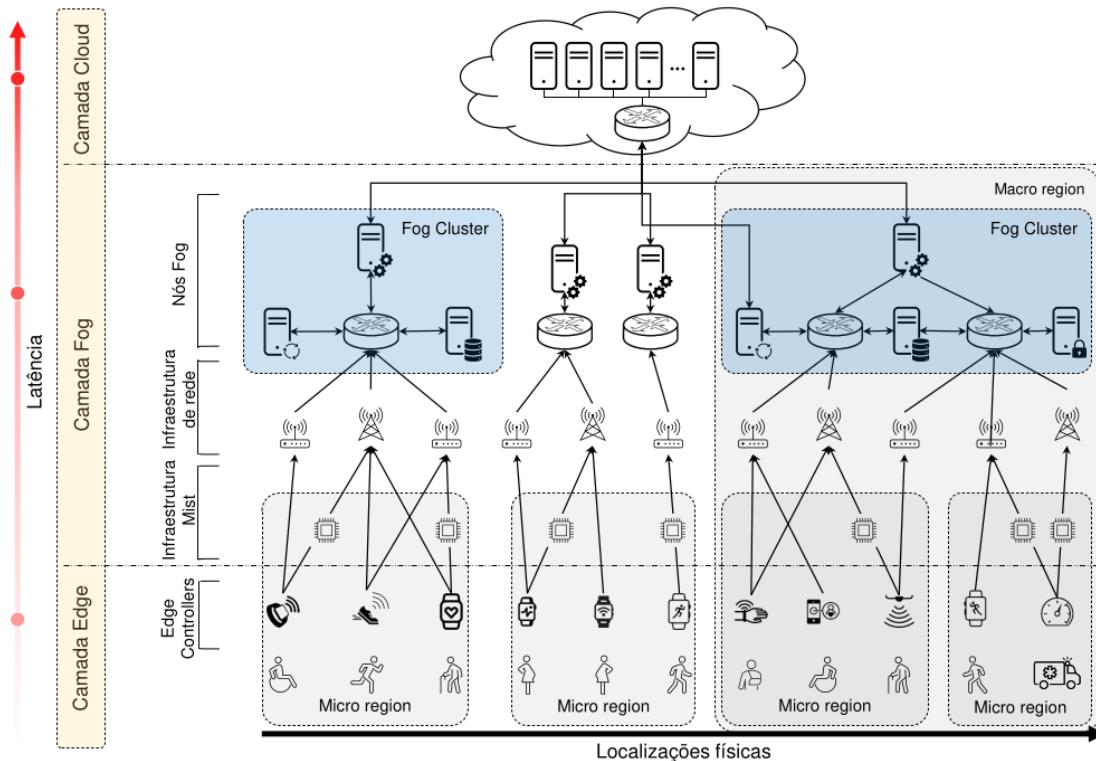
O modelo de gerenciamento de dados de sinais vitais apresentado neste trabalho integra o projeto Minha História Digital (MHD), financiado pela FAPERGS e desenvolvido por meio de uma cooperação científica entre a Unisinos, o Instituto Colaborativo de Blockchain (ICO-LAB) e uma aliança de pesquisa e inovação que reúne hospitais do Rio Grande do Sul.

O MHD busca criar uma solução computacional que integre arquitetura e algoritmos para rastrear e captar, em tempo hábil, os sinais vitais dos moradores da cidade. A iniciativa concentra-se no combate a doenças respiratórias, como a COVID-19, viabilizando a identificação da localização geográfica de cada indivíduo, a coleta pontual de seus sinais vitais e o armazenamento histórico dessas informações. Com isso, torna-se possível mapear a dispersão da doença a partir dos lugares visitados, avaliar a eficácia de medidas de quarentena e monitorar ou mesmo prever possíveis agravamentos nos parâmetros clínicos das pessoas acompanhadas.

Para alcançar o objetivo, o projeto contempla uma arquitetura em camadas que integra dispositivos *IoT*, nós de *Fog Computing* e serviços de *Cloud Computing*, buscando equilibrar latência, escalabilidade e consumo energético em aplicações de monitoramento de saúde. Essa solução tem como foco três objetivos fundamentais: (i) propor uma solução que abranja dispositivos vestíveis (pulseiras e *wearables*), leitores posicionados em locais estratégicos da cidade e sensores instalados em ambientes hospitalares; (ii) ajustar dinamicamente a quantidade de servidores e serviços nas camadas *fog* e *cloud* de acordo com a demanda computacional e o número de usuários, garantindo economia de energia e manutenção da qualidade de serviço (*QoS*); e (iii) definir, conforme a sensibilidade à latência, quais serviços e algoritmos precisam residir na *fog* e quais podem ser executados na *cloud*.

A Figura ?? ilustra a arquitetura do projeto MHD, organizada nas camadas *Edge–Fog–Cloud*. Na *edge*, indivíduos usam sensores de saúde disponíveis comercialmente para registrar sinais vitais ao longo das atividades diárias. Nós na camada de *fog*, distribuídos geograficamente, compõem uma infraestrutura próxima da fonte de captura, assumindo o processamento inicial dos dados. A *cloud* funciona como plataforma central que agrupa informações de toda a implantação e, apoiada por uma infraestrutura robusta, executa análises aprofundadas. Essa distribuição concede aos administradores de saúde uma visão abrangente de diferentes regiões e possibilita a detecção precoce de zonas de risco e possíveis surtos de doenças.

Figura 8: Visão da arquitetura da solução resultante da integração entre Internet das Coisas, Fog Computing e Cloud Computing.



Fonte: Adaptado do projeto Minha História Digital, de autoria de Rodrigo da Rosa Righi e Alexandre Luis de Andrade

No contexto do projeto MHD, o modelo proposto neste trabalho está situado na camada *edge*, onde executa a compressão adaptativa dos fluxos de sinais vitais de acordo com o escore clínico de cada paciente. Após a compactação, os dados são encaminhados às camadas *fog* e *cloud*, onde um serviço de avaliação recalcula continuamente esses escores com base no NEWS2 e, quando necessário, encaminha notificações de volta à *edge*. Esse ciclo de retroalimentação ajusta dinamicamente a frequência de coleta e os parâmetros de compressão, permitindo que pacientes em estado crítico sejam monitorados em maior resolução temporal enquanto se preserva largura de banda para casos estáveis, mantendo o equilíbrio entre eficiência operacional e relevância clínica.

3 TRABALHOS RELACIONADOS

Este capítulo tem como objetivo apresentar trabalhos relacionados com o modelo proposto. Com base no tema da pesquisa e nas questões apresentadas na Seção ??, foram analisados alguns trabalhos que seguem uma linha de pesquisa similar. Dessa forma, as próximas Seções descrevem como esses trabalhos foram selecionados e o estado-da-arte atual, descrevendo algumas abordagens a respeito de técnicas de compressão, transmissão e priorização de informações. Os artigos estão classificados em ordem crescente por ano de publicação e após por ordem alfabética dos nomes dos autores. Por fim, os trabalhos serão comparados e analisados para identificar oportunidades na área.

3.1 Metodologia de seleção

Esta seção descreve a metodologia de pesquisa adotada, os procedimentos utilizados e os trabalhos selecionados. Destacando as decisões tomadas por meio de uma revisão sistemática para fornecer uma visão geral das estratégias de compressão e priorização de dados para sinais vitais em sistemas de monitoramento de saúde. A abordagem apresentada identifica tecnologias, problemas e métodos utilizados neste campo. Além disso, a revisão inclui a identificação de estudos primários, a aplicação de critérios de inclusão e exclusão e a síntese dos resultados obtidos, buscando fornecer uma compreensão abrangente a cerca do tema explorado. O método de pesquisa foi utilizado por meio dos seguintes procedimentos, com base na revisão de (?):

- Estratégia de busca: Descreve as bibliotecas exploradas e a estratégia de coleta de dados.
- Seleção de artigos: Define como os artigos foram selecionados, apresentando os critérios de exclusão e as etapas para filtrar e selecionar os estudos.
- Software e ferramentas: Descreve as ferramentas e softwares utilizados para apoiar a pesquisa.

3.1.1 Estratégia de busca

Estudos foram selecionados de fontes de dados apropriadas para responder às questões de pesquisa, aumentando a probabilidade de encontrar informações pertinentes sobre compressão e priorização de dados na coleta de sinais vitais. A busca abrangeu as seguintes bases de dados eletrônicas:

- ACM Digital Library (<https://dl.acm.org/>)
- IEEE Xplore (<https://ieeexplore.ieee.org/>)
- Elsevier ScienceDirect (<https://www.sciencedirect.com/>)

- Springer Link (<https://link.springer.com/>)

Além disso, palavras-chave específicas foram definidas para criar uma sequência de busca, estruturada em unidades de busca e combinadas usando operadores booleanos. Essa sequência foi aplicada às bases de dados mencionadas, garantindo uma busca abrangente e detalhada.

```
(IoT OR sensors)
AND
(data compression OR compression)
AND
(prioritization OR prioritisation)
AND
monitoring
AND
(healthcare OR vital signs)
```

A sequência de busca abrange quatro grupos principais: IoT, compressão, priorização, monitoramento e saúde. Esses grupos são combinados usando o operador AND para garantir que os artigos selecionados abranjam todos esses tópicos. O primeiro grupo inclui termos relacionados a IoT. O segundo grupo concentra-se em palavras-chave associadas à compressão de dados. O terceiro grupo concentra-se em palavras-chave para priorização. Já o quarto grupo refere-se a monitoramento e, por fim, o quinto grupo refere-se a área da saúde. Essa estrutura de busca permite uma exploração abrangente de estudos relevantes para essas áreas de interesse.

3.1.2 Seleção dos artigos

A sequência de busca foi aplicada às bases de dados apresentadas na Seção ?? para selecionar os estudos. No total foram encontrados 748 artigos, incluindo todas as bases de dados. A busca se delimitou em artigos no idioma inglês publicados à partir de 2023 até maio de 2025. Essa escolha de período se deu por conta da pesquisa já realizada para a construção do VSAC, que se estendeu até 2023. Além disso, foram empregados critérios de exclusão compostos por três filtros (F1, F2 e F3) usados sequencialmente para selecionar as pesquisas mais relevantes.

Tabela 2: Filtros e critérios de inclusão

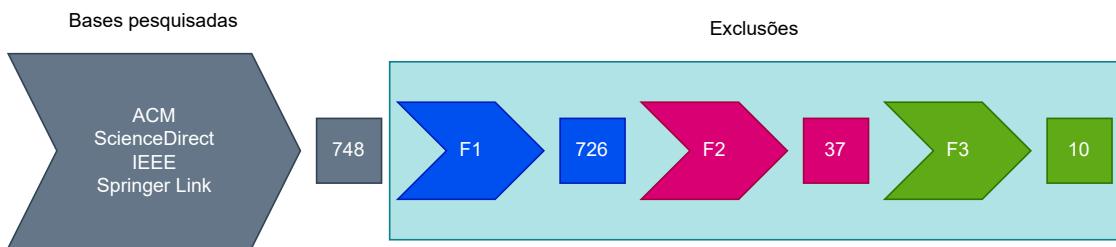
ID	Critério
F1	Eliminação de duplicadas
F2	Artigos que possuem a string de busca no título, resumo ou palavras-chave
F3	Leitura e análise do texto completo

-

Inicialmente, o F1 removeu todos os artigos duplicados. Após isso, o F2 excluiu artigos que não possuíam a string de busca no título, resumo ou palavras-chave. Nesse filtro a busca

foi adaptada, podendo o artigo tratar ou de compressão ou de priorização de dados. Por fim, no F3 os artigos foram lidos e analisados de maneira completa para selecionar os trabalhos voltados para compressão, priorização e/ou eficiência na transmissão de sinais vitais. A Figura ?? relata melhor os critérios de exclusão e os resultados obtidos. Já a Tabela ?? lista os trabalhos selecionados que serão detalhados na Seção ??.

Figura 9: Resultados da busca e filtragem



Fonte: Elaborado pelo autor

3.1.3 Software e ferramentas

Três aplicativos principais foram utilizados para auxiliar no processo de seleção dos artigos, além de planilhas do Google e arquivos CSV que serviram de suporte durante todo o processo. Primeiramente, para auxiliar na exportação dos resultados obtidos, foi utilizada a ferramenta Mendeley, que permitiu centralizar todos os artigos encontrados por meio da string de busca nas bases de dados. Após isso, os artigos selecionados foram exportados para a ferramenta Rayyan. Essa segunda ferramenta auxiliou na análise dos artigos, permitindo exclusões ou adições. Por fim, a última ferramenta foi um script desenvolvido pelo autor para automatizar a aplicação do filtro F2, buscando informações nos arquivos CSV exportados do Mendeley ou Rayyan. Com os resultados obtidos, os PDFs foram importados para a ferramenta Rayyan para melhor análise e leitura.

3.2 Estado-da-arte

A partir do resultado da pesquisa e filtragem dos trabalhos, foi possível obter 10 artigos de grande relevância para o entendimento do estado-da-arte. Tais trabalhos serão apresentados a seguir, destacando suas principais características, resultados e lacunas em aberto.

Tabela 3: Trabalhos relacionados

ID	Autor	Título	Ano
TR1	(?)	A lightweight and secure sensing model for body area networks in the Internet of Things in biological warfare applications	2023
TR2	(?)	Towards providing a priority-based vital sign offloading in healthcare with serverless computing and a fog-cloud architecture	2024
TR3	(?)	Lightweight Lossy/Lossless ECG Compression for Medical IoT Systems	2024
TR4	(?)	Compression of electrocardiogram signals using compressive sensing technique based on curvelet transform toward medical applications	2024
TR5	(?)	LPRLC: Linear Predictive Run Length Coding to Improve Energy Consumption of WBANs	2024
TR6	(?)	Blending lossy and lossless data compression methods to support health data streaming in smart cities	2025
TR7	(?)	IoT-driven remote health monitoring system with sensor fusion enhancing immediate medical assistance in distributed settings	2025
TR8	(?)	Research on Feistel Encryption Algorithm Based On Wireless Medical Sensor Networks	2025
TR9	(?)	F2multisense: a novel approach to fuzzy fusion in multisensor data to improve saving energy in WBSN	2025
TR10	(?)	A comprehensive evaluation of multiple video compression algorithms for preserving BVP signal quality	2025

Fonte: Elaborado pelo autor

3.2.1 TR1: A lightweight and secure sensing model for body area networks in the Internet of Things in biological warfare applications (?)

O estudo de (?) introduz um modelo leve e seguro de aquisição para *Body Area Networks* voltadas a cenários de guerra biológica. A proposta combina detecção comprimida adaptativa, cujo coeficiente de compressão é ajustado em função da esparsidade (entropia) dos dados, com uma etapa simultânea de cifragem baseada em matrizes sub-Gaussianas. Para sincronizar compressão e segurança, a chave criptográfica é derivada dinamicamente dos *inter-pulse intervals* do ritmo cardíaco do soldado, garantindo sigilo de ponta a ponta sem sobrecarga de processamento nos nós sensores. Avaliações em *MATLAB* demonstram economias energéticas de até 7 % sobre métodos de referência (*PALWSS*, *PEDTARA*, *DSCB*) e reduções equivalentes na latência de entrega dos pacotes, preservando relação sinal-ruído aceitável após reconstrução. Persistem, contudo, lacunas relevantes para aplicações clínicas amplas. A validação restringe-

se a simulações com quatro sensores e a um único cenário de paciente, não abrangendo sinais multivariados nem variações de criticidade clínica; além disso, o modelo supõe forte esparsidão dos sinais, apresentando alta complexidade de reconstrução quando essa premissa não se mantém. Falta ainda demonstração em hardware real, análise de escalabilidade em arquiteturas *Edge–Fog–Cloud* e integração de políticas de priorização baseadas em escores médicos para ajustar parâmetros de coleta em tempo de execução.

3.2.2 TR2: Towards providing a priority-based vital sign offloading in healthcare with serverless computing and a fog-cloud architecture (?)

O trabalho de (?) apresenta o modelo *SmartVSO*, uma arquitetura hierárquica *fog–cloud* que emprega computação *serverless* e uma fila na *cloud* para escalar o processamento de sinais vitais. O sistema calcula um *ranking* a partir da prioridade do usuário (estado clínico) e da prioridade do serviço (criticidade do algoritmo) e, com base nesse valor, decide se o sinal será executado localmente ou *offloaded*. Essa estratégia busca reduzir o tempo de resposta para pacientes em condições críticas, ao mesmo tempo em que preserva recursos nos nós de borda e garante escalabilidade em picos de carga, processando até 80000 sinais vitais em experimentos controlados. Persistem, contudo, lacunas relevantes: ausência de compressão dos sinais, o que pode sobrecarregar a rede em cenários de conectividade restrita; uso de limiares fixos de CPU para disparar o *offloading*, sem adaptação dinâmica a métricas como latência, consumo energético ou banda; inexistência de análise de segurança para criptografia e autenticação de dados sensíveis transmitidos entre as camadas *fog* e *cloud*.

3.2.3 TR3: Lightweight Lossy/Lossless ECG Compression for Medical IoT Systems (?)

O trabalho de (?) apresenta um compressor híbrido *lossy/lossless* leve para sinais de *ECG* em sistemas *IoT* médicos. A solução encadeia um modelo *lossy* baseado em otimização, implementado apenas com adições e multiplicações, a uma etapa *lossless* por *Huffman*, permitindo alcançar taxas de compressões médias superiores a 5:1 (CODE) e 4:1 (PTB-XL) com erro quadrático médio residual mínimo. Ao usar os sinais descompactados, o classificador realiza uma média de apenas 0,8 diagnósticos incorretos adicionais de um total de 402 casos, em comparação ao uso dos sinais originais não compactados, indicando preservação de informação diagnóstica mesmo após compressão agressiva. Tais resultados evidenciam potencial para reduzir custo de armazenamento local e tráfego de dados em monitoramento cardíaco contínuo em *edge* e *fog*. Persistem, contudo, algumas lacunas: o método restringe-se ao *ECG*, sem avaliar compressão multivariada de sinais vitais; inexistente mecanismo de priorização clínica que ajuste parâmetros de amostragem ou compressão conforme a criticidade do paciente; aspectos de segurança, latência de rede e consumo energético em arquiteturas *Edge–Fog–Cloud* não são mensurados em

ambiente real, limitando a extração para cenários de telemonitoramento escalável.

3.2.4 TR4: Compression of electrocardiogram signals using compressive sensing technique based on curvelet transform toward medical applications (?)

O trabalho de (?) apresenta uma abordagem de *compressive sensing* para sinais de ECG que emprega a *curvelet transform* a fim de gerar representações esparsas e, em seguida, codificação por *run-length* para minimizar o volume de dados transmitidos em *wireless sensor networks*. Avaliado com o repositório MIT-BIH, o método alcançou uma taxa de compressão de 15,7 e diferença média da raiz percentual de 2%, além de demonstrar redução de 91,35 % na carga de comunicação quando integrado a nós IoT, indicando impacto direto na eficiência energética e na largura de banda da rede. Entre as lacunas, podem ser citados: a pesquisa restringe-se a um único sinal vital (ECG), não avaliando compressão multivariada nem adaptações dinâmicas orientadas por escores clínicos; ausência de mecanismos de priorização que distingam dados críticos de pacientes em tempo hábil; e dependência de simulações em MATLAB, sem validação em arquiteturas *Edge–Fog–Cloud* reais ou análise de latência, segurança e consumo energético em múltiplas camadas.

3.2.5 TR5: LPRLC: Linear Predictive Run Length Coding to Improve Energy Consumption of WBANs (?)

O trabalho de (?) apresenta o esquema híbrido *LPRLC* (*Linear Predictive Run Length Coding*) para reduzir consumo energético e tráfego em *WBANs*. A solução combina *Linear Predictive Coding* (fase de predição) com *Run Length Encoding* (fase de compressão do erro residual), executando todo o processamento no nó sensor a fim de evitar custos de transmissão desnecessários. Avaliada em cinco sinais vitais (pressão arterial sistólica/diastólica, respiração, saturação de oxigênio e frequência cardíaca), a abordagem alcançou economia média de 98 % de energia e razão de compressão de até 70 : 1, superando algoritmos clássicos (*Huffman*, *Arithmetic*, *LZW*) em cenários simulados. Esses resultados evidenciam o potencial de técnicas leves de predição+compressão para prolongar a autonomia dos sensores e mitigar o gargalo de comunicação em sistemas de monitoramento corporal. Persistem, contudo, lacunas: ausência de mecanismos de priorização clínica que ajustem dinamicamente a taxa de amostragem conforme escores de criticidade; falta de validação em arquiteturas *Edge–Fog–Cloud*, sem métricas de latência, segurança ou confiabilidade ponta a ponta; e limitação a cenários simulados, sem testes *hardware-in-the-loop* que considerem interferências reais de rede e múltiplos pacientes.

3.2.6 TR6: Blending lossy and lossless data compression methods to support health data streaming in smart cities (?)

O trabalho de (?) propõe o modelo VSAC, destinado a otimizar o fluxo contínuo de sinais vitais em cenários de *smart cities*. A solução adota uma arquitetura *Edge–Fog–Cloud* e integra duas fases de compressão — *lossy* (adaptativa via *Swinging Door Tranding*) e *lossless* (Huffman ou LZW) — complementadas por um mecanismo dinâmico de priorização que ajusta parâmetros de amostragem e limiares de erro conforme notificações clínicas. Os autores demonstram reduções de até 46 % no volume transmitido em comparação com abordagens de compressão isolada, preservando níveis aceitáveis de distorção para uso em telessaúde. Persistem, entretanto, lacunas relevantes: a validação concentra-se em protótipos de laboratório e em um único sinal vital (*heart rate*), sem explorar a heterogeneidade multivariada comum na prática clínica; métricas de consumo energético e de impacto em rede real (latência, jitter) não são avaliadas; e priorização e apdatação na coleta de dados de pacientes de forma individual.

3.2.7 TR7: IoT-driven remote health monitoring system with sensor fusion enhancing immediate medical assistance in distributed settings (?)

O trabalho de (?) tem por objetivo principal desenvolver e validar um *Internet-of-Things–driven Health Monitoring System* (HMS) que integre sensores vestíveis, fusão de sensores e computação em nuvem para monitorar, em tempo hábil, temperatura corporal, frequência cardíaca, temperatura e umidade ambiente de pacientes em regiões remotas. A arquitetura proposta abrange três camadas — *device*, *cloud* e *user level* — conectadas por ZigBee/MQTT, possibilitando detecção de anomalias, emissão de alertas imediatos a profissionais de saúde e geração de relatórios por meio de um *Decision Support System*. A prova de conceito apresenta acurácia de 91,68 %, demonstrando viabilidade para telemedicina e redução da sobrecarga em unidades de saúde. Apesar dos avanços, permanecem lacunas relevantes: ausência de mecanismos de compressão que minimizem o volume de tráfego; e dependência de calibração manual dos limiares de alarme, o que pode gerar falsos positivos sob variações de contexto.

