

## **Proposta para validação de um kit de monitoramento remoto de pacientes – uma revisão de escopo**

### **Proposal for validation of a remote patient monitoring kit - a scoping review**

DOI:10.34117/bjdv7n10-214

Recebimento dos originais: 07/09/2021

Aceitação para publicação: 18 /10/2021

#### **Olávio Gonçalves de Almeida**

Graduação

Universidade Federal de Catalão – UFCAT

Av. Dr. Lamartine Pinto de Avelar, 1120, Setor Universitário, Catalão-GO, Brasil.

E-mail: olaviog.almeida@gmail.com

#### **Dalton Matsuo Tavares**

Doutorado

Unidade Acadêmica Especial de Biotecnologia – IBiotec / Universidade Federal de Catalão – UFCAT

Av. Dr. Lamartine Pinto de Avelar, 1120, Setor Universitário, Catalão-GO, Brasil.

E-mail: dalton\_tavares@ufg.br

#### **Stella Jacyszyn Bachega**

Doutorado

Faculdade de Engenharia – FENG / Universidade Federal de Catalão – UFCAT

Av. Dr. Lamartine Pinto de Avelar, 1120, Setor Universitário, Catalão-GO, Brasil.

E-mail: stella@ufg.br

#### **Samuel Henrique França Leite**

Graduação

Universidade Federal de Catalão – UFCAT

Av. Dr. Lamartine Pinto de Avelar, 1120, Setor Universitário, Catalão-GO, Brasil.

E-mail: samuelfranca@discente.ufg.br

#### **RESUMO**

Dispositivos para o monitoramento remoto de pacientes possuem capacidade para coletar dados do paciente. Tendo isso em vista, este trabalho tem como objetivo apresentar uma proposta para um modelo de validação dos dados coletados a partir de um sensor conectado a uma plataforma de monitoramento remoto. O modelo foi escolhido após uma revisão de escopo feita na literatura, buscando identificar quais métodos são utilizados para fazer a validação e análise de dados provenientes de sensores de monitoramento remoto. O modelo selecionado recebe como entrada os dados coletados pelo sistema e identifica os outliers presentes no conjunto de dados. Em relação ao modelo selecionado, foi possível identificar que o mesmo é capaz de lidar adequadamente com a detecção de outliers no cenário proposto.

**Palavras-chave:** métodos de validação, monitoramento remoto.

## ABSTRACT

Devices for remote patient monitoring have the ability to collect data from the patient. With this in mind, this paper aims to present a proposal for a validation model of the data collected from a sensor connected to a remote monitoring platform. The model was chosen after a scoping review of the literature to identify which methods are used to validate and analyze data from remote monitoring sensors. The selected model receives as input the data collected by the system and identifies the outliers present in the data set. In relation to the selected model, it was possible to identify that it is able to adequately deal with outlier detection in the proposed scenario.

**Keywords:** validation methods, remote monitoring.

## 1 INTRODUÇÃO

O presente trabalho se insere no contexto do monitoramento remoto de pacientes. Essa forma de monitoramento consegue conectar o paciente com o médico de forma remota, sendo possível transmitir várias informações (MORRISSEY, 2014). Isso permite a aquisição de dados como: atividade física, peso, qualidade de sono, frequência cardíaca, pressão sanguínea etc (YANG et al., 2018). Os dados são coletados por meio de dispositivos pequenos e vestíveis em contato com o corpo do paciente. Estes devem ser transmitidos para um local onde o profissional de saúde consiga fazer sua análise e formular uma visão da situação atual do paciente.

Os dados podem ser enviados para um computador, uma estação base, ou algum serviço em nuvem (MORRISSEY, 2014). Após a análise, existe a chance de o profissional de saúde detectar alguma anomalia nos dados e assim, identificar o melhor tratamento. Com isso, a prevenção de agravamento do problema pode evitar consequências severas (MORRISSEY, 2014). Com isso, é possível ocasionar uma redução nas filas e congestionamentos hospitalares (CARVALHO et al., 2010).

Um ponto importante a ser tratado com relação ao monitoramento remoto de pacientes é a confiabilidade e também a qualidade dos dados coletados. Eventos com o mesmo valor sendo coletado repetidamente e falhas de circuito, podem gerar anomalias no resultado (CHEN et al., 2015). Os dados que sofrem desvio além de um limiar pré-estabelecido, em relação aos dados já coletados, podem ser considerados anomalias.

O contexto geral deste trabalho é fazer uma análise dos dados coletados por um kit de monitoramento remoto de pacientes proprietário através de um método proposto pela literatura. Os dados analisados poderão ser usados em estudos futuros, como por exemplo, servindo como base de comparação para averiguar a precisão e a confiabilidade

de outro kit de monitoramento remoto aberto (*open source*). Para realizar essa análise, será feito um estudo de metodologias encontradas na literatura, em relação a validação e análise dos dados coletados por sensores. A partir disso, será escolhida a metodologia mais adequada às características descritas. Assim, o objetivo deste trabalho é identificar métodos para a validação de dados coletados por sensores e selecionar o mais apropriado.

O trabalho está organizado da seguinte maneira: na Seção 2 será apresentada a metodologia utilizada na pesquisa; a Seção 3 apresenta os métodos identificados na literatura; a seção 4 realiza uma discussão sobre os métodos introduzidos na Seção 3; e por fim, na Seção 5 é apresentada uma conclusão geral do trabalho.

## 2 METODOLOGIA

Para este trabalho, foi adotada uma metodologia denominada revisão de escopo. Com base no método de Levac, Colquhoun e O'Brien (2010), foi feita uma adaptação do protocolo de revisão de escopo que consiste na execução das seguintes etapas: (1) identificação da hipótese de pesquisa; (2) identificação das questões de pesquisa com base na hipótese; (3) identificação de estudos relevantes a partir de alguma base de conhecimento; (4) pré-seleção de estudos de acordo com critérios de inclusão/exclusão; (