3.2.8 TR8: Research on Feistel Encryption Algorithm Based On Wireless Medical Sensor Networks (?)

O estudo de (?) apresenta um esquema integrado para redes de *Wireless Medical Sensor* capaz de proteger e otimizar o fluxo de sinais vitais. A proposta une um algoritmo de criptografia em bloco inspirado na estrutura *Feistel* e em mapeamento caótico inteiro, recurso que resolve o problema de períodos curtos, aumenta a aleatoriedade da chave e reduz o consumo de energia. Esse algoritmo trabalha junto a um método de compressão baseado em modelos de regressão, que ajusta dinamicamente a taxa de amostragem e os coeficientes do modelo. Com-

plementarmente, o sistema atribui prioridades a cada sinal vital, garantindo que dados de maior criticidade sejam transmitidos em tempo hábil, enquanto os demais são compactados para poupar largura de banda e prolongar a autonomia dos nós. A validação, implementada em *Java* e conduzida com registros de pulso, temperatura e pressão arterial de oito pacientes, demonstra melhora significativa na eficiência energética da rede e na confiabilidade da transmissão. Persistem, contudo, lacunas: a avaliação empírica restringe-se a um conjunto limitado de pacientes e não quantifica a distorção introduzida pela compressão; a priorização baseia-se apenas na importância atribuída aos sinais, sem empregar escores clínicos dinâmicos.

3.2.9 TR9: F²multisense: a novel approach to fuzzy fusion in multisensor data to improve saving energy in WBSN (?)

O estudo de (?) introduz o *F²multisense*, um esquema bifásico para economizar energia em *Wireless Body Sensor Networks*. Na camada do nó, apenas amostras com escores *NEWS* relevantes são transmitidas; na camada do coordenador, um sistema de inferência difusa ajusta dinamicamente a taxa de amostragem de cinco sinais vitais, alcançando reduções de 40 % no tráfego e 64 % no consumo energético em simulações com dados do *MIMIC-II*. Tal abordagem demonstra que a priorização orientada por criticidade clínica pode prolongar a vida útil dos sensores sem sacrificar a confiabilidade do monitoramento. Persistem, contudo, lacunas que limitam a generalização do *F²multisense*. O trabalho não incorpora algoritmos de compressão para cenários de banda restrita, nem avalia latência, jitter ou sobrecarga em arquiteturas *Edge–Fog–Cloud* reais. A validação baseia-se em apenas cinco sensores e em simulações MATLAB, sem testes em hardware ou em ambientes heterogêneos; além disso, aspectos de segurança e autenticação dos dados clínicos não são considerados.

3.2.10 TR10: A comprehensive evaluation of multiple video compression algorithms for preserving BVP signal quality (?)

O trabalho de (?) objetiva quantificar o impacto que diferentes algoritmos de compressão de vídeo exercem sobre a qualidade do sinal de *Blood Volume Pulse* (BVP) captado por *remote photoplethysmography* (rPPG). Para tanto, os autores construíram o conjunto de dados *ZJXU-MOTION*, gravado sem perdas e abrangendo variações de iluminação, intensidade de movimento e modalidades de câmera (*RGB/infravermelho*). Sobre esse acervo, comprimiram-se os vídeos com seis codecs predominantes (H.264, H.265, AV1/VP9, MJPEG, ProRes e FFV1) e quatro modos de *bit-rate control* (CRF, CQP, VBR e All-I), avaliando-se a correlação de Pearson entre BVP previsto (POS e PhysNet) e referência, bem como *Similarity Index* (SSIM) e o *Peak Signal-to-Noise Ratio* (PSNR). O estudo conclui que codificação H.265 preserva melhor o BVP em cenas estáticas, enquanto FFV1 se destaca sob movimentos intensos. Persistem, contudo, lacunas significativas: (i) a investigação limita-se a cenários laboratoriais de curta du-

ração, sem considerar degradações adicionais de rede ou processamento em *Edge–Fog–Cloud*; não se avalia a influência de políticas dinâmicas que ajustem parâmetros de compressão em tempo hábil face a mudanças de criticidade clínica; e o enfoque exclusivo em BVP deixa em aberto a generalização para sinais multivariados (SpO_2 , pressão arterial, variabilidade cardíaca).

3.3 Análise e Oportunidades

A Tabela ?? evidencia que, dos dez trabalhos analisados, sete (70 %) recorrem a algum tipo de *compressão*, em geral híbrida, mesclando abordagens com e sem perdas ((????????)).

Já a coluna Priorização mostra que apenas cinco estudos oferecem algum mecanismo para diferenciar a importância dos sinais: dois baseiam-se em escores clínicos dinâmicos (fuzzy-NEWS em (?)) e ranking usuário + serviço em (?)); dois utilizam regras de limiar fixo ou pesos de importância ((??)); e um adota níveis de prioridade pré-definidos, podendo ser alterados automaticamente mas não implementado e testado, (?)).

Quanto ao Ajuste em tempo hábil, ou seja, capacidade de alterar taxa de amostragem, profundidade de compressão ou decisão de offload durante a execução, metade da amostra (5/10) apresenta alguma forma de adaptação: variação de esparsidate do sinal ((?)), ajustes de CD/Tm ((?)), tempo de amostragem sensível ao risco ((??)) ou heurística CPU + ranking em arquitetura *fog–cloud* ((?)).

Contudo, nenhum estudo combina, de maneira integrada, compressão multissinal adaptativa e priorização clínica em ciclo fechado. Com base na análise, as principais lacunas identificadas são:

- inexistência de um mecanismo de **compressão adaptativa multivariada** que permita configurar limiares de erro individuais por sinal e por paciente;
- ausência de **priorização fundamentada em escores clínicos padronizados** (p.,ex. NEWS2) executada na *fog–cloud* para ajustar, em tempo hábil, a relevância dos sinais vitais;
- falta de um **ciclo fechado de ajustes** que sincronize taxa de amostragem, fator de compressão e política de transmissão à medida que o estado clínico do paciente evolui.

Essas lacunas justificam a proposta do ViSPAC, cujo objetivo é justamente integrar compressão adaptativa multissinal, priorização dinâmica baseada em escores clínicos e um circuito fechado de notificações dentro de um ecossistema *Edge–Fog–Cloud*, maximizando eficiência de banda e energia sem perder oportunidade clínica.

Tabela 4: Síntese dos trabalhos relacionados

TR	Compressão	Priorização	Ajustes na coleta
(?)	Com perdas	—	Variação esparsidade
(?)	—	Ranking usuário + serviço	CPU + ranking
(?)	Com perdas + Sem perdas	—	—
(?)	Com perdas + Sem perdas	—	—
(?)	Com perdas + Sem perdas	—	—
(?)	Com perdas + Sem perdas	Nível prioridade	CD/Tm dinâmico
(?)	—	Limiares de sinais vitais	—
(?)	Com perdas	Importância dos sinais	Tempo amostragem
(?)	—	NEWS (fuzzy)	Risco NEWS contínuo
(?)	Com perdas + Sem perdas	—	—

Fonte: Elaborado pelo autor

4 MODELO VISPACE

A análise dos trabalhos relacionados mostra duas lacunas principais: (i) ausência de mecanismos de priorização dinâmica, baseados em escores clínicos, para ajuste rápido da compressão e da frequência de coleta; e (ii) uso restrito de compressão híbrida (lossy + lossless) para apenas um parâmetro, o que limita seu uso em cenários com múltiplos sinais. Este estudo propõe o modelo ViSPAC, que consolida FC e SpO_2 e integra o escore NEWS2 nas camadas fog–cloud para classificar riscos e acionar notificações que reconfiguram a camada edge. O ciclo adaptativo aumenta a resolução e reduz a compressão de dados críticos, enquanto dados menos urgentes são compactados, otimizando a banda e os recursos computacionais sem perder acurácia.

4.1 Caso de uso

Esta seção apresenta um caso prático do ViSPAC, demonstrando seu potencial efetivo. A arquitetura adaptativa do ViSPAC equilibra a otimização de recursos e a disponibilização de informações clínicas oportunas à equipe de saúde. Considere uma residência inteligente com quatro moradores: mãe, pai, filho adolescente e avó com doença respiratória crônica. Todos usam dispositivos vestíveis para monitorar a frequência cardíaca, a saturação de oxigênio (SpO_2), a frequência respiratória, a temperatura corporal e a pressão arterial. Os dados são transmitidos por um dispositivo local ao hospital municipal responsável. No começo, todos têm sinais vitais normais, conforme o escore NEWS2. A Figura ?? apresenta o fluxo deste caso de uso.

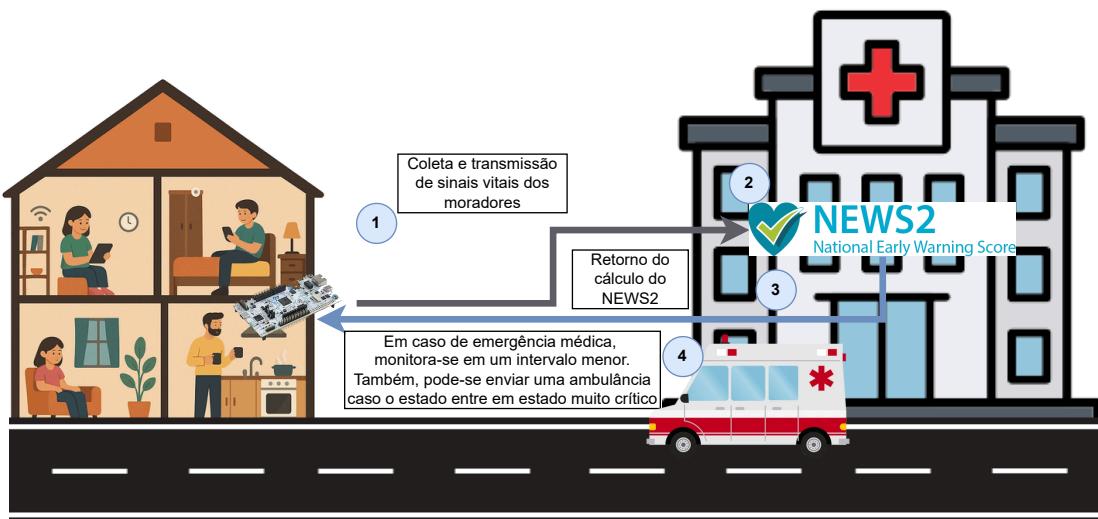
Enquanto os sinais vitais dos moradores permanecem estáveis, o ViSPAC envia relatórios menos frequentes, economizando bateria e largura de banda. Se houver queda na SpO_2 ou aumento da frequência respiratória, o sistema detecta e aumenta a transmissão de dados. Assim, os profissionais de saúde acompanham a evolução clínica e podem decidir entre o envio de ambulância e a orientação médica.

A avó, por ser de maior risco, é monitorada continuamente sem comprometer o desempenho da rede. O ViSPAC prioriza informações clínicas relevantes e aplica compressão aos dados menos importantes. Isso mantém vários sensores ativos sem saturar a internet. O exemplo mostra como o modelo adapta sua operação ao estado clínico de cada pessoa, reduzindo o tráfego quando possível e acelerando em situações de risco, resultando em respostas direcionadas e mantendo a eficiência e a escalabilidade.

4.2 Decisões de projeto

Conforme destacado na Seção ??, este trabalho integra o projeto MHD, propondo um método de compressão adaptativa de sinais vitais com priorização, desenvolvido na camada *Edge*. O modelo ViSPAC foi concebido para atender a necessidades clínicas e técnicas identificadas durante a revisão da literatura e em iniciativas anteriores, como o modelo VSAC, do qual herda

Figura 10: Caso de uso do modelo: Uma família tem seus sinais vitais monitorados e enviados ao hospital, que calcula o NEWS2 e, se necessário, retorna ajustes de parâmetros. Em casos extremos, a observação dos dados pode permitir ao hospital enviar uma ambulância para tratar de um caso urgente, por exemplo. O modelo funciona assim: 1) sinais vitais são coletados e analisados em casa e enviados ao hospital; 2) o hospital analisa os dados e calcula o escore NEWS2; 3) o escore NEWS2 é enviado de volta para ajustar as coletas; 4) em emergência, o hospital pode enviar uma ambulância ou orientar o paciente a ir ao pronto-socorro. Esse ciclo de coleta e envio exige compressão para reduzir o tamanho dos pacotes transmitidos.



Fonte: Elaborado pelo autor

diretrizes fundamentais para a compressão e o envio eficientes de dados clínicos.

Uma decisão fundamental foi escolher dois sinais vitais, FC e SpO_2 , para o monitoramento contínuo. A FC já havia sido validada no projeto VSAC (?). A SpO_2 foi escolhida por sua importância clínica em síndromes respiratórias agudas e em deteriorações rápidas da oxigenação. Trabalhos recentes dão suporte a essa escolha: (?) validaram compressão em cinco sinais vitais, incluindo FC e SpO_2 . Já (?) mostraram que esses parâmetros são determinantes no cálculo de escores de alerta precoce, como o NEWS. Ambos são usados no cálculo do escore NEWS2. Isso permite avaliar rapidamente o estado clínico do paciente e, com base nisso, ajustar parâmetros como o nível de compressão e o intervalo de coleta/transmissão. Os demais sinais do NEWS2 foram ajustados com os valores padrão de um paciente saudável (Tabela ??). Assim, é possível isolar o comportamento do modelo frente às variações de FC e de SpO_2 , sem o ruído de outros parâmetros.

O gerenciamento da coleta de dados utiliza um mecanismo dinâmico. O intervalo de aquisição (tempo entre coletas) varia conforme a gravidade clínica do paciente. Essa estratégia conta com respaldo em (?), cujo método F²multisense mostrou que ajustar a amostragem aos escores de criticidade reduz em 40% o tráfego e em 64% o consumo de energia. Assim, priorizam-se transmissões frequentes em casos de alto risco, assegurando vigilância contínua. Ao mesmo

Tabela 5: Valores padrão assumidos para sinais vitais não monitorados no cálculo do NEWS2

Parâmetro	Valor Padrão	Pontuação NEWS2	Justificativa
Frequência Respiratória	16 rpm	0	Valor central da faixa normal (12–20 rpm)
Temperatura	36,5 °C	0	Valor central da faixa normal (36,1–38,0 °C)
Pressão Arterial Sistólica	120 mmHg	0	Valor normotensivo típico
Nível de Consciência	Alerta (A)	0	Paciente orientado e responsável
Oxigênio Suplementar	Não	0	Paciente em ar ambiente

Fonte: Valores derivados das faixas normais do protocolo NEWS2 (?).

tempo, pacientes estáveis recebem um back-off exponencial (aumento progressivo do intervalo de coleta), o que reduz o uso de recursos de rede e de energia, sem prejudicar a detecção de eventos adversos.

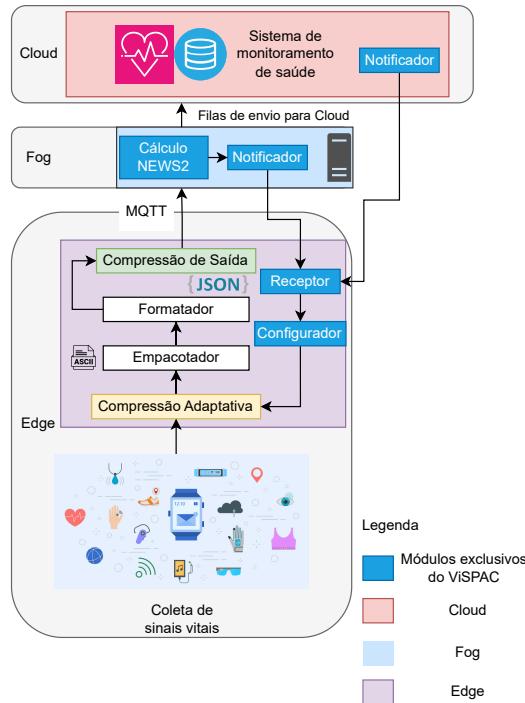
O ViSPAC adota uma arquitetura em três camadas (*Edge-Fog-Cloud*), conforme ilustrado na Figura ???. Essa escolha de arquitetura é respaldada pela literatura: (?) demonstraram que arquiteturas *fog-cloud* com escalonamento baseado em prioridade reduzem significativamente o tempo de resposta para pacientes críticos, enquanto (?) validaram a viabilidade de sistemas de monitoramento remoto em três camadas (*device-cloud-user*). Cada camada do ViSPAC possui papéis específicos para criar um sistema responsável e eficiente:

- Camada Edge: É onde a inteligência principal do ViSPAC opera. A coleta de sinais vitais por meio de sensores alimenta o módulo de compressão adaptativa. Os dados são então processados pelo Empacotador e Formato e, se necessário, passam pela Compressão de Saída antes de serem enviados à camada superior.
- Camada Fog: Recebe os dados via MQTT e calcula imediatamente o NEWS2, devolvendo o *feedback* ao *Edge* antes de qualquer comunicação com a nuvem. Os dados processados são então organizados em filas de prioridade por nível de risco (HIGH > MODERATE > LOW > MINIMAL) e encaminhados de forma assíncrona à *Cloud* por um *worker* dedicado. Esse desacoplamento garante que a latência percebida pelo *Edge* não depende da disponibilidade ou do tempo de resposta da nuvem.
- Camada Cloud: É responsável pelo armazenamento de longo prazo e pelas análises históricas. Embora o cálculo imediato ocorra na *Fog*, a nuvem mantém a capacidade de gerar notificações secundárias para a reconfiguração da *Edge* e realizar auditoria dos dados.

A escolha pelo protocolo MQTT fundamenta-se não apenas na sua arquitetura *publish-subscribe* (padrão de mensageria que desacopla publicadores de assinantes por meio de um intermediário), mas principalmente na redução drástica do *overhead* de cabeçalho. Diferentemente do HTTP, que exige cabeçalhos verbosos baseados em texto em cada requisição, o MQTT possui um cabeçalho fixo, binário, de apenas 2 bytes (?), o que reduz drasticamente o consumo de banda. Essa escolha é corroborada por (?), que utilizaram MQTT/ZigBee para

conectar sensores vestíveis a um sistema de monitoramento remoto, com acurácia de 91,68%. Além disso, o suporte a níveis de Qualidade de Serviço (QoS) permite ajustar a garantia de entrega conforme a criticidade do dado (e.g., QoS 1 para riscos altos, QoS 0 para monitoramento de rotina), otimizando o consumo de energia na transmissão.

Figura 11: Arquitetura macro do modelo ViSPAC.



Fonte: Elaborado pelo autor

A principal característica diferenciadora desta arquitetura é o ciclo de *feedback* fechado, identificado como uma lacuna na análise dos trabalhos relacionados (Seção ??). Enquanto (?) propuseram priorização baseada na importância atribuída aos sinais e (?) utilizaram inferência fuzzy para ajustar a taxa de amostragem, nenhum trabalho integra completamente a compressão adaptativa multissinal à priorização clínica em ciclo fechado. No ViSPAC, as notificações da nuvem são recebidas pelo módulo Receptor na *Edge*, que as repassa ao Configurador. Este, por sua vez, ajusta dinamicamente os parâmetros da Compressão Adaptativa e a frequência de coleta, adaptando o comportamento do sistema ao estado clínico atual do paciente. Os módulos destacados em azul na legenda representam as contribuições específicas do ViSPAC para a criação desse ciclo de priorização inteligente.

No que se refere à compressão, a abordagem híbrida combina o melhor de duas técnicas, seguindo uma tendência consolidada na literatura. (?) demonstraram que a combinação *lossy+lossless* para ECG alcança taxas de compressão superiores a 5:1 preservando a capacidade diagnóstica. A compressão *lossy*, por meio do algoritmo SDT, foi selecionada por sua eficiência

computacional ($O(n)$) e por sua adequação morfológica aos sinais monitorados. Sinais como a SpO_2 e a FC em repouso apresentam longos períodos de estabilidade (platôs) intercalados por mudanças pontuais; o SDT explora essa característica ao descartar pontos colineares dentro de uma margem de tolerância, preservando os picos e vales essenciais. (?) corroboram essa abordagem ao alcançar economia de 98% na energia com técnicas de predição e compressão leve em WBANs. Por outro lado, a compressão *lossless*, utilizando algoritmos como Huffman ou LZW, é aplicada aos dados já processados para garantir a integridade total em blocos críticos ou em reenvios. Essa combinação permite maximizar o ganho de compressão sem violar os limites de distorção aceitáveis para diagnósticos, seguindo a estratégia já validada no VSAC (?).

Finalmente, presume-se que os indivíduos monitorados estejam utilizando dispositivos vestíveis previamente associados ao sistema¹, com capacidade de transmissão contínua ou intermitente, conforme orientado pelo ViSPAC. Esses dispositivos devem operar de forma transparente no ambiente de uma cidade inteligente, permitindo o rastreamento contínuo e a emissão de alertas clínicos em tempo hábil.

Em resumo, as seguintes decisões de projeto compõem o modelo ViSPAC:

- Arquitetura: O modelo opera em uma estrutura de três camadas (*Edge-Fog-Cloud*), utilizando o protocolo MQTT para comunicação leve entre *Edge* e *Fog*. A camada *Edge* é responsável pela coleta e pela compressão adaptativa dos dados. Já a camada *Fog* assume o cálculo do escore de risco (NEWS2) para garantir uma resposta rápida à *Edge*. Por fim, a camada *Cloud* armazena e analisa os dados históricos e também pode gerar notificações para a camada *Edge*. Essa arquitetura encerra um ciclo de controle dinâmico.
- Sinais e Priorização: O sistema utiliza os sinais de FC e SpO_2 para calcular o escore NEWS2, que serve de base para priorizar os dados e ajustar dinamicamente a frequência de coleta e os parâmetros de compressão.
- Compressão Híbrida: A abordagem combina uma compressão com perdas (*lossy*) com o algoritmo SDT para otimizar a transmissão em períodos de estabilidade e uma compressão sem perdas (*lossless*) com Huffman/LZW para garantir a integridade de dados críticos.
- Serialização Binária: O ViSPAC adota o formato binário *MessagePack* para serialização dos dados entre *edge* e *fog*, em substituição ao JSON textual utilizado no VSAC. Essa escolha elimina o *overhead* de caracteres de delimitação e de codificação textual de números, além de permitir a operação direta dos algoritmos de compressão *lossless* sobre *bytes*, sem a necessidade de codificação Base64 intermediária.
- Validação de Hardware: Os testes devem ser executados com limitações de hardware para simular ambientes *edge*.

¹A integração física dos dispositivos vestíveis (protocolos de pareamento Bluetooth/Wi-Fi, autenticação e provisionamento) está fora do escopo desta dissertação. O modelo assume que a comunicação entre vestível e gateway *edge* já está estabelecida.

- Contexto Operacional: O modelo assume o uso contínuo de dispositivos vestíveis pelos pacientes em um ambiente de cidade inteligente, permitindo o monitoramento contínuo e a geração de alertas em tempo hábil.

4.3 Arquitetura

Conforme apresentado na seção anterior, a arquitetura do ViSPAC adota uma estrutura em três camadas (*Edge*, *Fog* e *Cloud*). Esta seção detalha os módulos internos e suas interações. A decisão de processar o escore NEWS2 na camada intermediária (*Fog*) fundamenta-se no princípio de computação ciente de localidade. Segundo (?), aplicações de IoT sensíveis à latência exigem que a inteligência computacional seja distribuída geograficamente, o que a aproxima da fonte dos dados. Ao realizar o cálculo do NEWS2 na *Fog*, o sistema minimiza o deslocamento de dados por meio da rede, reduzindo drasticamente o tempo de resposta em comparação a uma abordagem puramente baseada em nuvem.

O centro do ViSPAC reside em seis módulos executados na *Edge* e um executado nas camadas de *Fog/Cloud*. Cada um possui responsabilidades bem definidas e interage de forma orquestrada para garantir economia de recursos, sem perder a fidelidade aos dados críticos. Os módulos podem ser visualizados detalhadamente na Figura ??.

O grau de fidelidade e a frequência de monitoramento do ViSPAC são controlados dinamicamente por um conjunto de quatro parâmetros, definidos pelo Algoritmo 1 com base no risco clínico do paciente. Os parâmetros DC_s e $T_{SDT,s}^{\text{base}}$ ajustam a granularidade da compressão *lossy* para cada sinal $s \in \{\text{FC}, \text{SpO}_2\}$, controlando a precisão de forma individual. Esses valores são definidos de modo que $T_{SDT,s}^{\text{base}} = 2 \times IC_s^{\text{base}}$, garantindo a coerência entre o intervalo de coleta e o limiar de tempo do algoritmo SDT. Em paralelo, o IC e a ε controlam a frequência do ciclo de reavaliação do paciente. Valores reduzidos para estes parâmetros elevam a resolução e o consumo de banda; valores mais altos promovem a economia de recursos à custa da granularidade.

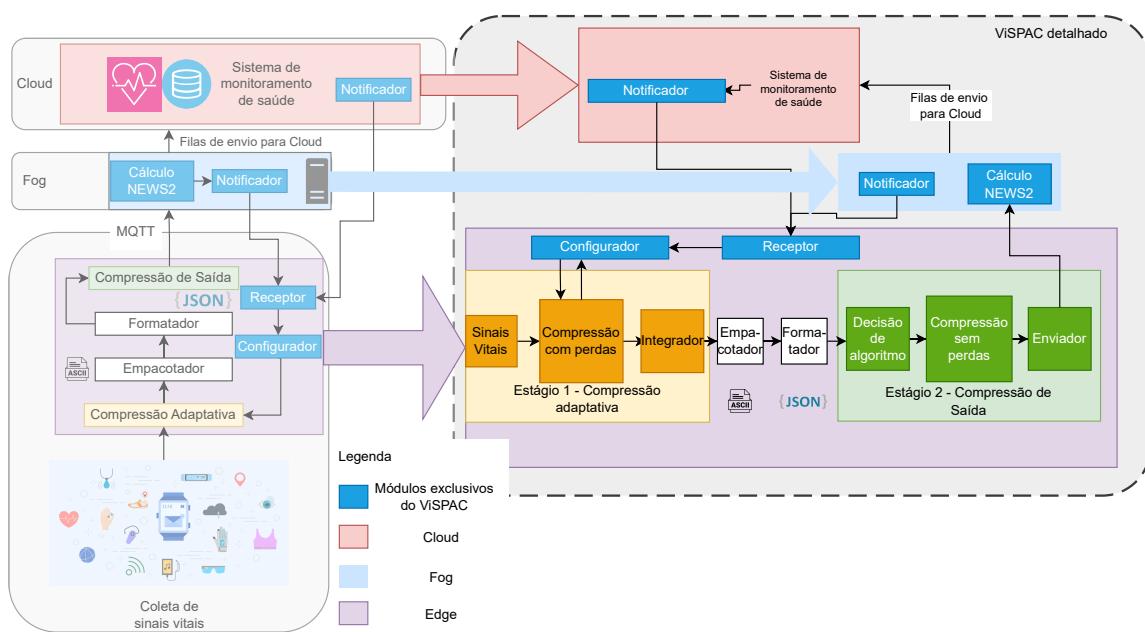
Abaixo, uma explicação de cada módulo do modelo:

- **Notificador:** presente tanto na *Fog* (para resposta rápida) quanto na *Cloud*. É responsável por emitir as notificações de reconfiguração para a *Edge* após o cálculo do NEWS2.
- **Receptor:** responsável por receber notificações das camadas de *Fog/Cloud*. Atua como ouvinte em uma fila de mensagens, filtrando as prioridades e repassando essas informações ao Configurador.
- **Configurador:** converte a prioridade clínica, expressa pelo escore NEWS2, em um conjunto de quatro parâmetros operacionais: IC , ε , DC e T_{SDT} . Os valores de DC e T_{SDT} são direcionados ao módulo de *Compressão Adaptativa*, enquanto IC e ε são utilizados para governar os algoritmos de controle de frequência (Algoritmos 2 e 3).

- **Compressão Adaptativa:** recebendo dinamicamente os parâmetros DC e T_{SDT} do Configurador, este módulo implementa o algoritmo SDT, realizando compressão *lossy* das séries de FC e de SpO_2 . Picos e vales que excedem DC ou que têm intervalos superiores a T_{SDT} são preservados; pontos redundantes são descartados. O módulo inclui a função de Integrador, que agrupa cada amostra ao identificador de origem para manter a rastreabilidade.
- **Empacotador:** recebe os pacotes comprimidos do estágio anterior e determina o destino imediato deles. Os pacotes são concatenadas até formar blocos de aproximadamente 1 MB ou até o tempo máximo definido (o que definirá se o pacote é de prioridade alta ou não). Também o módulo organiza os pacotes por risco clínico, agrupando-os em diferentes filas para posterior encaminhamento e envio.
- **Formatador:** serializa o bloco produzido pelo Empacotador em formato binário *MessagePack*, acrescentando campos de verificação de integridade e a versão do protocolo. Diferentemente do JSON textual empregado no VSAC, essa representação binária compacta reduz o tamanho do *payload* antes mesmo da compressão sem perdas e prepara o arquivo para o estágio seguinte.
- **Compressão de Saída:** aplica compressão *lossless*. Arquivos com tamanho até 33 kB usam o algoritmo de Huffman, reduzindo a latência de processamento; arquivos maiores utilizam o algoritmo LZW, obtendo taxas de compactação superiores. O resultado final é encaminhado para transmissão à *Fog* via MQTT pelo Enviador, onde novos cálculos do NEWS2 poderão originar notificações de retroalimentação, gerando, assim, um *loop* entre monitoramento e classificação.

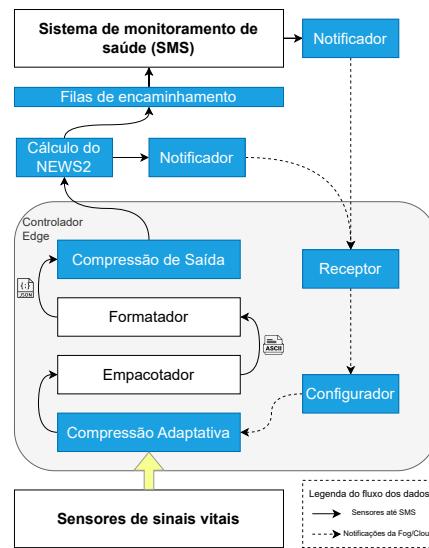
A Figura ?? compara os módulos do ViSPAC com os do VSAC. A principal diferença é que o ViSPAC implementa um ciclo completo de notificações da *fog-cloud*, ajustando dinamicamente a coleta, a compressão e a prioridade dos sinais vitais. Os módulos comunicam-se por interfaces definidas, formando um pipeline em que as decisões de prioridade controlam a granularidade dos dados transmitidos, economizando banda sem comprometer os dados clínicos nas camadas superiores.

Figura 12: Módulos do ViSPAC em micro. Na camada de Edge estão presentes os módulos de Compressão Adaptativa, compostos pela compressão com perdas e pelo Integrador; Empacotador e Formatador; Compressão de Saída, responsável por aplicar a compressão sem perdas e realizar o envio; Receptor, responsável por receber as notificações; e o Configurador, responsável por ajustar os parâmetros da compressão com perdas. Na camada de Fog, residem o cálculo primário do NEWS2 e o gerenciamento de filas, enquanto a *Cloud* armazena o histórico, ambos podendo enviar notificações pelo Notificador para a camada de Edge, continuando o ciclo fechado.

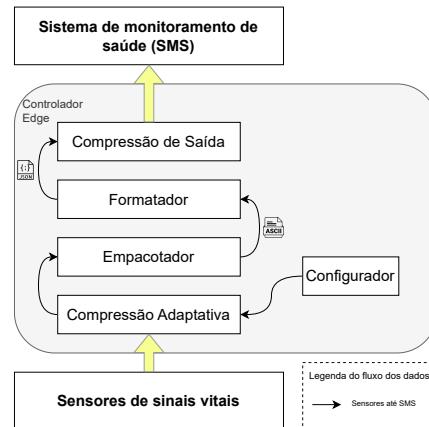


Fonte: Elaborado pelo autor

Figura 13: Comparação entre os módulos dos modelos ViSPAC e VSAC. Na primeira imagem podem-se observar os módulos adicionados ao ViSPAC em razão do VSAC (em azul). São eles: Notificador, Receptor, Configurador, Formatador e alterações na Compressão de Saída e na Compressão Adaptativa, para que comportem o novo modelo. Além disso, o ViSPAC propõe o uso de mais de um sinal vital e implanta ajustes automáticos no módulo Configurador. Ao receber uma notificação do Sistema de monitoração de saúde, o sistema adapta os parâmetros da Compressão Adaptativa aos novos valores, conforme definido na Seção ??.



(a) Módulos do ViSPAC



(b) Módulos do VSAC

Fonte: Figuras adaptadas de Andrade (2025) (?)

4.4 Funcionamento

Esta seção apresenta e detalha os três algoritmos centrais do modelo. Inicialmente, o sistema avalia o risco clínico do paciente com o escore NEWS2, que define a intensidade do monitoramento por meio do Algoritmo ???. Para pacientes estáveis, o Algoritmo ?? é ativado para otimizar recursos, aumentando gradualmente o intervalo entre as verificações por meio de um mecanismo de back-off exponencial (?). Essa abordagem, entretanto, pode acarretar risco em períodos de menor vigilância. Para mitigar essa limitação, o Algoritmo ?? implementa vigilância contínua, semelhante a um mecanismo de keep-alive (?), que permanece em constante observação. Ao detectar qualquer anomalia, o sistema aciona uma reavaliação completa e restabelece o monitoramento intensivo. Dessa forma, o sistema adapta-se de forma inteligente, ajustando o foco e a intensidade do monitoramento às condições clínicas de cada paciente.

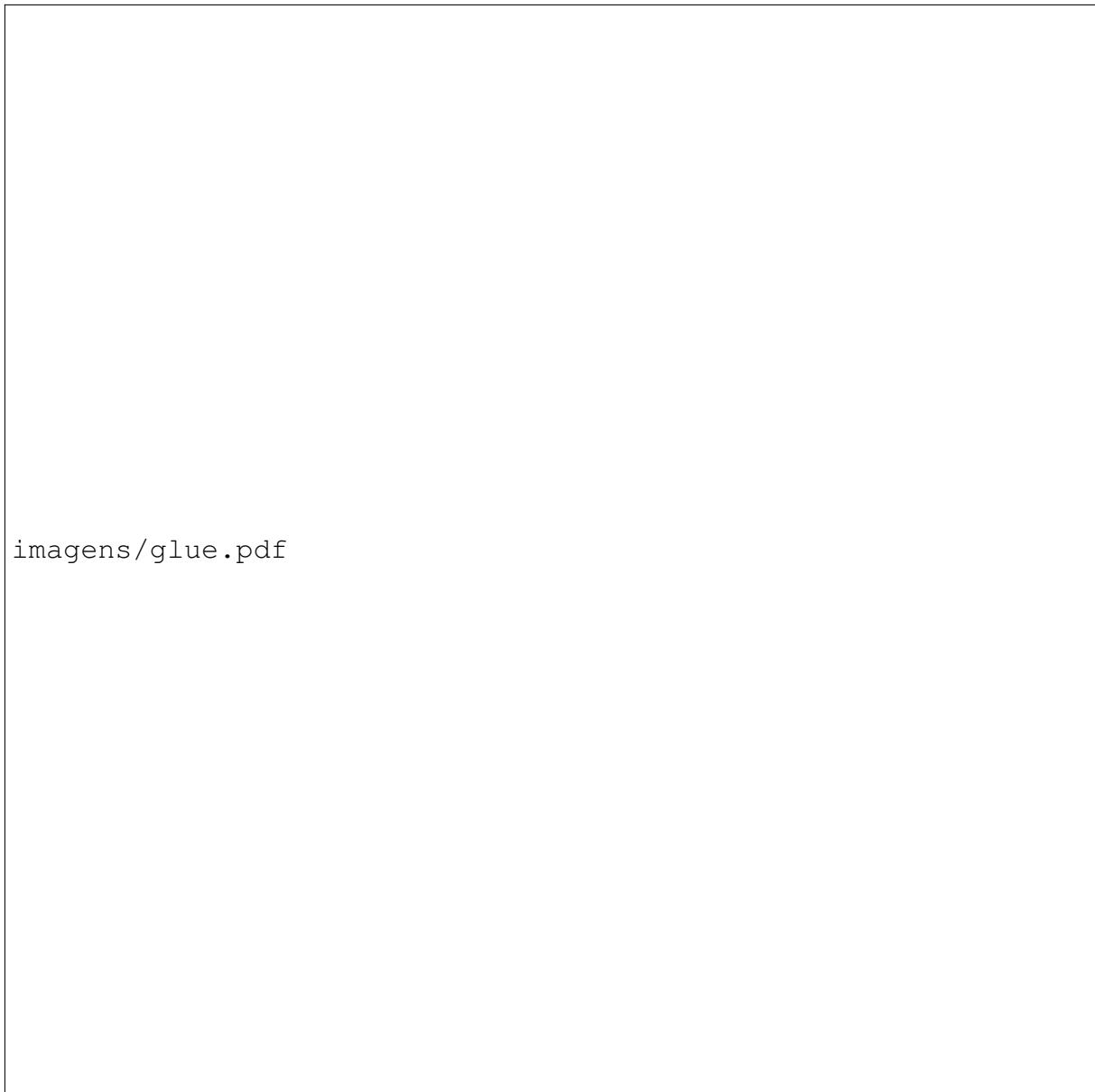
A Figura ?? ilustra o fluxo operacional e a interação entre os três algoritmos propostos do modelo ViSPAC. O Algoritmo 1 define os parâmetros de monitoramento com base no escore NEWS2. Com essa configuração, o Algoritmo 2 otimiza o sistema, aumentando o intervalo entre as coletas (*back-off*) durante períodos de estabilidade. Em paralelo, o Algoritmo 3 atua como uma vigilância contínua, realizando verificações rápidas e forçando a coleta imediata caso detecte uma variação fora do limite estabelecido nos sinais vitais. Essa dinâmica garante que o sistema seja eficiente em recursos, mas permaneça altamente responsivo a qualquer deterioração clínica.

4.4.1 Algoritmo 1 - Configurador por Sinal Vital

O Algoritmo 1: Configurador por Sinal Vital propõe um ajuste dinâmico da configuração de monitoramento de sinais vitais, com foco especial em FC e SpO_2 . O propósito é definir um conjunto completo de parâmetros operacionais, adaptados ao risco clínico do paciente, conforme determinado pelo escore NEWS2. Estes parâmetros incluem o intervalo de coleta (IC_s), que podemos comparar ao *time-to-live* (TTL) do protocolo DNS, no qual dados estáveis podem permanecer em cache por mais tempo, enquanto mudanças acionam atualizações imediatas (?). Além desses, incluem-se a margem de erro tolerável (ε_s), o desvio máximo de compressão (DC_s) e o intervalo máximo entre amostras ($T_{SDT,s}$), todos definidos individualmente por sinal. Em outras palavras, pacientes de maior risco terão seus sinais monitorados com maior frequência e menor compressão (com maior precisão), enquanto pacientes estáveis poderão ser acompanhados em intervalos maiores, economizando recursos computacionais sem comprometer a segurança e a qualidade das informações.

A literatura clínica demonstra que a frequência de monitorização de sinais vitais deve aumentar conforme o risco de deterioração do paciente. Diretrizes do NEWS2 recomendam monitorização mínima a cada 12 horas para pacientes com escore 0 (risco mínimo), passando para intervalos de 4–6 horas se o NEWS2 for entre 1 e 4 (risco baixo), e aumentando para pelo menos

Figura 14: Visão unificada do pipeline do ViSPAC do ponto de vista algorítmico. À coleta inicia, seguindo para a compressão adaptativa e envio para o fog. Quando a notificação chega, o Algoritmo 1 aplica novos parâmetros (se necessário) e o Algoritmo 2 verifica se o paciente está estável. Caso sim, aplica-se o back-off, caso contrário, mantém o tempo base configurado pelo Algoritmo 1. Com isso, o ciclo se repete. Em paralelo, o Algoritmo 3 fica monitorando, realizando coletas leves e comparando a margem de erro, os sinais vitais. Caso detecte uma anomalia, dispara uma coleta forçada e o ciclo se inicia novamente.



imagens/glue.pdf

Fonte: Elaborado pelo autor

1 vez por hora em pacientes com NEWS2 de 5 a 6 (risco moderado) ou com qualquer parâmetro individual muito alterado (pontuação 3). Para pacientes críticos ($\text{NEWS2} \geq 7$), indica-se monitoramento contínuo sempre que possível (?). Esses padrões refletem a necessidade de coletar dados vitais com maior frequência quando a probabilidade de mudanças clínicas relevantes é maior. Com base nessas informações, o algoritmo proposto adota uma abordagem mais conser-

vadora do que as diretrizes padrão. Foram estabelecidos tempos de coleta para cada categoria de risco, correspondentes à metade do intervalo máximo definido pelo NEWS2. Por exemplo, para um paciente com escore 0, em vez de monitoramento a cada 12 horas, a coleta é programada a cada 6 horas. O objetivo desta estratégia é ampliar a janela de detecção de eventos adversos, proporcionando uma camada adicional de segurança ao paciente.

Algoritmo 1: Configuração Dinâmica de Parâmetros por Sinal Vital

Entrada: Escore NEWS2 do paciente

Saída : Parâmetros:

$IC_{FC}, IC_{SpO_2}, \varepsilon_{FC}, \varepsilon_{SpO_2}, DC_{FC}, DC_{SpO_2}, T_{SDT,FC}, T_{SDT,SpO_2}, risco$

// Classificação do risco clínico

```

1 if NEWS2 ≥ 7 then
2   | risco ← ALTO
3 else if 5 ≤ NEWS2 ≤ 6 then
4   | risco ← MODERADO
5 else if 1 ≤ NEWS2 ≤ 4 then
6   | risco ← BAIXO
7 else
8   | risco ← MÍNIMO

```

// Parâmetros operacionais e de compressão conforme risco

// $T_{SDT} = 2 * IC$ para cada sinal

```

9 switch risco do
10   case ALTO do
11     |  $IC_{FC} \leftarrow 15\text{ s}; \varepsilon_{FC} \leftarrow \pm 2\text{ bpm}; DC_{FC} \leftarrow 2\text{ bpm}; T_{SDT,FC} \leftarrow 30\text{ s};$ 
12     |  $IC_{SpO_2} \leftarrow 15\text{ s}; \varepsilon_{SpO_2} \leftarrow \pm 1\%; DC_{SpO_2} \leftarrow 1\%; T_{SDT,SpO_2} \leftarrow 30\text{ s};$ 
13     | // Compressão mínima, alta fidelidade
14   case MODERADO do
15     |  $IC_{FC} \leftarrow 2\text{ min}; \varepsilon_{FC} \leftarrow \pm 4\text{ bpm}; DC_{FC} \leftarrow 4\text{ bpm}; T_{SDT,FC} \leftarrow 4\text{ min};$ 
16     |  $IC_{SpO_2} \leftarrow 3\text{ min}; \varepsilon_{SpO_2} \leftarrow \pm 1\%; DC_{SpO_2} \leftarrow 1\%; T_{SDT,SpO_2} \leftarrow 6\text{ min};$ 
17   case BAIXO do
18     |  $IC_{FC} \leftarrow 5\text{ min}; \varepsilon_{FC} \leftarrow \pm 6\text{ bpm}; DC_{FC} \leftarrow 6\text{ bpm}; T_{SDT,FC} \leftarrow 10\text{ min};$ 
19     |  $IC_{SpO_2} \leftarrow 10\text{ min}; \varepsilon_{SpO_2} \leftarrow \pm 2\%; DC_{SpO_2} \leftarrow 2\%; T_{SDT,SpO_2} \leftarrow 20\text{ min};$ 
20   case MÍNIMO do
21     |  $IC_{FC} \leftarrow 10\text{ min}; \varepsilon_{FC} \leftarrow \pm 10\text{ bpm}; DC_{FC} \leftarrow 10\text{ bpm}; T_{SDT,FC} \leftarrow 20\text{ min};$ 
22     |  $IC_{SpO_2} \leftarrow 15\text{ min}; \varepsilon_{SpO_2} \leftarrow \pm 3\%; DC_{SpO_2} \leftarrow 3\%; T_{SDT,SpO_2} \leftarrow 30\text{ min};$ 
23     | // Compressão máxima, baixa fidelidade

```

22 return $\langle IC_{FC}, IC_{SpO_2}, \varepsilon_{FC}, \varepsilon_{SpO_2}, DC_{FC}, DC_{SpO_2}, T_{SDT,FC}, T_{SDT,SpO_2}, risco \rangle$

A Tabela ?? resume a lógica dos intervalos de coleta (IC) definidos. Conforme detalhado, os valores foram dimensionados para o contexto de monitoramento contínuo, a fim de garantir vigilância em tempo hábil (próximo ao real). Vale ressaltar que **todos os parâmetros apresentados são configuráveis pelo administrador do sistema**, permitindo a adaptação do modelo a diferentes políticas institucionais ou protocolos clínicos específicos.

Tabela 6: Parâmetros do Algoritmo 1 e justificativa

Risco	NEWS2	Diretriz NEWS2	IC_{ViSPAC} (FC)	Justificativa
ALTO	≥ 7	Contínuo	15 s	Máxima vigilância; aproxima monitoramento contínuo
MODERADO	5–6	$\geq 1 \times / \text{hora}$	2 min	Vigilância frequente para detecção precoce de deterioração
BAIXO	1–4	4–6 horas	5 min	Equilíbrio entre economia de recursos e segurança clínica
MÍNIMO	0	12 horas	10 min	Máxima economia; paciente estável dispensa monitoramento intensivo

Fonte: Elaborado pelo autor.

Uma decisão de projeto importante diz respeito à relação entre IC_s e $T_{SDT,s}$. O parâmetro T_{SDT} define o intervalo máximo que o algoritmo SDT pode manter sem registrar um ponto, mesmo que o sinal permaneça dentro da tolerância DC_s . Para garantir coerência operacional, adotou-se a relação $T_{SDT,s}^{\text{base}} = 2 \times IC_s^{\text{base}}$ para cada sinal. Essa proporção assegura que o SDT force um registro apenas após pelo menos duas coletas consecutivas estáveis, evitando registros redundantes em períodos de baixa variabilidade fisiológica. Além disso, como FC e SpO_2 possuem intervalos de coleta distintos (como apresentado no Algoritmo 1), cada sinal recebe seu próprio valor de T_{SDT} , permitindo um ajuste fino na compressão de acordo com as características clínicas de cada parâmetro.

Diferentes sinais podem ter impacto distinto no escore: uma alteração sutil na SpO_2 pode elevar consideravelmente o NEWS2 (por exemplo, SpO_2 caindo de 96% para 92% adiciona 2 pontos), enquanto a FC exige variações mais amplas para alterar a pontuação (FC saindo de 80 para 109 bpm adiciona 1 ponto) (?). Além disso, para medir o SpO_2 em relógios e pulseiras, o sensor calcula a média do sinal ao longo de 8–12 s. Isso reduz o ruído, mas também atrasa, por alguns segundos, a exibição de uma queda real na saturação (?). A FC, por sua vez, pode ser mostrada batimento a batimento. Assim, aceitar um erro de até 5 bpm na FC quase nunca compromete a análise (?), enquanto um erro de 3–4% na SpO_2 pode esconder o início de um caso clínico mais grave. Com base nessas tolerâncias clínicas, os valores de ε (margem de erro) apresentados no Algoritmo 1 foram definidos: para FC, as margens variam de ± 2 bpm (risco ALTO) a ± 10 bpm (risco MÍNIMO); para SpO_2 , as margens são mais rígidas, variando de $\pm 1\%$ (riscos ALTO e MODERADO) a $\pm 3\%$ (risco MÍNIMO). Essa parametrização garante confiabilidade onde ela é mais crítica, enquanto permite maior tolerância em pacientes estáveis.

4.4.2 Algoritmo 2 - Back-off Exponencial Condicionado ao Risco

Depois que o Algoritmo 1 define todos os parâmetros operacionais para cada paciente p e para cada sinal $s \in \{\text{FC}, \text{SpO}_2\}$, o Algoritmo 2 entra em ação para ajustar dinamicamente o intervalo de coleta. Ele trabalha com as leituras que passam pelo processo de compressão, usando o intervalo-base $IC_{s,p}^{\text{base}}$ e a margem de erro ε_s para implementar um mecanismo de

back-off exponencial. Esse conceito foi inspirado em protocolos de rede como Ethernet e TCP (?).

A escolha por uma estratégia exponencial (2^n) em vez de linear justifica-se pela necessidade de reduzir rapidamente o congestionamento da rede e o consumo de bateria quando o paciente está estável por períodos prolongados. A lógica é simples: se a variação do sinal vital permanece dentro do limite ϵ por K ciclos seguidos, a chance de uma mudança significativa no próximo ciclo diminui, o que permite intervalos maiores entre as coletas.

Porém, diferente dos protocolos de rede (em que o tempo de espera pode dobrar após uma única transmissão bem-sucedida), o Algoritmo 2 adota uma abordagem mais conservadora, apropriada ao contexto clínico. Ele introduz um limiar de estabilidade K , que determina quantas leituras consecutivas dentro da margem ϵ_s são necessárias antes de aplicar o *back-off*. O valor $K = 3$ foi adotado como padrão, representando um equilíbrio entre responsividade e robustez: um valor menor (ex.: $K = 1$) permitiria *back-off* prematuro baseado em leituras isoladas potencialmente ruidosas, enquanto um valor maior (ex.: $K = 5$) atrasaria excessivamente a economia de recursos em pacientes genuinamente estáveis. Apenas após K medições estáveis, o intervalo é dobrado. Isso funciona como um mecanismo de confiança que evita o relaxamento do monitoramento baseado em apenas algumas leituras isoladas.

Essa estratégia só funciona enquanto o paciente está em risco MODERADO, BAIXO ou MÍNIMO. Se o risco subir para ALTO ou se qualquer leitura ultrapassar ϵ_s , o algoritmo restaura imediatamente o IC padrão. Além do ajuste dinâmico do intervalo de coleta (IC_s^{curr}), o Algoritmo 2 também atualiza proporcionalmente o parâmetro $T_{\text{SDT},s}^{\text{curr}}$. Quando o *back-off* aumenta o intervalo de coleta, o T_{SDT} acompanha essa expansão seguindo a fórmula:

$$T_{\text{SDT},s}^{\text{curr}} = \max(T_{\text{SDT},s}^{\text{base}}, 2 \times IC_s^{\text{curr}}) \quad (4.1)$$

Essa estratégia mantém a coerência do modelo: se o sistema confia que o paciente está estável (aumentado IC), ele também relaxa a exigência de registros forçados no SDT. Ambos os parâmetros indicam o mesmo nível de confiança na estabilidade clínica. Nos casos de *reset* (leitura instável ou risco ALTO), tanto IC quanto T_{SDT} retornam aos valores base definidos pelo Algoritmo 1.

Por exemplo, considere um paciente com risco MÍNIMO monitorando SpO_2 . Inicialmente, $IC_{SpO_2}^{\text{curr}} = 15$ min e $T_{\text{SDT},SpO_2}^{\text{curr}} = 30$ min. Após três leituras consecutivas estáveis ($K = 3$), o primeiro *back-off* dobra ambos os valores para $IC = 30$ min e $T_{\text{SDT}} = 60$ min. Esse processo continua até atingir o teto de $IC^{\text{max}} = 6$ h, momento em que T_{SDT} atinge 12 h. Caso ocorra uma variação que exceda ϵ_{SpO_2} , ambos os parâmetros retornam imediatamente aos valores base (15 min e 30 min, respectivamente).

Contudo, essa estratégia de alongar o intervalo introduz um risco latente: um paciente pode sofrer uma deterioração aguda após um longo período de silêncio (ex.: 8 horas), e o sistema só detectaria o problema na próxima coleta agendada. Para mitigar essa vulnerabilidade crítica, propõe-se, no Algoritmo 3, uma vigilância contínua adicional, que funciona como uma verifi-

Algoritmo 2: Back-off Exponencial Condicionado ao Risco

Entrada: Leitura $\langle p, s, v, risco_p \rangle$
Saída : $IC_{s,p}^{\text{curr}}, T_{\text{SDT},s,p}^{\text{curr}}$

```

// Obtém valores base do Algoritmo 1 para o sinal s
1  $IC_s^{\text{base}} \leftarrow \text{PARAMS}[risco_p][IC_s]; T_{\text{SDT},s}^{\text{base}} \leftarrow \text{PARAMS}[risco_p][T_{\text{SDT},s}];$ 
// Etapa 0 - Só executa em MODERADO, MÍNIMO ou BAIXO
2 if  $risco_p = \text{ALTO}$  then
3    $stableCnt_{s,p} \leftarrow 0; IC_{s,p}^{\text{curr}} \leftarrow IC_s^{\text{base}}; T_{\text{SDT},s,p}^{\text{curr}} \leftarrow T_{\text{SDT},s}^{\text{base}};$  return  $IC_{s,p}^{\text{curr}}, T_{\text{SDT},s,p}^{\text{curr}};$ 
// Etapa 1 - Verifica estabilidade da leitura
4 if  $|v - v_{\text{ult}}| \leq \varepsilon_s$  then
5    $stableCnt_{s,p} + +;$ 
6 else
7    $stableCnt_{s,p} \leftarrow 0; IC_{s,p}^{\text{curr}} \leftarrow IC_s^{\text{base}}; T_{\text{SDT},s,p}^{\text{curr}} \leftarrow T_{\text{SDT},s}^{\text{base}};$  return  $IC_{s,p}^{\text{curr}}, T_{\text{SDT},s,p}^{\text{curr}};$ 
// Etapa 2 - Aplica back-off exponencial após K leituras
// estáveis
8 if  $stableCnt_{s,p} \geq K$  then
9   if  $risco_p = \text{MODERADO}$  then
10    |  $IC^{\text{max}} \leftarrow 30 \text{ min};$ 
11   else if  $risco_p = \text{BAIXO}$  then
12    |  $IC^{\text{max}} \leftarrow 2 \text{ h};$ 
13   else
14    |  $IC^{\text{max}} \leftarrow 6 \text{ h};$ 
15    $IC_{s,p}^{\text{curr}} \leftarrow \min(IC^{\text{max}}, 2 \times IC_{s,p}^{\text{curr}}); stableCnt_{s,p} \leftarrow 0;$ 
// Etapa 3 - Ajusta  $T_{\text{SDT}}$  proporcionalmente ao  $IC^{\text{curr}}$ 
16    $T_{\text{SDT},s,p}^{\text{curr}} \leftarrow \max(T_{\text{SDT},s}^{\text{base}}, 2 \times IC_{s,p}^{\text{curr}});$ 
17 return  $IC_{s,p}^{\text{curr}}, T_{\text{SDT},s,p}^{\text{curr}};$ 

```

cação de pulso (*keep-alive*), garantindo vigilância mínima contínua sem sacrificar a eficiência energética.

Em resumo, o Algoritmo 2 atua como um elástico que alonga o tempo entre as coletas completas. Enquanto o paciente permanece estável em risco MODERADO, BAIXO ou MÍNIMO, ele dobra o intervalo de coleta, respeitando um teto máximo. Se qualquer leitura ultrapassar ε_s ou se o escore NEWS2 subir, o intervalo volta instantaneamente ao valor-base. Essa lógica é supervisionada pelo Algoritmo 3, que garante que, mesmo durante os intervalos mais longos, uma verificação de segurança seja realizada.

4.4.3 Algoritmo 3 - Vigilância Contínua

Para mitigar o risco de deterioração não detectada durante os longos intervalos de *back-off*, o Algoritmo 3 implementa vigilância contínua. Ele opera em paralelo, executando uma

coleta leve em um intervalo de verificação fixo (IC_{verif}). O valor de $IC_{\text{verif}} = 2$ minutos foi escolhido por corresponder ao intervalo base de risco MODERADO definido no Algoritmo 1; essa escolha garante que, mesmo pacientes em *back-off* máximo (até 6 horas para risco MÍNIMO), seus sinais sejam verificados com frequência equivalente à de um paciente de risco moderado, estabelecendo um piso de segurança clínica. Este método é diretamente análogo ao mecanismo de TCP *keep-alives*, um recurso fundamental de protocolos de rede projetado para verificar a integridade de uma conexão ociosa. Assim como o TCP *keep-alives* envia pacotes de sonda para garantir que a outra extremidade ainda está ativa, a coleta leve atua como uma sonda fisiológica para garantir que o estado do paciente não tenha se alterado silenciosamente (?).

Algoritmo 3: Vigilância Contínua

Entrada: Leitura leve $\langle p, s, v \rangle$ executada a cada IC_{verif}
Saída : Disparo da ColetaCompletaImediata (p, s) se a verificação falhar

```

// Executa em paralelo, independentemente do risco
// Compara leitura leve com a última leitura válida
1 if  $|v - v_{\text{ult}}| > \varepsilon_s$  then
    // Variação anômala detectada, disparar ação
2   ColetaCompletaImediata  $(p, s)$ ;
```

Uma coleta leve é uma operação de baixo custo computacional: ela apenas mede o sinal vital e o compara à última leitura válida, sem acionar o fluxo completo de compressão e envio às camadas de *fog-cloud*. Se essa verificação rápida detectar uma variação que exceda a margem de erro ε_s , ela força a interrupção imediata do temporizador de *back-off* do Algoritmo 2, disparando uma coleta completa para uma reavaliação imediata do estado do paciente. Dessa forma, o sistema combina eficiência energética com vigilância contínua, garantindo uma resposta rápida a mudanças clínicas súbitas.

4.4.4 Resumo da interdependência dos algoritmos

Os três algoritmos apresentados formam um sistema de controle integrado, onde cada componente depende e complementa os demais: o **Algoritmo 1** atua como configurador inicial, traduzindo o escore NEWS2 em parâmetros operacionais (IC , ε , DC , T_{SDT}); o **Algoritmo 2** otimiza dinamicamente o intervalo de coleta quando há estabilidade prolongada, respeitando os limites definidos pelo Algoritmo 1; e o **Algoritmo 3** supervisiona continuamente o paciente, podendo interromper o *back-off* do Algoritmo 2 e forçar uma reavaliação pelo Algoritmo 1 caso detecte deterioração. Essa arquitetura em cascata garante que o sistema seja eficiente em condições estáveis, sem abrir mão da vigilância necessária para detectar eventos críticos.

A Tabela ?? resume os parâmetros que regem o comportamento dos algoritmos, distinguindo-os pela sua natureza (default ou dinâmica) e pela entidade que os define ou atualiza. Essa separação explicita uma arquitetura de controle em duas camadas. A primeira é a camada *Default*,

composta por valores imutáveis em tempo de execução que codificam diretrizes clínicas e políticas do sistema. As configurações definidas ou alteradas pelo sistema podem ser modificadas pelo administrador. A segunda é a camada Dinâmica: nela, o Escore NEWS2, calculado em tempo hábil pelo sistema de saúde, é traduzido pelos algoritmos em parâmetros operacionais adaptativos (IC_s^{base} , ε_s , DC_s , T_{SDT}) e, posteriormente, refinado pelo mecanismo de *back-off*, que ajusta IC_s^{curr} . Todos os pacientes iniciam com os mesmos valores e, após as coletas iniciais, são ajustados. Essas configurações garantem que o sistema permaneça ancorado em premissas robustas e seja responsivo às mudanças no estado clínico do paciente.

Tabela 7: Parâmetros do ViSPAC: natureza e origem

Parâmetro	Descrição	Natureza	Definido/alterado por
Escore NEWS2	Pontuação clínica aferida periodicamente	Entrada	Sistema automatizado de cálculo
Limites de Risco	Faixas do NEWS2 que definem MÍNIMO–ALTO	Default	Diretriz clínica
Regras de Config. (Alg. 1)	Mapeia risco $\rightarrow \{IC, \varepsilon, DC, T_{\text{SDT}}\}$ por sinal	Default	Sistema (pré-deploy)
K	Leituras estáveis necessárias antes do back-off	Default	Sistema
IC_{max}	Teto de IC após back-off (por nível de risco)	Default	Sistema (Alg. 2)
IC_{verif}	Intervalo da vigilância contínua (keep-alive)	Default	Sistema (Alg. 3)
Risco Clínico	Classe de risco derivada do NEWS2	Dinâmico	Alg. 1
IC_s^{base}	Intervalo base de coleta por sinal s	Dinâmico	Alg. 1
ε_s	Margem de erro tolerável por sinal s	Dinâmico	Alg. 1
DC_s	Desvio máx. de compressão por sinal s	Dinâmico	Alg. 1
$T_{\text{SDT},s}^{\text{base}}$	T_{SDT} base por sinal s ($= 2 \times IC_s^{\text{base}}$)	Dinâmico	Alg. 1
stableCnt _{s,p}	Contador de leituras estáveis consecutivas	Dinâmico	Alg. 2
IC_s^{curr}	Intervalo corrente após back-off	Dinâmico	Alg. 2
$T_{\text{SDT},s}^{\text{curr}}$	T_{SDT} corrente ($= \max(T_{\text{SDT},s}^{\text{base}}, 2 \times IC_s^{\text{curr}})$)	Dinâmico	Alg. 2

5 METODOLOGIA E EXPERIMENTO

Esta seção descreve a metodologia para avaliar o modelo ViSPAC. A estratégia experimental visa validar a adaptação dinâmica do sistema diante de variações clínicas do paciente, bem como mensurar a eficiência no uso de recursos da rede e no processamento. A abordagem fundamenta-se em simulação controlada na nuvem, utilizando dados reais de sinais vitais para garantir a relevância dos resultados.

Não foi identificado na literatura um *benchmark* padronizado para avaliação de modelos de compressão adaptativa de sinais vitais em arquiteturas *Edge-Fog-Cloud*. Diante dessa lacuna, a metodologia foi construída utilizando o **modelo VSAC** (?) como *baseline* de comparação (representado pelo Cenário 2: Compressão Estática). Complementarmente, adotaram-se métricas consolidadas na literatura sobre compressão de sinais biomédicos: Taxa de Compressão (T_C), Distorção por PRD (*Percentage Root-Mean-Square Difference*) e Latência do ciclo de *feedback* (T_{loop}), conforme detalhado na Seção ??.

5.1 Design do Experimento

O experimento foi estruturado em um fluxo de processamento distribuído em três camadas: *edge*, *fog* e *cloud*, reproduzindo uma infraestrutura de Cidades Inteligentes voltada à saúde. O fluxo operacional do experimento é dividido em quatro estágios principais:

1. **Ingestão e Simulação:** Dados provenientes de *datasets* consolidados são injetados nas unidades *edge*, nas quais cada instância simula um conjunto de pacientes com perfis de risco distintos (alto, moderado, baixo e mínimo).
2. **Processamento Adaptativo na Borda:** A camada *edge* realiza a compressão dos dados. Dependendo do risco identificado pelo ciclo de *feedback*, o modelo ajusta os parâmetros de coleta (IC, ε) e de compressão (DC, T_{SDT}).
3. **Análise e Feedback na Fog:** A camada *fog* atua como um nó intermediário responsável pela descompressão, pelo cálculo do escore NEWS2 e pela retransmissão imediata dos novos parâmetros operacionais para a *edge*, fechando o laço de controle.
4. **Persistência e Análise Histórica:** Os dados processados são encaminhados para a *cloud* por meio de uma fila por risco, onde são armazenados em um banco de dados para permitir a análise pelos agentes de saúde.

A arquitetura foi projetada para ser testada sob diferentes cargas, variando a composição de pacientes em cada nó de borda, a fim de avaliar o desempenho do modelo em cenários de alta variabilidade clínica.

5.2 Materiais e Dados

Para a validação experimental do protótipo, foram utilizados dados reais de sinais vitais, especificamente FC (já utilizado no VSAC) e SpO_2 . A seleção dos dados buscou representar cenários clínicos opostos, incluindo pacientes críticos e indivíduos saudáveis, permitindo avaliar a resposta do modelo ViSPAC em extremos de variabilidade fisiológica. Os dados foram obtidos a partir de duas fontes públicas distintas:

1. **BIDMC PPG and Respiration Dataset (??)¹**: Disponibilizado no PhysioNet, este conjunto contém dados fisiológicos coletados de pacientes internados em Unidades de Terapia Intensiva (UTI) no *Beth Israel Deaconess Medical Center* (Boston, EUA).
 - **Aplicação no Experimento:** Utilizado para simular os perfis de risco **ALTO** e **MODERADO**. Foram empregados os registros de 53 pacientes, em que cada série temporal possui mais de 400 amostras contínuas. A natureza instável destes sinais desafia a capacidade do modelo de manter a fidelidade dos dados sob compressão.
2. **Biosensor Student Health Fitness Data (?)**: Disponível na plataforma Kaggle, este conjunto de dados reúne dados biométricos de estudantes universitários saudáveis, caracterizados por alta estabilidade e baixo risco clínico.
 - **Aplicação no Experimento:** Utilizado para simular os perfis de risco **BAIXO** ou **MÍNIMO**.
 - **Pré-processamento e Agrupamento:** Originalmente, este *dataset* apresenta leituras pontuais (uma única aferição por indivíduo), o que inviabilizaria a simulação de um monitoramento contínuo ao longo do tempo. Para contornar essa limitação metodológica, foi aplicada uma técnica de concatenação de registros: os dados de diferentes estudantes foram agrupados de forma sequencial para sintetizar pacientes virtuais. Logo, os dados dos estudantes foram agrupados em 8 pacientes, gerando de 80 a 90 registros por paciente virtual, o que permitiu o experimento.

Como apenas FC e SpO_2 foram utilizados nos testes, os demais sinais vitais foram fixados e têm valores padrão. Essa composição heterogênea de dados permitiu testar o algoritmo de adaptação dinâmica diante de padrões de sinais vitais com entropias distintas, validando a generalização do modelo.

5.3 Ambiente de Avaliação

A validação experimental do modelo ViSPAC foi conduzida em um ambiente de nuvem totalmente virtualizado na plataforma da Amazon Web Services (AWS). A concepção da infraestrutura buscou simular um cenário realista de Cidades Inteligentes, em que a distribuição

¹<https://physionet.org/content/bidmc/1.0.0/>

geográfica dos recursos e as limitações de hardware são fatores críticos. Para tanto, adotou-se uma arquitetura multirregião, distribuindo as camadas de processamento entre diferentes centros de dados da AWS para introduzir latências de rede representativas de um ambiente de produção em larga escala.

A Figura ?? apresenta o diagrama detalhado da topologia de rede implementada, evidenciando a segmentação geográfica, a organização das sub-redes e os fluxos de comunicação entre os componentes.

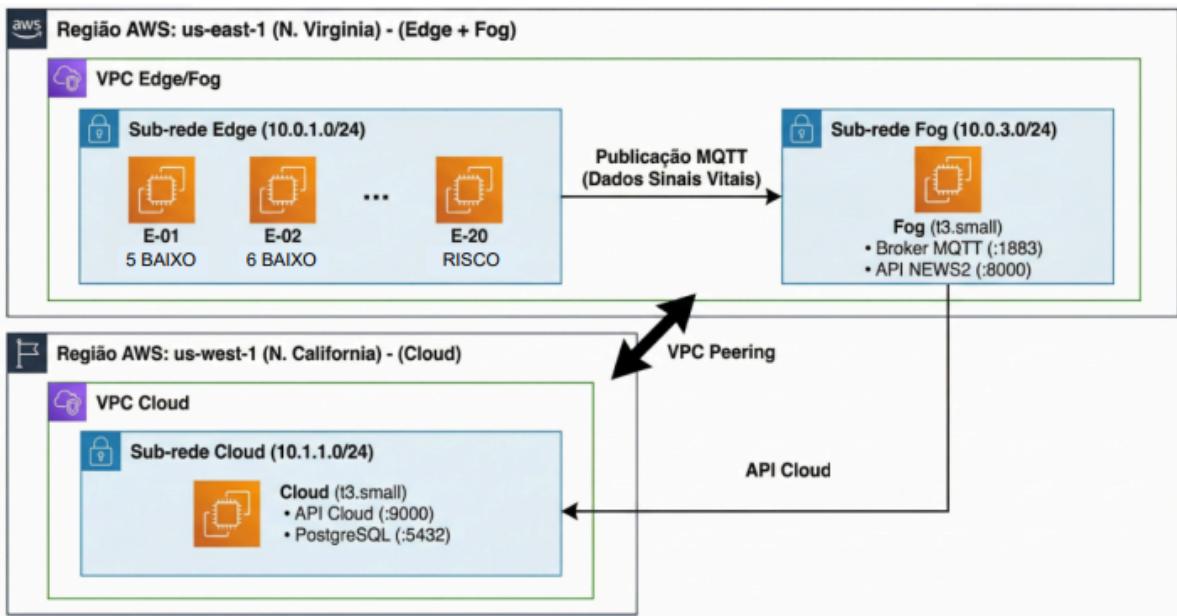


Figura 15: Diagrama da topologia da infraestrutura experimental distribuída na AWS. A arquitetura é segmentada em duas regiões geográficas distintas: us-east-1 (Norte da Virgínia), que hospeda as camadas *Edge* e *Fog*, e us-west-1 (Norte da Califórnia), que hospeda a camada *Cloud*. Uma conexão entre duas VPCs, chamada de emparelhamento de VPCs (*VPC Peering*), interliga as redes regionais. Na parte superior, observa-se a *VPC Edge/Fog* contendo a sub-rede *Edge* com 20 dispositivos simulados (ex.: E-01, E-02) e a sub-rede *Fog* com o nó concentrador. O fluxo de dados principal ocorre via MQTT da *Edge* para a *Fog*. Na parte inferior, a *VPC Cloud* contém uma API e um banco de dados PostgreSQL, recebendo dados processados da *Fog* por meio de uma API HTTP via peering.

Para garantir a reproduzibilidade do experimento e a gerência eficiente dos 22 nós computacionais envolvidos, toda a infraestrutura foi definida por meio de código (*Infrastructure as Code* - IaC) com a ferramenta Terraform. Os arquivos de configuração utilizados para provisionar este ambiente encontram-se disponíveis no repositório do projeto².

Cada dispositivo Edge (identificado por seu EDGE_ID) foi configurado para monitorar um conjunto fixo e reproduzível de pacientes, selecionados aleatoriamente a partir dos datasets *BIDMC PPG and Respiration Dataset* e *Biosensor Student Health Fitness Data*. A reproduzibilidade é garantida por meio de *seeding* determinístico baseado no EDGE_ID, permitindo

²<https://github.com/mateusrovedaa/dissertacao/tree/terraform>

a replicação exata dos experimentos. A variabilidade entre edges reflete cenários realistas de implantação distribuída em múltiplas instalações.

Abaixo são detalhados os componentes de cada região, conforme ilustrado na Figura ??:

5.3.1 Região us-east-1 (Edge e Fog)

Esta região concentra os componentes responsáveis pela coleta e pelo processamento inicial dos dados, simulando a infraestrutura local de uma cidade.

- **Camada Edge (Sub-rede Edge):** Composta por 20 instâncias EC2 de categoria base *t2.small* (1 vCPU, 1GB RAM). Um aspecto metodológico fundamental foi a imposição de restrições artificiais de hardware nestas instâncias para simular dispositivos IoT de baixo custo. Por meio de mecanismos do sistema operacional (Cgroups), a memória RAM foi limitada a **256 MB** e a utilização da CPU foi restringida a **50%** da capacidade de um núcleo. Conforme mostrado na figura, estas instâncias foram configuradas com diferentes densidades de pacientes (ex.: nó E-01 simulando 5 pacientes de baixo risco) para criar cargas de trabalho heterogêneas. A comunicação de saída para a camada *Fog* é realizada exclusivamente pelo protocolo MQTT.

A escolha específica destes limites fundamenta-se na equivalência com dispositivos *Single Board Computer* (SBC) amplamente utilizados, como o **Raspberry Pi Zero** (que possui 512 MB de SDRAM e um processador quad-core ARM Cortex-A53). Visto que instâncias de nuvem operam com processadores de alto desempenho (arquitetura x86_64), a redução de 50

- **Camada Fog (Sub-rede Fog):** Composta por uma instância *t3.small* (2 vCPU, 2GB RAM), atuando como *gateway* regional. Este nó executa um *broker* MQTT para receber os dados comprimidos da borda, realiza a descompressão, calcula o escore NEWS2 e retorna os parâmetros de feedback para a borda. Simultaneamente, ele atua como ponte para a nuvem, encaminhando dados processados para a região remota.

5.3.2 Região us-west-1 (Cloud) e Interconexão

Esta região simula um datacenter centralizado para armazenamento de longo prazo e análises intensivas.

- **Camada Cloud (Sub-rede Cloud):** Hospeda uma instância *t3.small* (2 vCPU, 2GB RAM) executando um banco de dados PostgreSQL para persistência histórica dos sinais vitais e uma API para recebimento de dados.
- **VPC Peering:** A comunicação entre a camada Fog (us-east-1) e a camada Cloud (us-west-1) é viabilizada por meio de uma conexão de emparelhamento de VPCs (*VPC Peering*)

ring). O tráfego flui pelo *backbone* da AWS, mas está sujeito às latências inter-regionais naturais, cruciais para a avaliação realista das métricas de tempo de resposta do sistema.

5.4 Implementação do Modelo ViSPAC

O software do sistema foi implementado em Python, estruturado de forma modular para operar de forma distribuída. A lógica de processamento divide-se entre a execução leve na borda (centrada na redução de dados) e a execução analítica na camada de *fog* (centrada na tomada de decisão clínica).

5.4.1 Pipeline de Compressão na Borda

Cada instância de borda executa um ciclo contínuo de coleta e processamento. O fluxo de dados foi desenhado para priorizar a latência em situações críticas e a eficiência de banda em situações estáveis. O processamento ocorre em dois estágios:

1. **Compressão Primária (Lossy):** Todos os sinais vitais brutos (FC e SpO_2) são submetidos ao algoritmo *Swinging Door Trending* (SDT). Este algoritmo foi selecionado por sua complexidade computacional linear ($O(n)$), o que o torna ideal para o hardware restrito simulado. A agressividade da compressão é determinada pelo parâmetro *DC* (Desvio de Compressão), ajustado dinamicamente.
2. **Compressão Secundária (Lossless):** A saída do SDT pode passar por um segundo estágio de compressão sem perdas (Huffman ou LZW) para maximizar a redução do tamanho do pacote. No entanto, implementou-se uma regra de **bypass para risco crítico**: caso o paciente esteja classificado em risco ALTO, o estágio secundário é ignorado e os dados são transmitidos imediatamente após o SDT. Esta estratégia visa eliminar o *overhead* computacional da segunda compressão, reduzindo a latência ponta-a-ponta em detrimento de uma pequena economia de banda.

Para a implementação dos algoritmos, foram adaptadas bibliotecas de código aberto disponíveis na comunidade, otimizando-as para o fluxo de dados do experimento:

- **SDT:** Baseado na implementação de *Swinging Door Trending*, disponível em repositório público³, adaptada para suportar janelas de tempo variáveis (T_{SDT}).
- **LZW e Huffman:** Utilizados como algoritmos de entropia para a compressão secundária, com base nas implementações de Gupta⁴ e Bhrigu⁵.

³<https://github.com/chelaxe/SwingingDoor>

⁴<https://github.com/adityagupta3006/LZW-Compressor-in-Python>

⁵<https://github.com/bhrigu123/huffman-coding>

5.4.2 Serialização Binária

O ViSPAC adota o formato binário *MessagePack*⁶ para a serialização dos lotes transmitidos da *edge* para a *fog*, em substituição ao JSON textual utilizado no VSAC. A escolha fundamenta-se na avaliação de (?), que demonstrou que o *MessagePack* apresenta desempenho de serialização superior ao JSON em microcontroladores embarcados de 8 a 32 bits, além de tamanho de *payload* significativamente reduzido. O *MessagePack* produz representações equivalentes ao JSON, porém em formato binário compacto: números inteiros e de ponto flutuante são armazenados em sua forma nativa (1–9 bytes), enquanto no JSON cada dígito consome um byte de texto. Além disso, estruturas como mapas e arrays são delimitadas por cabeçalhos de poucos bytes, eliminando os caracteres { }, [], : e , exigidos pela notação textual. Essa representação resulta em uma redução no tamanho do *payload* antes da aplicação dos algoritmos *lossless*, sem custo computacional relevante. Adicionalmente, os algoritmos de Huffman e LZW foram adaptados para operar diretamente sobre sequências de bytes, dispensando a codificação Base64 intermediária necessária quando a entrada era textual.

5.4.3 Mecanismo de Adaptação e Feedback (Fog)

A inteligência do sistema reside na camada *Fog*. Ao receber um lote de dados via MQTT, o nó *fog* realiza a descompressão, reconstrói os sinais vitais e calcula o escore NEWS2. Esse escore atua como gatilho para o ciclo de *feedback*: o módulo Configurador, executado na *edge*, recebe a classificação do NEWS2 e aplica os novos valores conforme o Algoritmo ???. Uma piora no estado clínico provoca uma reconfiguração imediata para alta fidelidade, enquanto a estabilização permite o relaxamento dos parâmetros para economizar recursos.

O encaminhamento dos dados processados para a *cloud* é configurável em dois modos. No modo **síncrono**, o fog encaminha os dados à *cloud* via HTTP e aguarda a confirmação antes de responder ao *edge*, de modo que a latência percebida pelo *edge* inclui o tempo de ida e volta até a nuvem. No modo **assíncrono**, o fog responde ao *edge* imediatamente após o processamento e enfileira os dados para encaminhamento posterior. As filas são organizadas por nível de risco clínico, com um *worker* dedicado que as drena, respeitando a ordenação HIGH > MODERATE > LOW > MINIMAL, garantindo que os dados de pacientes críticos cheguem antes à camada de armazenamento. Nesse modo, o tempo de resposta da nuvem não afeta a responsividade do ciclo de *feedback*.

5.5 Métricas de Avaliação

A eficácia do modelo ViSPAC é avaliada sob três perspectivas fundamentais: eficiência de rede, latência e fidelidade clínica. A escolha destas métricas fundamenta-se em uma revisão

⁶<https://msgpack.org/>

sistemática dos trabalhos relacionados, conforme resumido na Tabela ??, garantindo que a avaliação seja comparável ao estado da arte.

Tabela 8: Métricas de avaliação recorrentes na literatura sobre monitoramento remoto de saúde e compressão de dados.

Trabalho Relacionado	Taxa de Compressão	Latência / Tempo	Distorção / Qualidade
Zhou et al. (?)	Sim	-	PSNR, RMSE, SSIM
Esmaeili et al. (?)	Sim	Sim	-
Andrade et al. (?)	Sim	-	PRD
Hassan et al. (?)	Sim	-	PSNR, PRD, MSE
Shankani et al. (?)	-	Sim	-
Chang et al. (?)	Sim	-	PRD, MSE

Fonte: Elaborado pelo autor.

Com base nesta análise, definem-se as seguintes métricas para o experimento:

1. **Taxa de Compressão (T_C):** Quantifica a redução do volume de dados trafegados na rede. É definida pela razão entre o tamanho do arquivo comprimido final (S_f) e o tamanho dos dados brutos originais (S_i):

$$T_C = \left(\frac{S_f}{S_i} \right) \times 100\%$$

Quanto menor o valor de T_C , maior a eficiência do sistema em economizar largura de banda.

2. **Latência do Ciclo de Feedback (T_{loop}):** Métrica crítica para sistemas adaptativos mede o tempo total de resposta do sistema a uma mudança clínica. É calculada pela diferença entre o instante de coleta da amostra na borda (t_{coleta}) e o instante em que os novos parâmetros de configuração, gerados após o processamento na *fog*, são efetivamente aplicados no dispositivo (t_{ajuste}):

$$T_{loop} = t_{ajuste} - t_{coleta}$$

Um T_{loop} baixo indica alta responsividade, essencial para pacientes em deterioração rápida.

3. **Distorção do Sinal (PRD):** Avalia a perda de qualidade decorrente da compressão *lossy* (SDT). Utiliza-se a Diferença Percentual da Raiz Quadrada Média (*Percentage Root-Mean-Square Difference*), que compara o sinal original (X) com o sinal reconstruído (Y) ao longo de N amostras:

$$PRD = \sqrt{\frac{\sum_{i=1}^N (X_i - Y_i)^2}{\sum_{i=1}^N X_i^2}} \times 100$$

Valores de PRD próximos de zero indicam alta fidelidade. Para os sinais clínicos, busca-se um equilíbrio em que o PRD seja minimizado especificamente nos trechos de anomalia clínica.

5.6 Cenários de Teste

Para isolar os benefícios de cada componente da arquitetura, o experimento foi dividido em três cenários distintos. Cada cenário foi executado por um período contínuo de **12 horas**. Esta janela temporal (que é duas vezes o limite definido para que o risco mínimo seja coletado, caso o *back-off* atue) foi selecionada para capturar a variabilidade natural dos sinais vitais e o comportamento do tráfego de rede em diferentes estados de carga. As configurações de cada cenário foram controladas por meio de variáveis de ambiente no Terraform. Todos os cenários utilizaram a arquitetura proposta e o mesmo código, porém com desvios conforme o cenário (como a desativação de módulos de compressão e a priorização).

5.6.1 Cenário 1: Linha de Base (Baseline)

Configurado en Terraform como `scenario1_baseline`. Este cenário representa a abordagem tradicional de monitoramento contínuo sem inteligência na borda.

- **Configuração:** Módulos de compressão desativados. Encaminhamento fog–cloud em modo síncrono.
- **Dinâmica:** Os dados são coletados e enviados em sua forma bruta (*raw data*) com intervalo de coleta fixo ($IC = 1$ s). Este valor foi escolhido para simular o pior cenário de consumo de banda, em que cada leitura é transmitida sem qualquer otimização.
- **Objetivo:** Estabelecer o pior caso de consumo de banda e de armazenamento, servindo como referência para o cálculo do ganho de eficiência (T_C) e para a validação da distorção zero ($PRD = 0$) dos dados originais.

5.6.2 Cenário 2: Compressão Estática

Configurado en Terraform como `scenario2_static`. Avalia o ganho obtido puramente pela introdução de algoritmos de compressão, sem sensibilidade ao contexto clínico. Este cenário se equipara ao VSAC.

- **Configuração:** Algoritmos SDT (Lossy) e Huffman/LZW (Lossless) ativados. Encaminhamento fog–cloud em modo síncrono.
- **Dinâmica:** Os parâmetros de compressão (DC , T_{SDT}) e de coleta (IC) são estáticos e prefixados. Não há ciclo de feedback ativo; o sistema ignora as flutuações do escore NEWS2.
- **Objetivo:** Quantificar a redução de dados que pode ser obtida apenas por meio da compressão. Este cenário permite diferenciar o ganho do algoritmo (matemático) do ganho da estratégia adaptativa (arquitetural).

5.6.3 Cenário 3: ViSPAC (Adaptativo Completo)

Configurado no Terraform como `scenario3_vispac`. Representa a solução proposta em sua totalidade.

- **Configuração:** Todos os módulos ativados, incluindo o ciclo de feedback *Edge-Fog-Cloud* e o encaminhamento assíncrono com filas de prioridade por risco na camada *fog*.
- **Dinâmica:** O sistema ajusta dinamicamente IC , ε , DC e T_{SDT} em resposta ao cálculo do NEWS2 na camada *fog*. Pacientes em estado crítico acionam o modo de alta fidelidade (menor compressão, transmissão imediata), enquanto pacientes estáveis operam em modo de alta eficiência. O *fog* responde ao *edge* imediatamente após o processamento, sem aguardar a confirmação da *cloud*, reduzindo a latência do ciclo de feedback.
- **Objetivo:** Demonstrar a capacidade do ViSPAC de gerenciar o *trade-off* entre latência e eficiência. Espera-se que este cenário apresente uma T_C competitiva em relação ao Cenário 2, porém com uma PRD substancialmente menor nos momentos críticos e uma latência (T_{loop}) otimizada para eventos de risco.

6 RESULTADOS E DISCUSSÃO

Este capítulo apresenta a avaliação experimental da arquitetura *Edge-Fog-Cloud* proposta. O objetivo central é validar a eficácia do modelo ViSPAC na otimização do monitoramento remoto de sinais vitais, quantificando ganhos de eficiência de rede, de armazenamento e de manutenção da QoS e da fidelidade clínica. A discussão subsequente analisa as dinâmicas dos algoritmos de compressão adaptativa e de priorização baseada em risco, respondendo às questões de pesquisa levantadas no início deste trabalho. Em suma, a Figura ?? apresenta um resumo dos testes executados e dos que serão debatidos nas próximas seções.

Resumo Comparativo

Métrica	Cenário 1	Cenário 2	Cenário 3
Transmissões	6,726,598	567,108	214,505
Dados Brutos (MB)	353.96	366.61	358.16
Dados Comprimidos (MB)	353.96	95.12	65.41
Compressão Total (%)	0.0%	74.1%	81.7%
PRD Média (%)	0.0000%	3.8911%	1.2114%
↳ PRD Risco Alto (%)	0.0000% (n=0)	0.4330% (n=216,060)	0.6053% (n=194,274)
↳ PRD Risco Moderado (%)	0.0000% (n=0)	1.1197% (n=44,859)	1.2545% (n=6,046)
↳ PRD Risco Baixo (%)	0.0000% (n=0)	6.4011% (n=92,293)	8.4899% (n=4,205)
↳ PRD Risco Mínimo (%)	0.0000% (n=0)	6.8824% (n=213,896)	9.9169% (n=9,980)
Latência Média (ms)	1380.4 ± 245.5	1271.2 ± 140.8	1182.5 ± 78.0
↳ Latência Risco Alto (ms)	1313.2 (n=36,863)	1228.9 (n=20,858)	1172.3 (n=28,930)
↳ Latência Risco Moderado (ms)	1538.6 (n=4,292)	1372.8 (n=4,849)	1208.8 (n=3,269)
↳ Latência Risco Baixo (ms)	1558.1 (n=2,817)	1317.0 (n=2,907)	1229.5 (n=1,320)
↳ Latência Risco Mínimo (ms)	1809.4 (n=3,022)	1423.2 (n=1,686)	1284.9 (n=1,434)
CPU Média (%)	0.2%	0.2%	0.2%
Memória Média (MB)	89.0	87.5	87.9

* n = número de amostras coletadas para cálculo da métrica

Figura 16: Resumo comparativo das execuções realizadas nos 3 cenários de testes. Cada coluna apresenta os resultados detalhados, exibindo a latência e a distorção para cada nível de risco. Nos cenários 1 e 2, a latência (*tloop*) foi apenas medida, sem adaptação com novos parâmetros.
Fonte: Elaborado pelo autor.

Os dados detalhados e os gráficos interativos desta análise podem ser consultados no dashboard online do projeto: <https://roveda.dev/dissertacao/results/dashboard.html>.

6.1 Visão Geral e Comparativo Qualitativo

Os experimentos foram conduzidos em um ambiente de simulação distribuída. Foram utilizados 20 nós de borda e uma população de pacientes com perfis de risco heterogêneos, conforme detalhado na Seção ???. Para isolar as contribuições de cada componente da arquitetura, três cenários distintos foram avaliados e estão descritos na Tabela ???.

O **Cenário 1 (Linha de Base)** estabelece o pior caso de consumo, simulando dispositivos IoT burros que transmitem dados brutos continuamente. O **Cenário 2 (Estático)** representa o

Tabela 9: Comparativo qualitativo dos cenários experimentais

Característica	Cenário 1: Linha de Base	Cenário 2: Estático (VSAC)	Cenário 3: (Proposto)	ViSPAC
Abordagem	Transmissão Bruta	Compressão Estática	Está-	Compressão Adaptativa ao Risco
Coleta	Contínua (1s)	Fixa (15s)		Dinâmica (Algoritmo Back-off)
Compressão	Nenhuma	SDT + man/LZW	Huff-	SDT Adaptativo + Huffman/LZW
Priorização	Nenhuma	Nenhuma		Filas de Prioridade por Risco
Feedback	Unidirecional	Unidirecional		Bidirecional (Fog → Edge)
Objetivo	Integridade total	Redução de dados		Eficiência e Contexto Clínico

Fonte: Elaborado pelo autor.

estado da arte anterior (modelo VSAC), com compressão na borda, mas sem sensibilidade ao contexto. O Cenário 3 (ViSPAC) concretiza a proposta deste trabalho ao integrar a priorização baseada no risco clínico (NEWS2), permitindo a adaptação dinâmica dos parâmetros de coleta e de transmissão.

A Tabela ?? consolida as métricas de desempenho obtidas ao longo de 12 horas de simulação contínua, demonstrando a evolução da eficiência do modelo proposto.

Tabela 10: Métricas de desempenho globais da simulação (20 nós edge, 12 horas)

Métrica	Linha de Base	Estático	ViSPAC
Transmissões Totais (pacotes)	6.642.048	564.237	220.403
Volume de Dados Brutos (MB)	245,11	246,96	245,01
Volume Trafegado Final (MB)	245,11	61,65	45,21
Taxa de Compressão Média (%)	0,00	75,0 ± 10,6	81,6 ± 9,7
Distorção – PRD Médio Global (%)	0,00	3,62 ± 4,94	1,16 ± 2,67
Latência Média (ms)	1.358,3 ± 249,5	1.271,2 ± 149,0	1.048,9 ± 3,8

Fonte: Resultados experimentais obtidos da simulação. Valores de \pm representam o desvio padrão entre os 20 nós edge ($\sigma, n = 20$).

É importante notar que o **Volume de Dados Brutos** é praticamente idêntico entre os três cenários (≈ 245 MB), pois representa os mesmos dados fisiológicos serializados em *MessagePack* antes de qualquer compressão. A pequena variação do Cenário Estático (+0,75%) decorre do *overhead* de metadados de lote (nível de risco, escore NEWS2, flags de compressão), que não é totalmente compensado pela amortização em lotes maiores nos *edges* de alto risco. Contudo, o indicador real de eficiência é o **Volume Trafegado Final**, no qual o ViSPAC demonstra sua superioridade.

6.2 Análise de Redução de Transmissões

A métrica de maior impacto sistêmico observada foi a drástica redução no número de interações com a rede. Comparado à Linha de Base, o ViSPAC reduziu em 96,7% a frequência de envios, caindo de mais de 6,6 milhões de transmissões para apenas 220.403 (Figura ??).

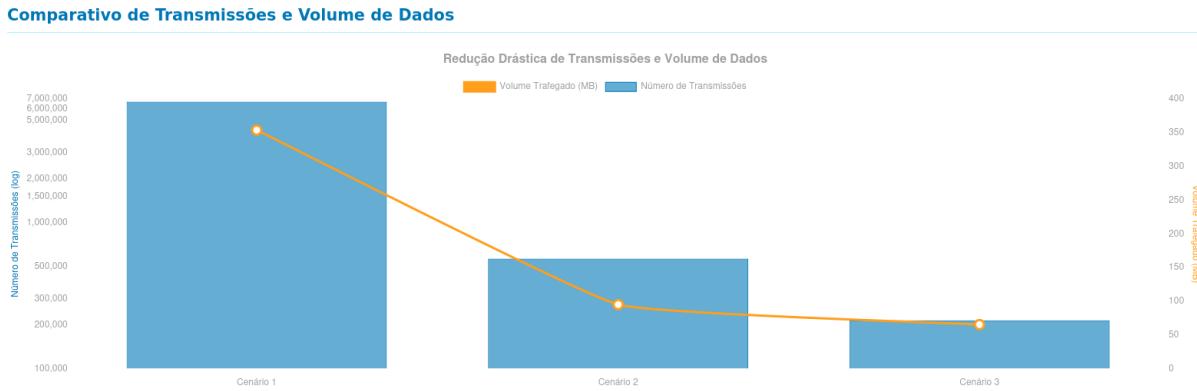


Figura 17: Gráfico comparativo do número total de transmissões (eixo primário, barras) e do volume de dados trafegados (eixo secundário, linha). A redução dos pacotes no ViSPAC impacta diretamente a longevidade da bateria dos dispositivos de borda.

Fonte: Elaborado pelo autor.

Essa supressão massiva de tráfego é atribuída à atuação do **Algoritmo de Back-off (Algoritmo 2)** e à adaptação ao risco do paciente. Ao identificar períodos de estabilidade clínica e desacoplar a coleta da transmissão, o ViSPAC permite que o dispositivo de borda permaneça em modos de baixo consumo por períodos prolongados, despertando apenas para verificações locais (Algoritmo 3) ou quando o intervalo de envio é atingido. Diferentemente do Cenário Estático, que opera com janela fixa, o ViSPAC compra tempo de silêncio na rede, custeado pela estabilidade do paciente.

6.3 Eficiência de Compressão e Volume de Dados

O modelo ViSPAC alcançou uma **Taxa de Compressão Média de 81,6%**, superando a de 75,0% do Cenário Estático. Como consequência, o volume final trafegado foi de apenas 45,21 MB, em comparação aos 245,11 MB originais. A adoção do *MessagePack* como formato de serialização (Seção ??) contribuiu com uma redução de aproximadamente 30% no volume bruto dos *payloads* antes mesmo da compressão *lossless*, pois a representação binária de números e estruturas é significativamente mais compacta do que a notação textual JSON. A Figura ?? apresenta esses dados de forma resumida.

Este ganho em relação ao modelo Estático (VSAC) não decorre de um novo algoritmo matemático de compressão, mas sim da estratégia arquitetural. Ao espaçar as coletas de pacientes estáveis, o ViSPAC gera séries temporais mais redundantes e lineares, um cenário ideal para o algoritmo SDT. Além disso, ao evitar pacotes de alta frequência para pacientes sem alterações,

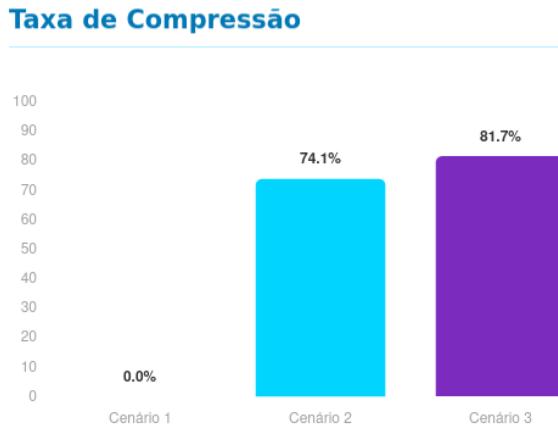


Figura 18: Comparação da compressão média entre os cenários. Como o cenário 1 não possui compressão, ela está zerada. Os cenários 2 e 3 apresentam uma compressão muito próxima, com o cenário 3 tendo vantagem.

Fonte: Elaborado pelo autor.

reduz-se o overhead de cabeçalhos de rede (TCP/IP e MQTT), o que corresponderia a uma parte significativa do volume de transmissões pequenas e frequentes.

6.4 Fidelidade Clínica e Distorção (PRD)

Um dos resultados mais significativos deste estudo reside na relação entre a compressão e a qualidade clínica. O ViSPAC obteve maior compressão geral mantendo uma distorção (PRD) média global de apenas **1,16%**, consideravelmente menor que os **3,62%** do Cenário Estático (Figura ??). Esse resultado, à primeira vista contraintuitivo (maior compressão com menor erro), explica-se pela **seletividade clínica** do modelo. No Cenário Estático, uma compressão média é aplicada a todos os pacientes. No ViSPAC:

- **Pacientes Estáveis:** Sofrem compressão agressiva e *back-off*. O PRD é elevado localmente, mas clinicamente irrelevante, pois o sinal é constante.
- **Pacientes Críticos:** O sistema desativa a compressão agressiva e o *back-off*. O sinal é transmitido com fidelidade quase total.

A Tabela ?? decompõe o PRD por nível de risco clínico e revela um padrão importante: quando analisados individualmente, os valores de PRD do ViSPAC são **superiores** aos do Cenário Estático em todas as faixas de risco. Isso ocorre porque os limiares SDT adaptativos do ViSPAC aplicam uma compressão *lossy* mais agressiva do que o limiar fixo do modelo estático.

Entretanto, a coluna de amostras (*n*) evidencia o mecanismo que explica a média global inferior. No Cenário Estático, o intervalo de coleta é fixo; por isso, pacientes de risco baixo e mínimo acumulam mais de 316 mil amostras com PRD elevado (5,59%–6,21%), dominando a média ponderada. No ViSPAC, o *backoff* exponencial espaça progressivamente as coletas desses mesmos pacientes estáveis, reduzindo suas amostras para cerca de 14 mil. Assim, embora

Distorção do Sinal (PRD)

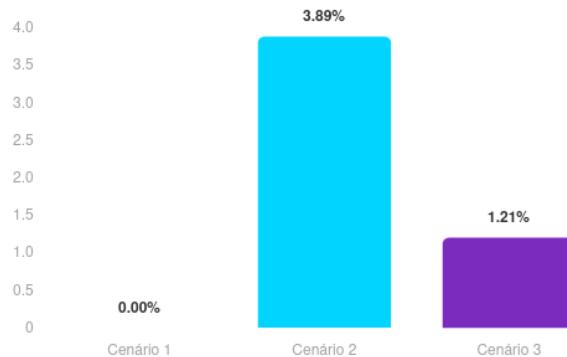


Figura 19: Comparação da Distorção PRD média. O ViSPAC apresenta menor distorção global, pois preserva a integridade quase total dos dados em momentos de crise, concentrando a perda de dados apenas em momentos de irrelevância clínica.

Fonte: Elaborado pelo autor.

Tabela 11: Distorção PRD por nível de risco clínico

Nível de Risco	Estático		ViSPAC	
	PRD (%)	Amostras	PRD (%)	Amostras
Alto	0,50	207.742	0,56	200.503
Moderado	0,70	40.396	1,27	5.429
Baixo	5,59	85.310	8,25	5.401
Mínimo	6,21	230.789	9,86	9.070
Média Global	3,62	564.237	1,16	220.403

Fonte: Elaborado pelo autor.

cada amostra individual apresente PRD ligeiramente mais alto, essas amostras representam uma fração muito menor do total, e a média global cai para 1,16%. Em resumo: o ViSPAC comprime mais agressivamente quando comprime, porém comprime **muito menos vezes**, e o saldo líquido é uma distorção média global substancialmente inferior.

O elevado desvio padrão do PRD entre nós ($\sigma = 2,67$) reflete diretamente essa adaptação: nós com pacientes exclusivamente de risco mínimo apresentam PRD próximo de 9,86%, enquanto nós com pacientes de alto risco mantêm PRD de apenas 0,56%. Os valores observados enquadram-se dentro dos limiares reportados na literatura para compressão de sinais biomédicos: PRD inferiores a 5% são considerados adequados para visualização e inferiores a 2% para diagnóstico (?). Em todos os cenários de risco alto ou moderado, onde a informação clínica é mais crítica, o PRD do ViSPAC permanece abaixo de 1,27%, bem dentro da faixa diagnóstica. Contudo, uma validação formal exigiria a avaliação por profissionais de saúde, com base nos sinais reconstruídos, para decisões clínicas, o que representa uma oportunidade para trabalhos futuros.

6.5 Análise de Latência e Decomposição por Camada

A latência percebida pelo *edge* (T_{loop}) é uma métrica composta pela soma de três componentes: o tempo de rede (transmissão e recepção), o processamento do NEWS2 na *fog* e, nos cenários de encaminhamento síncrono, o tempo de envio e de confirmação da *cloud*. Para compreender as diferenças entre os cenários, a Tabela ?? decompõe a latência média com base nos logs instrumentados do fog.

Tabela 12: Decomposição da latência média por componente e cenário

Componente	Linha de Base	Estático	ViSPAC
Processamento NEWS2 (<i>fog</i>)	2,31 ms	1,65 ms	0,88 ms
Encaminhamento fog–cloud	230,06 ms	207,73 ms	assíncrono
Rede + overhead edge	1.125,9 ms	1.061,8 ms	1.048,0 ms
Latência total percebida pelo edge	1.358,3 ms	1.271,2 ms	1.048,9 ms

Fonte: Calculado a partir dos logs do fog e das métricas dos edges.

A decomposição evidencia três fatores que contribuem para a redução de latência no ViSPAC. O primeiro e mais expressivo é o **desacoplamento do encaminhamento para a cloud**. Nos Cenários 1 e 2, o fog aguarda a confirmação da *cloud* antes de responder ao *edge*, o que adiciona 230 ms e 208 ms ao caminho crítico, respectivamente. No ViSPAC, esse encaminhamento ocorre de forma assíncrona via filas de prioridade, eliminando essa parcela da latência percebida pelo *edge*. O segundo fator é a **redução do volume dos pacotes transmitidos**. A compressão adaptativa, combinada com a serialização binária *MessagePack*, resulta em pacotes menores, que demandam menos tempo de serialização e de transmissão na rede, contribuindo para a redução da componente de rede de 1.126 ms (Linha de Base) para 1.048 ms (ViSPAC). O terceiro fator, de menor magnitude, é a **eficiência do processamento no fog**: com lotes menores e menos frequentes, o tempo de processamento do NEWS2 cai de 2,31 ms para 0,88 ms.

Em termos práticos, dos 309,4 ms de redução entre a Linha de Base e o ViSPAC, aproximadamente 230 ms decorrem da decisão de arquitetura de encaminhamento assíncrono, e os 79 ms restantes resultam da combinação de compressão, serialização binária e menor volume de dados na rede. No Cenário 3, a componente de rede responde por 1.048,0 dos 1.048,9 ms de latência total (99,9%), indicando que o processamento do fog é negligenciável em relação ao tempo de comunicação. Esse resultado reforça a premissa de que futuras otimizações devem concentrar-se na inteligência distribuída na borda, conforme discutido na Seção ??.

6.6 Resumo por Cenário

A Tabela ?? consolida as características e os resultados de cada cenário, permitindo uma comparação direta dos *trade-offs* envolvidos.

O **Cenário 1 (Linha de Base)** representa o pior cenário operacional, com o máximo consumo de banda e sem perda de dados. O **Cenário 2 (Estático/VSAC)** demonstra os ganhos

Tabela 13: Resumo consolidado por cenário de teste

Característica	Linha de Base	Estático (VSAC)	ViSPAC
Transmissões	6,6M	564K	220K
Taxa de Compressão	0%	75,0%	81,6%
PRD Médio Global	0%	3,62%	1,16%
Latência Média	1.358 ms	1.271 ms	1.049 ms
Adaptação ao Risco	Não	Não	Sim
Vigilância Contínua	Não	Não	Sim

Fonte: Resultados experimentais.

puramente algorítmicos da compressão, reduzindo substancialmente o volume sem considerar o contexto clínico. O **Cenário 3 (ViSPAC)** mostra que a integração da priorização clínica não apenas mantém os ganhos de compressão, mas os amplia enquanto simultaneamente reduz a distorção média.

6.7 Comportamento Adaptativo: Estudos de Caso

Para validar a hipótese de que o sistema reaja corretamente a diferentes perfis clínicos, analisamos o comportamento temporal de três pacientes distintos durante a simulação.

6.7.1 Paciente de Risco Alto (Paciente 09)

O Paciente 09 (nó `edge-08`) apresentou instabilidade contínua ($NEWS2 \geq 7$). Conforme ilustrado na Figura ??, o ViSPAC identificou um risco elevado e manteve o sistema em estado de alerta. O mecanismo de *back-off* foi desativado, e as transmissões ocorreram com frequência máxima e compressão mínima. O PRD resultante para este nó foi de 0,42%, garantindo que a equipe médica remota visualizasse cada nuance das oscilações cardíaca e respiratória.

6.7.2 Paciente de Risco Estável (Paciente 08)

Em contraste, o Paciente 08 (nó `edge-06`) manteve sinais vitais dentro da normalidade ($NEWS2 = 0$), conforme demonstrado na Figura ???. A Figura ??, por sua vez, demonstra a atuação do Algoritmo 2. Observa-se uma escada nos intervalos de coleta: após confirmar a estabilidade, o sistema dobrou sucessivamente o tempo entre envios até atingir o teto configurado. O nó transmitiu apenas uma fração dos pacotes em comparação ao Paciente 09. Esta economia inteligente libera largura de banda da rede para priorizar o tráfego de pacientes em estado crítico.

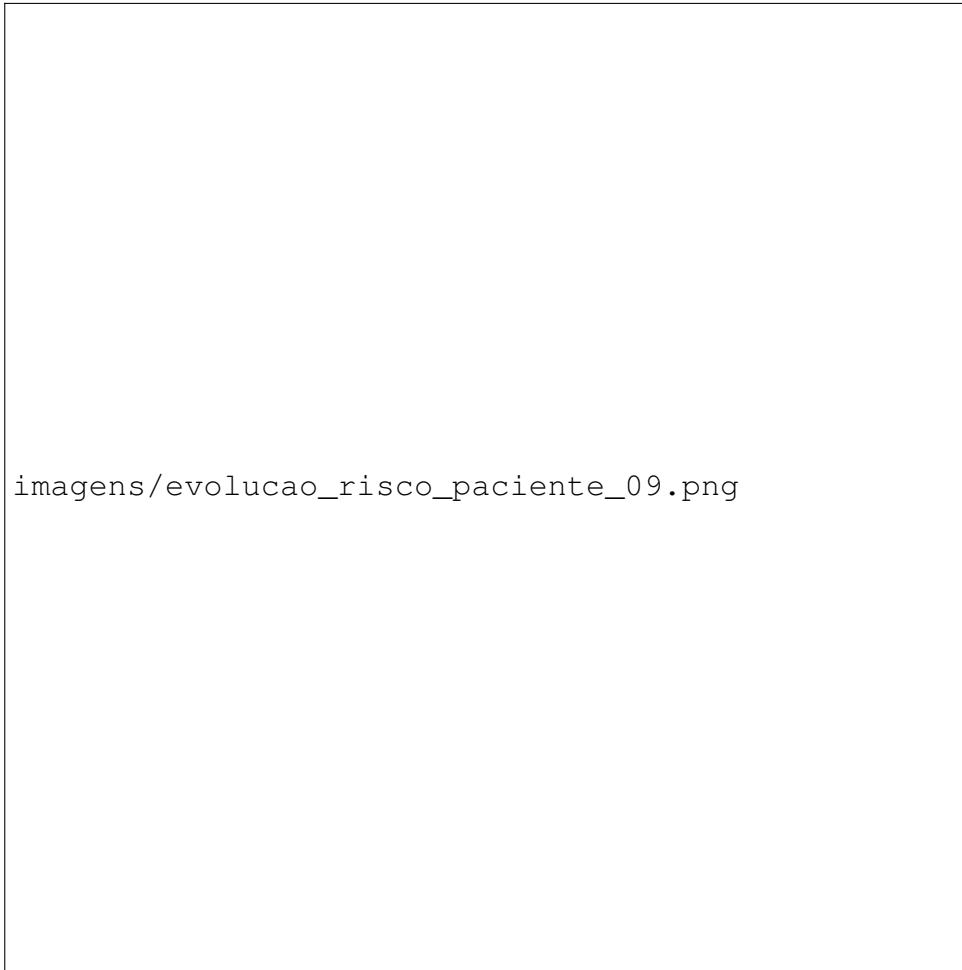


Figura 20: Evolução do risco do Paciente 09. O sistema manteve-se em modo de alta fidelidade durante todo o período.

Fonte: Elaborado pelo autor.

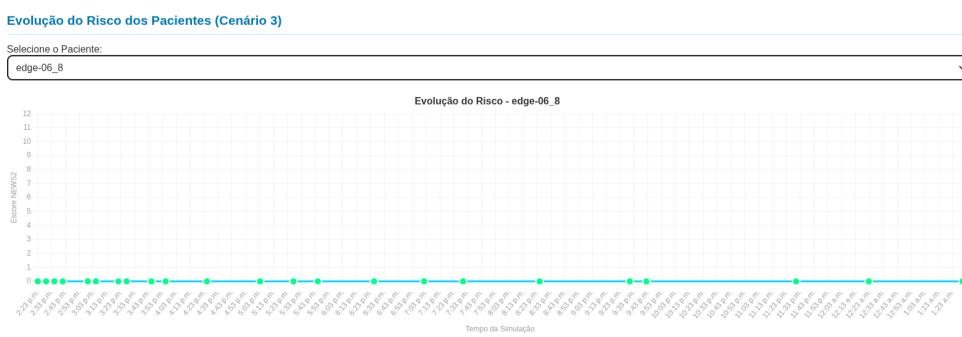


Figura 21: Nível de risco do Paciente 08. Podemos notar que o risco se manteve estável.

Fonte: Elaborado pelo autor.

6.7.3 Paciente de Risco Variável (Paciente 37)

Um dos cenários mais desafiadores e representativos da proposta do ViSPAC pode ser observado no Paciente 37 (*edge-07*, que alternou entre estados de estabilidade e deterioração rápida. Inicialmente, o paciente apresentava risco BAIXO. Porém, logo após o início da simulação, o

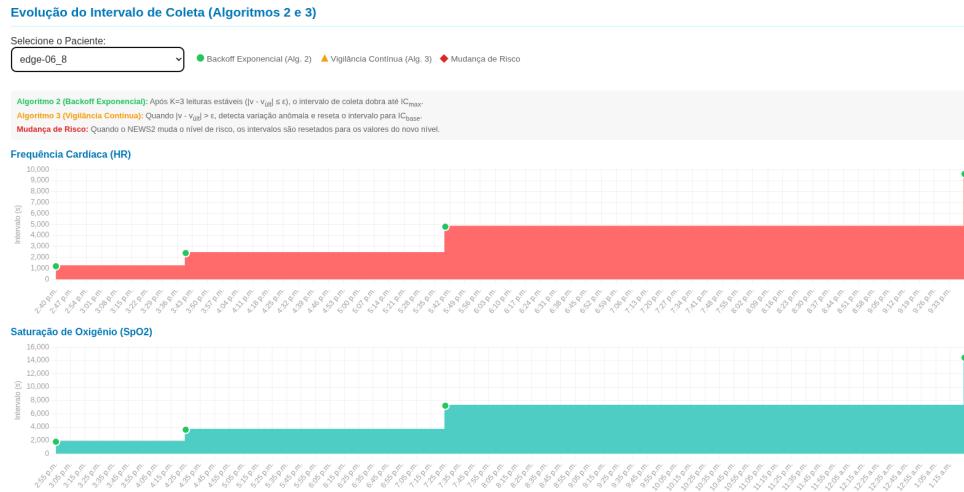


Figura 22: Evolução do intervalo de coleta do Paciente 08. Os pontos verdes indicam os momentos em que o algoritmo de back-off aumentou o intervalo devido à estabilidade prolongada.

Fonte: Elaborado pelo autor.

paciente teve seu risco alterado para MODERADO, o que levou o ViSPAC a ajustar o tempo de coleta e os parâmetros de compressão. Ao longo da jornada, o paciente oscilou entre riscos moderados e baixos, atingindo, inclusive, o risco ALTO em determinado momento. Todos os algoritmos foram acionados ao longo da simulação:

- 1. Detecção (Algoritmo 3):** A vigilância contínua detectou que a leitura atual divergia da anterior, ultrapassando o limiar ε .
- 2. Interrupção:** O temporizador de *back-off* foi cancelado imediatamente quando o algoritmo 3 atuou, forçando uma transmissão fora de hora.
- 3. Reconfiguração:** A camada *Fog* calculou o novo NEWS2 e enviou novos parâmetros à borda.
- 4. Monitoramento Intensivo:** O nó passou a coletar e transmitir com frequência elevada, reduzindo a compressão.
- 5. Retorno à Estabilidade:** Quando o paciente voltou a ficar estável, o *back-off* retomou a atuação.

As Figuras ?? e ?? demonstram, respectivamente, o risco e os algoritmos em atuação ao longo desse caso.

Este caso comprova a eficácia do ciclo fechado *Edge-Fog-Cloud*: o sistema economizou recursos quando possível, mas reagiu em tempo hábil à deterioração, comportando-se como um monitor de UTI quando necessário e como um monitor ambulatorial leve nos demais momentos.



Figura 23: Nível de risco do Paciente 37. Podemos notar que o risco permanece instável.

Fonte: Elaborado pelo autor.

6.8 Discussão e Resposta às Questões de Pesquisa

A análise dos resultados permite responder diretamente às questões de pesquisa formuladas na Seção ??, validando as hipóteses do trabalho.

6.8.1 QP1: Teste com múltiplos sinais vitais

A expansão do modelo para contemplar simultaneamente FC e SpO_2 demonstrou que a abordagem de compressão adaptativa multivariada é viável. O algoritmo SDT conseguiu processar ambos os sinais em paralelo na camada de borda, consumindo menos de 0,2% de CPU na instância limitada. A correlação clínica entre os sinais foi tratada na camada *Fog* por meio do NEWS2, o que prova que o modelo suporta a complexidade de múltiplos parâmetros fisiológicos sem degradação de desempenho.



Figura 24: Evolução do intervalo de coleta do Paciente 37. Os pontos verdes indicam os momentos em que o algoritmo de back-off aumentou o intervalo devido à estabilidade. Os pontos vermelhos indicam onde houve uma mudança de risco e, por fim, os pontos amarelos indicam onde o algoritmo 3 precisou atuar, forçando uma coleta em razão da mudança brusca.

Fonte: Elaborado pelo autor.

6.8.2 QP2: Adaptação via notificações Fog/Cloud

O experimento comprovou a eficácia do mecanismo de *feedback*. A latência média do ciclo (T_{loop}) de 1.048,9 ms é adequada para aplicações de monitoramento remoto (não invasivo/não cirúrgico), sendo inferior aos 5,3 segundos relatados por Cassel et al. (?). A comparação é aproximada: Cassel reporta o tempo de processamento para 60% dos sinais muito críticos (percentil), enquanto o ViSPAC reporta a média global de latência de todos os pacientes. Conforme detalhado na Tabela ??, essa latência contempla o tempo de rede, a descompressão e o cálculo do NEWS2, sendo beneficiada pelo encaminhamento assíncrono à *cloud*. Ademais, o sistema foi capaz de reconfigurar dinamicamente os parâmetros IC , ε e DC na borda. A arquitetura baseada em MQTT garantiu que essas notificações de controle fossem leves, sem concorrerem com o tráfego de dados dos sinais vitais.

6.8.3 QP3: Impacto na qualidade dos dados

Os resultados de PRD médio global (1,16% no ViSPAC vs 3,62% no Estático) respondem positivamente a esta questão. O modelo demonstrou que é possível reduzir o volume de dados *melhorando* a qualidade da informação clínica relevante. A estratégia de vincular a agressividade da compressão ao risco clínico (e não a um parâmetro fixo) garante que a distorção ocorra apenas onde é tolerável (sinais estáveis), preservando a morfologia do sinal em eventos críticos.

6.8.4 QP4: Impacto computacional e gestão de coleta

A avaliação de recursos computacionais mostrou que a sobrecarga introduzida pelos algoritmos de decisão (Configurador, Back-off) é negligenciável em relação aos benefícios de rede. O uso de memória manteve-se estável (87 MB) e o uso de CPU permaneceu mínimo. O impacto principal foi positivo: ao reduzir em 96,7% o número de ativações para transmissão, o modelo sugere uma extensão significativa da vida útil da bateria em dispositivos reais, um requisito fundamental para IoT em saúde.

6.8.5 Consulta a Especialista Clínica

Para avaliar a pertinência clínica das decisões de projeto do ViSPAC, foi realizada uma consulta estruturada a uma profissional da área da saúde. A respondente possui formação em Enfermagem, com título de Mestre e Doutora, além de especializações em Saúde Pública, Enfermagem do Trabalho e Auditoria dos Serviços de Saúde. Atua como monitora de simulação clínica em ambiente hospitalar simulado, utilizando protocolos como o NEWS2 e o MEWS em cenários de treinamento com pacientes adultos.

A consulta abordou seis dimensões relevantes para o modelo proposto, cujas respostas são sintetizadas a seguir.

Escolha do escore clínico. A especialista declarou utilizar o NEWS2 e o MEWS em sua prática profissional, confirmando que o escore adotado pelo ViSPAC é reconhecido e aplicado em contextos clínicos reais. Seu nível de confiança na avaliação por escores foi de 6 em uma escala de 1 a 7, indicando elevada confiança no instrumento, ainda que reconhecendo que escores não substituem a avaliação clínica completa.

Seleção de sinais vitais. Questionada sobre quais sinais vitais escolheria para monitoramento remoto residencial limitado a dois ou três parâmetros, a respondente indicou **frequência respiratória, saturação de oxigênio e frequência cardíaca**, por serem “os mais sensíveis para detectar deterioração precoce”. Dois dos três sinais indicados (FC e SpO_2) correspondem exatamente aos utilizados no experimento do ViSPAC, enquanto a frequência respiratória é apontada como extensão prioritária para trabalhos futuros (Seção ??).

Aceitabilidade da latência. A latência de aproximadamente 1,05 s do ciclo de *feedback* do

ViSPAC foi apresentada como referência (arredondada para 1,5 s na formulação da pergunta, para avaliar um limiar mais conservador). A especialista considerou esse tempo “perfeitamente adequado para monitoramento não invasivo”, corroborando a adequação da arquitetura proposta para cenários de vigilância remota.

Confiança em sistemas automatizados. A respondente afirmou que utilizaria um sistema de alerta automático “para apoiar a tomada de decisão, mas nunca como substituto da avaliação clínica”, pois “a tecnologia amplia nossa capacidade de vigilância e antecipa riscos”, embora não capte “nuances como queixas subjetivas e contexto social”. Essa perspectiva reforça o posicionamento do ViSPAC como um **sistema de auxílio à tomada de decisão clínica**: ao fornecer dados priorizados e alertas baseados no NEWS2, o modelo amplia a capacidade de vigilância do profissional sem pretender substituir o julgamento clínico, que incorpora variáveis qualitativas inacessíveis a sensores remotos.

Sensibilidade versus especificidade. Diante da escolha entre receber muitos alertas (com falsos positivos) ou poucos alertas (com risco de omissão), a especialista optou pela **maior sensibilidade**, ressaltando a necessidade de “estratégia para evitar fadiga de alarmes”. O mecanismo de vigilância contínua do ViSPAC (Algoritmo 3) opera nessa direção: mantém a sensibilidade de detecção mesmo durante intervalos longos de *back-off*, enquanto a classificação por nível de risco nas filas de prioridade organiza os alertas para reduzir a sobrecarga informacional.

Critério para intervenção remota. A especialista indicou que enviaria uma ambulância com base exclusivamente em dados remotos apenas “quando percebesse evidência forte de risco iminente de vida e impossibilidade de contato”, o que reforça a importância da classificação por níveis de risco implementada pelo ViSPAC: o sistema prioriza a transmissão de alta fidelidade precisamente nos casos em que o escore NEWS2 indica deterioração grave.

Em síntese, as respostas da especialista corroboram as principais decisões arquiteturais do modelo: a adoção do NEWS2 como escore de referência, a seleção de FC e *SpO₂* como sinais prioritários, a adequação da latência obtida para monitoramento não invasivo e o equilíbrio entre sensibilidade de detecção e eficiência de recursos. A consulta, embora limitada a uma única respondente, fornece uma perspectiva clínica que complementa a avaliação quantitativa apresentada nas seções anteriores.

6.8.6 Comparação com Trabalhos Relacionados

A Tabela ?? apresenta uma comparação quantitativa entre os resultados do ViSPAC e os valores relatados nos trabalhos relacionados. Essa análise permite posicionar a contribuição desta dissertação no contexto do estado da arte.

Os resultados permitem identificar os seguintes posicionamentos: **Taxa de Compressão**. O ViSPAC alcança 81,6% de compressão média, valor comparável aos 80% (5:1) de Chang et al. (?) e superior aos 46% do VSAC (?). Trabalhos como Hassan et al. (?) (15,7:1) e Vakil et al. (?) (70:1) reportam taxas superiores; contudo, esses estudos focam exclusivamente em

Tabela 14: Comparação quantitativa entre ViSPAC e trabalhos relacionados

Trabalho	Compressão	Distorção/Qualidade	Latência/Eficiência
ViSPAC (proposto)	81,6%	PRD médio global 1,16%	T_{loop} 1,05 s; 96,7% redução transmissões
VSAC (?)	46%	Menor que 36%	—
Chang et al. (?)	80%	0,8 erro/402 casos	—
Hassan et al. (?)	93,6%	PRD 2%	91,35% redução tráfego
Vakil et al. (?)	98,6%	—	98% economia energia
Cassel et al. (?)	—	—	5,3 s (P60 críticos)
Shankani et al. (?)	—	—	40% redução tráfego; 64% energia

Fonte: Elaborado pelo autor.

ECG, com métodos de *compressive sensing* ou de predição linear, não avaliando múltiplos sinais vitais nem a adaptação dinâmica. A vantagem do ViSPAC reside na **seletividade clínica**: a compressão agressiva é aplicada apenas a pacientes estáveis, enquanto pacientes críticos recebem transmissão de alta fidelidade. **Distorção (PRD)**. O PRD médio global de 1,16% do ViSPAC é consideravelmente inferior aos 2% reportados por Hassan et al. (?), mesmo com taxas de compressão comparáveis. Essa diferença decorre da estratégia adaptativa: ao concentrar a perda de informação nos períodos de estabilidade clínica e preservar integralmente os dados em momentos críticos, o modelo minimiza a distorção média global sem sacrificar a capacidade diagnóstica. **Latência e Responsividade**. A latência de ciclo de 1.048,9 ms posiciona o ViSPAC favoravelmente em relação aos 5,3 segundos de Cassel et al. (?). Embora as métricas não sejam diretamente comparáveis (Cassel reporta percentil 60 de sinais críticos; ViSPAC reporta média global), a melhoria relativa é atribuída a três fatores: o processamento primário do NEWS2 na camada *fog* (em vez de depender exclusivamente da *cloud*), a leveza do protocolo MQTT e o encaminhamento assíncrono para a nuvem, que remove a espera pela confirmação da nuvem do caminho crítico. **Eficiência de Transmissão**. A redução de 96,7% no número de transmissões supera os 40% de Shankani et al. (?) e os 91,35% de Hassan et al. (?). Vakil et al. (?) reportam 98% de economia energética. Vale notar que a redução de transmissões e a economia de energia são métricas correlatas (menos transmissões implicam menor uso do rádio), porém não equivalentes, pois o consumo energético depende também de fatores como o processamento local e o modo de *sleep*. O ViSPAC oferece uma vantagem qualitativa adicional: a garantia de que pacientes em deterioração clínica recebem monitoramento intensivo imediato, algo não contemplado nos demais trabalhos.

6.8.7 Vantagens e Limitações do ViSPAC

A análise crítica dos resultados permite identificar os seguintes pontos fortes e fracos do modelo proposto:

Vantagens:

- **Eficiência adaptativa:** O modelo ajusta automaticamente os parâmetros de compressão

e de coleta conforme o estado clínico, eliminando a necessidade de configuração manual por paciente.

- **Preservação de dados críticos:** Ao concentrar a compressão agressiva apenas em pacientes estáveis, o ViSPAC garante alta fidelidade nos momentos mais importantes para a tomada de decisão clínica.
- **Vigilância contínua:** O mecanismo de *keep-alive* (Algoritmo 3) oferece uma rede de segurança que detecta deteriorações mesmo durante longos intervalos de *back-off*.
- **Arquitetura distribuída:** O processamento do NEWS2 na *fog* reduz a latência e a dependência da conectividade com a *cloud*.

Limitações:

- **Ambiente simulado:** A validação foi conduzida em ambiente de nuvem com restrições artificiais de hardware, não reproduzindo todas as condições de um ambiente IoT real (interferências, variabilidade de bateria, etc.).
- **Segurança e privacidade dos dados:** O protótipo atual não implementa criptografia ponta a ponta na comunicação MQTT entre as camadas. Em uma implantação real envolvendo dados clínicos sensíveis, o sistema estaria sujeito a regulamentações como a LGPD (Lei Geral de Proteção de Dados) e a HIPAA (Health Insurance Portability and Accountability Act). Para adequação, seria necessário usar de estratégias como a pseudonimização dos identificadores de pacientes (e de quaisquer dados sensíveis ou identificadores) antes da transmissão.
- **Tolerância a falhas no nó Fog:** A arquitetura atual centraliza o processamento do NEWS2 em um único nó *fog*, criando um ponto único de falha. Caso o *fog* fique indisponível, os dispositivos *edge* continuam operando com os últimos parâmetros recebidos, o que pode ser inadequado caso o estado clínico do paciente mude de forma relevante. Em uma implantação de produção, recomenda-se: (i) replicação do nó *fog* com balanceamento de carga; (ii) cálculo local de estimativa do NEWS2 no *edge* como *fallback*; e (iii) em um cenário extremo, a camada *cloud* poderia assumir temporariamente as funções do *fog*, embora com maior latência.
- **NEWS2 parcial com apenas 2 dos 7 sinais:** Os demais parâmetros do NEWS2 (freqüência respiratória, temperatura, pressão arterial sistólica, nível de consciência e uso de oxigênio suplementar) foram fixados em valores de um paciente saudável. Essa simplificação introduz uma limitação clínica significativa: pacientes com alterações nesses parâmetros fixos seriam classificados incorretamente. Em uma implantação real, todos os 7 parâmetros do NEWS2 devem ser monitorados (ou do sistema de escore escolhido).

6.9 Análise de Escalabilidade

Os resultados apresentados até aqui demonstram a eficácia do ViSPAC em 20 nós de borda e em 97 pacientes. Contudo, a viabilidade de um modelo de monitoramento em Cidades Inteligentes depende da capacidade de atender a populações em maior número. Esta seção investiga o limite de escalabilidade da arquitetura proposta a partir de uma análise empírica do componente central do sistema: o nó *fog*.

A escolha do *fog* como objeto de análise decorre de uma observação arquitetural direta. Cada nó *edge* opera de forma autônoma, processando compressão e *back-off* localmente sem depender dos demais. A camada de borda, portanto, escala horizontalmente por construção: adicionar um novo dispositivo *edge* não impacta os demais dispositivos. O nó *fog*, por outro lado, concentra o processamento de todos os lotes recebidos: descomprime os dados, calcula o escore NEWS2, devolve o *feedback* ao *edge* emissor e enfileira os dados para encaminhamento assíncrono à *cloud*. É, por definição, o gargalo potencial da arquitetura. Determinar sua capacidade máxima equivale a determinar a do sistema como um todo.

Para viabilizar essa análise, o código do *fog* foi instrumentado para registrar, a cada lote MQTT recebido, três intervalos de tempo: o tempo de processamento efetivo (*process_ms*), que abrange a descompressão e o cálculo do NEWS2; o tempo de encaminhamento (*forward_ms*); e o tempo total do ciclo (*total_ms*). No Cenário 3, o encaminhamento à *cloud* opera de forma assíncrona por meio de filas de prioridade, de modo que o *forward_ms* registrado é zero: o *fog* responde ao *edge* sem aguardar a nuvem. O tempo de serviço efetivo, portanto, corresponde apenas ao processamento local. Ao longo das 12 horas de simulação, o *fog* processou **35.513 lotes**, fornecendo uma base estatística robusta para a análise.

6.9.1 Utilização Observada do Fog

A Tabela ?? consolida as métricas de tempo extraídas dos logs do *fog* no Cenário 3 (encaminhamento assíncrono).

Tabela 15: Métricas de desempenho do nó *fog* no Cenário 3 (35.513 lotes, 12h)

Métrica	Média	Mediana	Máximo
Processamento NEWS2 (<i>process_ms</i>)	0,88 ms	0,19 ms	102,37 ms
Encaminhamento à <i>cloud</i> (<i>forward_ms</i>)	assíncrono (fora do caminho crítico)		
Total por lote (<i>total_ms</i>)	1,14 ms	0,43 ms	105,78 ms

Fonte: Extraído dos logs do *fog* (Cenário 3, modo assíncrono).

O cálculo do escore NEWS2 consome, em média, **0,88 ms** por lote, com mediana de 0,19 ms. Ao longo de 12 horas e 35.513 lotes, o *fog* dedicou apenas 31 segundos de CPU ao processamento clínico. O tempo total por lote (1,14 ms) inclui a sobrecarga de desserialização *MessagePack* e de publicação da resposta via MQTT, mas permanece na ordem de 1 milissegundo.

Para ilustrar o impacto do encaminhamento assíncrono, nos Cenários 1 e 2, que usam encaminhamento síncrono, o tempo médio de serviço foi de 232,5 ms e 209,5 ms por lote, respectivamente. Nesses cenários, o fog aguardava confirmação da cloud antes de continuar, então a latência inter-regional (208–230 ms) dominava o tempo. No Cenário 3, ao desacoplar o encaminhamento, o tempo de serviço despencou para pouco mais de 1 ms, uma redução de mais de duas ordens de magnitude.

A implementação do *fog* utiliza o MQTT com uma única *thread* de processamento: enquanto um lote é processado, os demais aguardam na fila do *broker*. Esse modelo sequencial configura um sistema de servidor único, cuja utilização (ρ) corresponde à fração do tempo em que o servidor esteve efetivamente ocupado (?):

$$\rho = \frac{n \times \bar{t}fog}{Tsim} = \frac{35.513 \times 0,00114 \text{ s}}{43.200 \text{ s}} = 0,00094 \quad (0,094\%) \quad (6.1)$$

onde n é o número de lotes processados, $\bar{t}fog$ é o tempo médio de serviço por lote (1,14 ms) e $Tsim$ é a duração da simulação em segundos.

O resultado indica que, com 20 nós de borda e 97 pacientes, o fog utilizou menos de **0,1% da sua capacidade de processamento**. A título de comparação, caso o encaminhamento fosse síncrono (como nos Cenários 1 e 2), a utilização seria de 16–25%, limitando a capacidade a 65–103 edges por fog. O modo assíncrono, portanto, não apenas reduz a latência percebida no edge, mas também amplia a capacidade do fog em mais de duas ordens de magnitude.

6.9.2 Projeção de Capacidade e Limites

A independência entre os nós *edge* sustenta uma hipótese de escalamento linear: cada edge adicionado contribui com uma parcela incremental e constante de carga para o *fog*. Sob essa premissa, a utilização projetada para N_e nós de borda é:

$$\rho(N_e) = \frac{N_e}{20} \times 0,00094 \quad (6.2)$$

Em modelos clássicos de filas (e.g., M/M/1), o tempo médio de resposta cresce como $E[T] = \frac{1}{\mu(1-\rho)}$ (?), onde μ é a taxa de serviço. Essa relação hiperbólica implica que, a partir de $\rho \approx 0,8$, pequenos acréscimos de carga causam aumentos desproporcionais na latência: a $\rho = 0,8$, o tempo de resposta é cinco vezes o tempo de serviço; a $\rho = 0,9$, dez vezes. Por isso, adota-se, para essa projeção, $\rho \leq 0,8$ como limiar conservador de operação estável, pois operar próximo desse valor permite absorver variações súbitas de carga e diferenças no perfil de risco dos pacientes, além de evitar a aproximação dos pontos de saturação do sistema. Aplicando esse limiar:

$$N_e^{max} = 20 \times \frac{0,8}{0,00094} \approx 17.000 \text{ edges} \quad (6.3)$$

Esse valor, equivalente a aproximadamente **82.500 pacientes**, é um limite teórico de CPU.

Na prática, outros componentes da infraestrutura se tornariam gargalos antes que o processamento atingisse sua capacidade máxima. A coluna msg/s da Tabela ?? reflete uma segunda dimensão de dimensionamento: a taxa de chegada de mensagens ao *broker*. Essa taxa é derivada da observação dos cenários de teste realizados: o *fog* recebeu 35.513 lotes em 43.200 segundos, resultando em uma taxa média de $\lambda_{edge} = 35.513 / (43.200 \times 20) \approx 0,041$ msg/s por *edge*, linearmente projetada para N_e *edges*. A tabela combina ambas as projeções, indicando, para cada faixa de escala, os componentes de infraestrutura que demandariam atenção.

Tabela 16: Projeção de escalabilidade do ViSPAC (encaminhamento assíncrono)

Edges	Pacientes	ρ (1 fog)	msg/s	Observação
20	97	0,094%	0,8	Experimento realizado
100	485	0,47%	4,1	Margem ampla para uso (manter como experimento)
500	2.425	2,4%	20,5	Margem ampla para uso (manter como experimento)
1.000	4.850	4,7%	41,1	Margem ampla para uso (manter como experimento)
5.000	24.250	24%	205	Monitorar o uso e verificar banda
10.000	48.500	47%	411	Monitorar o uso e verificar banda, podendo ter um <i>broker</i> dedicado
17.000	82.450	80%	698	Límite teórico de CPU ($\rho \approx 0,8$), não recomendado operar nessa margem pois mudanças súbitas em grande escala podem deixar o sistema inoperável

Fonte: Projeção linear a partir da utilização observada no Cenário 3.

Os limites práticos decorrem de três fatores. O primeiro é a capacidade do *broker* MQTT, responsável pela comunicação entre os *edges* e o *fog*: um servidor Mosquitto de instância única suporta, tipicamente, dezenas de milhares de mensagens por segundo para payloads pequenos, o que acomodaria confortavelmente até milhares de *edges* nas taxas observadas. O segundo fator é a memória consumida pelas filas de prioridade: caso a *cloud* apresente lentidão temporária, os dados se acumulam na fila do *fog* até que o *worker* consiga drená-los. Em cenários com milhares de *edges*, picos de enfileiramento podem exigir dimensionamento de memória. O terceiro fator é a largura de banda da conexão HTTP entre o *fog* e a *cloud*, que deve ser suficiente para escoar o volume agregado de dados encaminhados pelo *worker* assíncrono.

Uma consideração adicional diz respeito à composição de risco da população monitorada. No experimento, pacientes de risco alto predominaram e geraram aproximadamente 14 vezes mais tráfego do que os de risco baixo (0,424 pacotes por segundo por *edge* contra 0,030). Essa assimetria decorre da própria lógica do ViSPAC: o *back-off* suprime a maior parte das transmissões de pacientes estáveis, enquanto pacientes críticos transmitem continuamente. Em uma implantação em que a maioria da população apresenta risco baixo ou mínimo, a carga por *edge* seria substancialmente menor, e a capacidade do *fog* poderia superar as projeções da tabela. Em cenários com alta concentração de pacientes críticos, o dimensionamento de múltiplos *fogs* regionais seria necessário.

Em síntese, o encaminhamento assíncrono com filas de prioridade transformou o perfil de escalabilidade do *fog*. O gargalo deixou de ser o processamento (que consome menos de 0,1% da capacidade) e passou a depender de fatores de infraestrutura (broker, memória, rede), que são dimensionáveis de forma convencional. Para demandas superiores, a arquitetura suporta escalamento horizontal: a adição de F nós fog regionais, cada um com seu próprio *broker* MQTT, distribui a carga de forma proporcional.

7 CONCLUSÃO

Esta dissertação apresentou o **ViSPAC**, um modelo voltado ao monitoramento remoto de sinais vitais em arquiteturas *Edge–Fog–Cloud*. O modelo foi concebido como uma extensão do VSAC, incorporando mecanismos de priorização dinâmica baseados no escore clínico NEWS2 e expandindo a capacidade de processamento para múltiplos sinais vitais (FC e SpO_2).

O principal diferencial do ViSPAC reside na integração de um ciclo fechado de *feedback* entre as camadas *fog-cloud* e *edge*, permitindo ajustar dinamicamente os parâmetros de compressão e frequência de coleta conforme o estado clínico do paciente. Essa abordagem foi implementada por meio de três algoritmos interdependentes: (i) o Configurador por Sinal Vital, que configura os parâmetros operacionais com base no risco clínico; (ii) o *Back-off* Exponencial, que otimiza os intervalos de coleta durante períodos de estabilidade dos pacientes; e (iii) a Vigilância Contínua, que garante a detecção de deteriorações clínicas mesmo durante longos intervalos de silêncio.

Os experimentos foram conduzidos em um ambiente distribuído na AWS, simulando uma infraestrutura de Cidades Inteligentes, com 20 nós de borda operando sob restrições de hardware para se assemelharem às de dispositivos IoT. Os resultados demonstraram a eficácia do modelo proposto em múltiplas dimensões:

- **Eficiência de rede:** O ViSPAC reduziu em 96,7% o número de transmissões em comparação à linha de base, passando de mais de 6,6 milhões para 220.403 pacotes em 12 horas de simulação. Essa redução drástica impacta diretamente o consumo de energia de dispositivos vestíveis.
- **Compressão:** A taxa de compressão média atingiu 81,6%, superando os 75,0% do cenário estático (VSAC). A adoção da serialização binária *MessagePack* reduziu o volume bruto dos *payloads* em aproximadamente 30% em relação ao JSON textual, e o volume final trafegado foi reduzido de 245,11 MB para 45,21 MB.
- **Fidelidade clínica:** Apesar da maior compressão, o ViSPAC apresentou distorção PRD média global de apenas 1,16%, substancialmente inferior aos 3,62% do modelo estático. Esse resultado aparentemente contraditório é explicado pela seletividade clínica: a compressão agressiva é aplicada apenas em momentos de estabilidade, preservando a integridade dos dados em situações críticas.
- **Responsividade:** A latência média do ciclo de *feedback* foi de 1.048,9 ms, adequada para aplicações de monitoramento remoto não invasivo.

Retomando a questão de pesquisa principal formulada na Seção ?? — *Como aplicar priorização dinâmica com base em escores de múltiplos sinais vitais, de modo a otimizar a compressão, a coleta e a transmissão de dados clínicos críticos em um ambiente Edge–Fog–Cloud sem*

prejudicar a qualidade das informações? — este trabalho demonstra que a resposta reside na **integração de escores clínicos ao ciclo de controle distribuído**. Ao vincular os parâmetros de compressão (DC , T_{SDT}) e de coleta (IC) diretamente ao nível de risco do paciente, o modelo ViSPAC consegue: (i) reduzir em 81,6% o volume de dados e em 96,7% as transmissões para pacientes estáveis; além de (ii) preservar a fidelidade clínica (PRD médio global 1,16%) concentrando a compressão mínima nos pacientes críticos. A arquitetura *Edge–Fog–Cloud* permite que esse ajuste ocorra em tempo hábil (latência média de 1,05 s), viabilizando uma solução que **otimiza recursos sem comprometer a qualidade da informação clínica**.

Os experimentos corroboram as hipóteses levantadas pelas questões secundárias de pesquisa. Validou-se a viabilidade técnica de utilizar mais de um sinal vital, mantendo o desempenho estável. Paralelamente, o fluxo de notificações *fog-cloud-edge* comprovou sua eficiência via MQTT, e a qualidade dos dados beneficiou-se da compressão orientada ao contexto clínico. Por fim, verificou-se que a sobrecarga computacional decorrente dos algoritmos de decisão é marginal quando comparada aos benefícios auferidos.

Em síntese, o ViSPAC apresenta-se como uma solução equilibrada para o desafio do monitoramento de saúde em larga escala no contexto das Cidades Inteligentes. Ao distribuir a inteligência entre as camadas (processamento do sinal na borda e avaliação do risco na *fog*), o modelo consegue otimizar recursos de rede e energia sem comprometer a capacidade de resposta a eventos clínicos adversos.

7.1 Contribuições

As principais contribuições desta pesquisa são:

1. **Modelo ViSPAC:** Um modelo de compressão adaptativa e priorização inteligente que gerencia múltiplos sinais vitais com ajuste dinâmico baseado no escore NEWS2, integrando compressão híbrida (*lossy + lossless*) a um ciclo fechado de controle *Edge–Fog–Cloud*.
2. **Estratégia de Controle Integrada:** A proposição de três algoritmos interdependentes que, em conjunto, equilibram eficiência e segurança clínica: (i) configuração dinâmica por sinal vital; (ii) *back-off* exponencial condicionado ao risco; e (iii) vigilância contínua de baixo custo (*keep-alive*).

Complementarmente, a avaliação experimental, conduzida em ambiente de nuvem distribuído, com restrições de hardware simuladas para dispositivos IoT, demonstrou a viabilidade técnica do modelo em cenários realistas. O código-fonte e a infraestrutura como código estão disponíveis publicamente, permitindo a reprodução dos experimentos em diferentes contextos.

7.2 Transferência de Tecnologia

A arquitetura e os algoritmos desenvolvidos nesta dissertação têm potencial direto de aplicação no mercado de saúde digital, transcendendo o ambiente acadêmico. A transferência dessa tecnologia para a sociedade pode ser visualizada em três frentes de produtos complementares, cobrindo toda a cadeia de atendimento, conforme ilustrado na Figura ?? e que poderia ser testada em um ambiente de cidade inteligente (como São Leopoldo).



Figura 25: Visão de transferência de tecnologia: da pesquisa acadêmica aos produtos de saúde digital. A camada *Edge* opera em dispositivos vestíveis e em gateways domésticos, coletando dados, comprimindo-os e encaminhando-os para a camada *Fog*, que atua como radar de risco em postos de saúde e ambulâncias do SAMU. Por fim, a camada *Cloud* oferece gestão estratégica às secretarias de saúde. As setas indicam a comunicação contínua entre as camadas.

Fonte: Elaborado pelo autor.

- **Camada Edge (Gateway Doméstico):** Desenvolvimento de um dispositivo concentrador (*System on Chip* ou aplicativo para smartphone) que recebe dados de múltiplos vestíveis (smartwatches, cintas cardíacas, oxímetros) via Bluetooth. Este gateway executaria o ViSPAC, comprimindo os sinais antes de encaminhá-los à camada *Fog*. Como diferencial de mercado, a solução ofereceria uma extensão significativa da vida útil da bateria dos vestíveis (que apenas transmitem via Bluetooth de curto alcance) e redução do consumo de dados móveis. *Exemplo prático:* Um paciente diabético poderia usar seu smartwatch convencional conectado a um gateway ViSPAC em casa, que decide inteligentemente quando transmitir dados para a nuvem, reduzindo os custos do plano de dados em até 90%.
- **Camada Fog (Aplicação de Triagem):** Uma solução de software para instalar em servidores locais das Unidades Básicas de Saúde (UBS) e no SAMU (Serviço de Atendimento Móvel de Urgência). Esta aplicação funcionaria como um radar de risco, permitindo que equipes de socorro visualizem, em tempo hábil, apenas os pacientes em deterioração clínica na sua área de atuação, com latência mínima e independência da nuvem central. *Exemplo prático:* Em um cenário de desastre com múltiplas vítimas, a central do SAMU poderia visualizar um mapa com pontos coloridos (verde, amarelo e vermelho) indicando a criticidade de cada paciente monitorado na região, permitindo priorizar o despacho de ambulâncias aos casos mais graves primeiro.
- **Camada Cloud (Plataforma de Gestão):** Um serviço SaaS (*Software as a Service*) voltado a secretarias de saúde, planos de saúde e hospitais. Diferente da visão tática da *Fog*, esta camada ofereceria uma visão estratégica: análise de tendências da população, auditoria de eventos passados e integração com prontuários eletrônicos, valorizando o armazenamento eficiente dos dados históricos comprimidos pelo ViSPAC. *Exemplo prático:* Uma Secretaria Municipal de Saúde poderia identificar, por meio de dashboards analíticos, que determinado bairro apresenta aumento de 30% nos escores médios do NEWS2 durante ondas de calor, direcionando campanhas preventivas e o reforço de equipes naquela região.

7.3 Trabalhos Futuros

A partir dos resultados obtidos e das limitações identificadas, emergem três direções principais para pesquisas futuras:

1. **Inteligência na borda:** A latência média de 1.048,9 ms corresponde, em sua grande parte, ao tempo de rede entre a edge e a fog. Para reduzi-la, trabalhos futuros podem explorar o cálculo de risco do NEWS2 diretamente no dispositivo edge, utilizando modelos leves de aprendizado de máquina ou o cálculo bruto. Pacientes estáveis poderiam

ser gerenciados localmente, reservando a comunicação com o fog apenas para eventos de mudança de risco ou para validação periódica.

2. **Novos algoritmos de compressão:** O SDT demonstrou-se eficaz, mas outras técnicas podem ser avaliadas para sinais mais variáveis. Algoritmos baseados em wavelets, auto-encoders variacionais e métodos de compressão sensíveis ao contexto clínico representam alternativas promissoras para maximizar a taxa de compressão mantendo a fidelidade diagnóstica.
3. **Algoritmos alternativos de priorização:** A arquitetura modular do ViSPAC permite substituir o NEWS2 por outros escores. Trabalhos futuros podem avaliar escores especializados (PEWS, MEWS), modelos de *machine learning* treinados no histórico do próprio paciente ou a fusão de múltiplos indicadores clínicos em um escore composto.

REFERÊNCIAS

- ALI, A.; MONTANARO, T.; SERGI, I.; CARRISI, S.; GALLI, D.; DISTANTE, C.; PATRONO, L. An Innovative IoT and Edge Intelligence Framework for Monitoring Elderly People Using Anomaly Detection on Data from Non-Wearable Sensors. **Sensors**, [S.l.], v. 25, n. 6, 2025.
- American Heart Association. **Target Heart Rates Chart**. Last reviewed 12 Aug 2024, <https://www.heart.org/en/healthy-living/fitness/fitness-basics/target-heart-rates>.
- ANDRADE, A.; da Costa, C. A.; ROEHR, A.; MUCHALUAT-SAADE, D.; da Rosa Righi, R. Blending lossy and lossless data compression methods to support health data streaming in smart cities. **Future Generation Computer Systems**, [S.l.], v. 167, p. 107748, 2025.
- ANDRÉ SETTI CASSEL, G.; da Rosa Righi, R.; André da Costa, C.; Rosecler Bez, M.; PASIN, M. Towards providing a priority-based vital sign offloading in healthcare with serverless computing and a fog-cloud architecture. **Future Generation Computer Systems**, [S.l.], v. 157, p. 51–66, 2024.
- ANWAR, M.; MASUD, F.; BUTT, R. A.; IDRUS, S. M.; AHMAD, M. N.; BAJURI, M. Y. Traffic Priority-Aware Medical Data Dissemination Scheme for IoT-Based WBASN Healthcare Applications. **Computers, Materials & Continua**, [S.l.], v. 71, n. 3, 2022.
- AWS Pricing Calculator — calculator.aws. [Accessed 19-09-2025], <https://calculator.aws/#/createCalculator>.
- BANKS, A.; GUPTA, R. **MQTT Version 3.1.1**. [S.l.]: OASIS, 2014. OASIS Standard, Acessado em: 2025-05-10.
- Beta Systems. **Rising Demands, Rising Challenges: the evolution of data centers**. Acessado em: 10 dez. 2025, <https://www.betasystems.com/resources/blog/rising-demands-rising-challenges-the-evolution-of-data-centers/>.
- BONOMI, F.; MILITO, R.; ZHU, J.; ADDEPALLI, S. Fog computing and its role in the internet of things. In: FIRST EDITION OF THE MCC WORKSHOP ON MOBILE CLOUD COMPUTING, 2012, New York, NY, USA. **Proceedings...** Association for Computing Machinery, 2012. p. 13–16. (MCC '12).
- BRADEN, R. T. **Requirements for Internet Hosts - Communication Layers**. [S.l.]: RFC Editor, 1989. n. 1122. (Request for Comments).
- BRISTOL, E. H. Swinging door trending: adaptive trend recording. In: ISA NATIONAL CONFERENCE, 1990. **Proceedings...** [S.l.: s.n.], 1990. p. 749–753.
- CHANG, Y.; SOBELMAN, G. E. Lightweight Lossy/Lossless ECG Compression for Medical IoT Systems. **IEEE Internet of Things Journal**, [S.l.], v. 11, n. 7, p. 12450–12458, 2024.
- DAMERA, V. K.; CHERIPELLI, R.; PUTTA, N.; SIRISHA, G.; KALAVALA, D. Enhancing remote patient monitoring with AI-driven IoMT and cloud computing technologies. **Scientific Reports**, [S.l.], v. 15, n. 1, p. 24088, Jul 2025.

DESAI, R. **Dr Rajiv Desai.** 2025.

EDDY, W. **Transmission Control Protocol (TCP).** [S.l.]: RFC Editor, 2022. n. 9293. (Request for Comments).

ESMAEILI, S.; SHAMSI, R. A lightweight and secure sensing model for body area networks in the Internet of Things in biological warfare applications. **Internet of Things**, [S.l.], v. 22, p. 100717, 2023.

FIELDING, R. T.; NOTTINGHAM, M.; RESCHKE, J. **HTTP/1.1.** [S.l.]: RFC Editor, 2022. n. 9112. (Request for Comments).

FRIESEL, D.; SPINCZYK, O. Data Serialization Formats for the Internet of Things. **Electronic Communications of the EASST**, [S.l.], v. 80, 2021. Avaliação analítica de formatos de serialização binária (MessagePack, Protocol Buffers, XDR, Avro) em microcontroladores IoT de 8 a 32 bits.

GENEVA, I. I.; CUZZO, B.; FAZILI, T.; JAVAID, W. Normal Body Temperature: a systematic review. **Open Forum Infectious Diseases**, [S.l.], v. 6, n. 4, p. ofz032, Apr. 2019. Published 2019 Apr 9.

GREENHALGH, T.; KOH, G. C. H.; CAR, J. Covid-19: a remote assessment in primary care. **BMJ**, [S.l.], v. 368, 2020.

GU, B.; KIM, H. S.; KIM, H.; YOO, J. I. Advancements in Wearable Sensor Technologies for Health Monitoring in Terms of Clinical Applications, Rehabilitation, and Disease Risk Assessment: systematic review. **JMIR Mhealth Uhealth**, [S.l.], v. 14, p. e76084, jan 2026. PMID: 41511829; PMCID: PMC12831105.

HARCHOL-BALTER, M. **Performance Modeling and Design of Computer Systems:** queueing theory in action. Cambridge: Cambridge University Press, 2013.

HASSAN, A. M. A.; MOHSEN, S. Compression of electrocardiogram signals using compressive sensing technique based on curvelet transform toward medical applications. **Multimedia Tools and Applications**, [S.l.], v. 84, n. 12, p. 11203–11219, 2024.

Internet Protocol. [S.l.]: RFC Editor, 1981. n. 791. (Request for Comments).

IoT Analytics. **Global IoT Market Forecast:** 23.9 billion connected devices by 2026. Dados extraídos do relatório Global Cellular IoT Connectivity Tracker & Forecast, Relatório de Mercado.

JACOBSON, V. Congestion avoidance and control. In: **SYMPORIUM PROCEEDINGS ON COMMUNICATIONS ARCHITECTURES AND PROTOCOLS**, 1988, New York, NY, USA. **Anais...** Association for Computing Machinery, 1988. p. 314–329. (SIGCOMM '88).

KADHIM, A. N.; MANAA, M. E. Design an efficient internet of things data compression for healthcare applications. **Bulletin of Electrical Engineering and Informatics**, [S.l.], v. 11, n. 3, p. 1678–1686, 2022.

KANEYAGE, J. T.; JONES, J. M.; BURNS, J. C.; JAIN, S.; SUN, X.; JIMENEZ-FERNANDEZ, S.; BERRY, E.; PANCHERI, J. M.; JAGGI, P.; RAMILO, O.; TREMOULET, A. H. Axillary, Oral and Rectal Routes of Temperature Measurement During

Treatment of Acute Kawasaki Disease. **Pediatric Infectious Disease Journal**, [S.l.], v. 35, n. 1, p. 50–53, Jan. 2016.

KIM, K. B.; BAEK, H. J. Photoplethysmography in Wearable Devices: a comprehensive review of technological advances, current challenges, and future directions. **Electronics**, [S.l.], v. 12, n. 13, 2023.

KLEINROCK, L. **Queueing Systems, Volume 1: theory**. New York: Wiley-Interscience, 1975.

BLAZEK, V.; KUMAR V., J.; LEONHARDT, S.; MUKUNDA RAO, M. (Ed.). Pulse Oximetry for the Measurement of Oxygen Saturation in Arterial Blood. In: _____. **Studies in Skin Perfusion Dynamics: photoplethysmography and its applications in medical diagnostics**. Singapore: Springer Singapore, 2021. p. 51–78.

KWAK, K. S.; ULLAH, S.; ULLAH, N. An overview of IEEE 802.15.6 standard. In: **INTERNATIONAL SYMPOSIUM ON APPLIED SCIENCES IN BIOMEDICAL AND COMMUNICATION TECHNOLOGIES (ISABEL 2010)**, 2010., 2010. **Anais...** [S.l.: s.n.], 2010. p. 1–6.

LI, C.; WANG, J.; WANG, S.; ZHANG, Y. A review of IoT applications in healthcare. **Neurocomputing**, [S.l.], v. 565, p. 127017, 2024.

MELEVEEDU, K.; MISKOVSKY, J.; MEHARG, J. et al. Tocilizumab for severe COVID-19 related illness - A community academic medical center experience. **Cytokine X**, [S.l.], v. 2, n. 4, p. 100035, 2020.

MOHAPATRA, A. G.; MOHANTY, A.; NAYAK, S.; MENFASH, H. A.; ALQAHTANI, H.; AL-SHARAEI, A. M.; ALLAF, R.; NAFIE, F. M. IoT-driven remote health monitoring system with sensor fusion enhancing immediate medical assistance in distributed settings. **Alexandria Engineering Journal**, [S.l.], v. 120, p. 627–636, 2025.

MORGAN, S. **The World Will Store 200 Zettabytes Of Data By 2025**. Acessado em: 10 dez. 2025, <https://cybersecurityventures.com/the-world-will-store-200-zettabytes-of-data-by-2025/>.

NGUYEN, N.; LEE, G.; HOLDERREAD, B.; HOLMAN, T.; PLETCHER, S.; SCHWARTZ, R. Scaling Wireless Continuous Vital Sign Monitoring Across an 8-Hospital Health System: digital health implementation report. **JMIR Med Inform**, [S.l.], v. 14, n. 1, p. e78216, 2026.

NHS England. **National Early Warning Score (NEWS)**. Acessado em 8 maio 2025, <https://www.england.nhs.uk/ourwork/clinical-policy/sepsis/nationalearlywarningscore/>.

PARIHAR, A.; PRAJAPATI, J. B.; PRAJAPATI, B. G.; TRAMBADIYA, B.; THAKKAR, A.; ENGINEER, P. Role of IOT in healthcare: applications, security & privacy concerns. **Intelligent Pharmacy**, [S.l.], v. 2, n. 5, p. 707–714, 2024.

PIMENTEL, M. A. F.; JOHNSON, A. E. W.; CHARLTON, P. H.; BIRRENKOTT, D.; WATKINSON, P. J.; TARASSENKO, L.; CLIFTON, D. A. **BIDMC PPG and Respiration Dataset**. [S.l.]: PhysioNet, 2016.

PIMENTEL, M. A. F.; JOHNSON, A. E. W.; CHARLTON, P. H.; BIRRENKOTT, D.; WATKINSON, P. J.; TARASSENKO, L.; CLIFTON, D. A. Toward a Robust Estimation of Respiratory Rate from Pulse Oximeters. **Annals of Biomedical Engineering**, [S.l.], v. 45, n. 8, p. 1914–1927, 2017.

REDDY, C.; DSOUZA, V.; THANGARAJAN, A.; PAWEŁCZAK, P.; KAWSAR, F.; MONTANARI, A. BioPulse: towards enabling perpetual vital signs monitoring using a body patch. , [S.l.], p. 103–108, 02 2025.

REICHL, J.; LEIFKE, M.; WEHRLI, S. et al. Pilot study for the development of an automatically generated and wearable-based early warning system for the detection of deterioration of hospitalized patients of an acute care hospital. **Archives of Public Health**, [S.l.], v. 82, n. 1, p. 179, 2024. Published 2024 Oct 8.

REINSEL, D.; GANTZ, J.; RYDNING, J. **Worldwide IDC Global DataSphere Forecast, 2024–2028**: ai everywhere, but upsurge in data will take time. [S.l.]: International Data Corporation (IDC), 2024. Market Forecast.

RICCALTON, V.; THRELFALL, L.; ANANTHAKRISHNAN, A. e. a. Modifications to the National Early Warning Score 2: a scoping review. **BMC Medicine**, [S.l.], v. 23, p. 154, 2025.

ROBBEN, S.; ENGLEBIENNE, G.; KRÖSE, B. Delta Features From Ambient Sensor Data are Good Predictors of Change in Functional Health. **IEEE Journal of Biomedical and Health Informatics**, [S.l.], v. 21, n. 4, p. 986–993, July 2017. Epub 2016 Jul 22.

Royal College of Physicians. **National Early Warning Score (NEWS) 2**: standardising the assessment of acute-illness severity in the nhs. London: Royal College of Physicians, 2017. Updated report of a working party.

SAHU, M. L.; ATULKAR, M.; AHIRWAL, M. K.; AHAMAD, A. Vital Sign Monitoring System for Healthcare Through IoT Based Personal Service Application. **Wireless Personal Communications**, [S.l.], v. 122, n. 1, p. 129–156, 2022.

SALOMON, D. **Data Compression**: the complete reference. 4. ed. [S.l.]: Springer, 2007.

SHANKANI, R.; KHALIFAVI, M.; SHIRMOHAMMADI, Z.; NIKOOFARD, A. F²mulsense: a novel approach to fuzzy fusion in multisensor data to improve saving energy in wbsn. **The Journal of Supercomputing**, [S.l.], v. 81, n. 6, p. 794, 2025.

TANENBAUM, A. S.; WETHERALL, D. J. **Computer Networks**. 5th. ed. USA: Prentice Hall Press, 2010.

TAO, Y.; WANG, C.; QIN, H. Research on Feistel Encryption Algorithm Based On Wireless Medical Sensor Networks. **Procedia Computer Science**, [S.l.], v. 259, p. 888–896, 2025. Sixth International Conference on Futuristic Trends in Networks and Computing Technologies (FTNCT06), held in Uttarakhand, India.

TOUW, H.; VERHEUL, M.; TUINMAN, P. et al. Photoplethysmography respiratory rate monitoring in patients receiving procedural sedation and analgesia for upper gastrointestinal endoscopy. **Journal of Clinical Monitoring and Computing**, [S.l.], v. 31, n. 4, p. 747–754, 2017.

- VAKIL, M.; SHIRMOHAMMADI, Z. LPRLC: linear predictive run length coding to improve energy consumption of wbans. , [S.I.], 05 2024.
- VAN ANH, D.; NGUYEN, V.-H. Leveraging Priority Queueing in IoT-Edge-Fog-Cloud Infrastructures for Efficient Healthcare Monitoring. **IEEE Access**, [S.I.], v. 13, p. 80461–80477, 2025.
- VITAZKOVA, D.; FOLTAN, E.; KOSNACOVA, H.; MICJAN, M.; DONOVAL, M.; KUZMA, A.; KOPANI, M.; VAVRINSKY, E. Advances in Respiratory Monitoring: a comprehensive review of wearable and remote technologies. **Biosensors**, [S.I.], v. 14, n. 2, 2024.
- WEI, S.; XIONG, D.; WANG, J.; LIANG, X.; WANG, J.; CHEN, Y. The Accuracy of the National Early Warning Score 2 in Predicting Early Death in Prehospital and Emergency Department Settings: a systematic review and meta-analysis. **Annals of Translational Medicine**, [S.I.], v. 11, n. 2, p. 95, 2023.
- ZHANG, L.; CAO, B.; LI, Y.; PENG, M.; FENG, G. A Multi-Stage Stochastic Programming-Based Offloading Policy for Fog Enabled IoT-eHealth. **IEEE Journal on Selected Areas in Communications**, [S.I.], v. 39, n. 2, p. 411–425, 2021.
- ZHOU, C.; YE, X.; WEI, Y.; De Florio, V.; SUN, H.; ZHAN, X.; LI, Y.; WANG, C.; ZHANG, X. A comprehensive evaluation of multiple video compression algorithms for preserving BVP signal quality. **Biomedical Signal Processing and Control**, [S.I.], v. 103, p. 107445, 2025.
- ZIYA, O. **Biosensor Student Health Fitness Data**. Dados biométricos de estudantes universitários saudáveis, incluindo FC e SpO2, Kaggle Dataset.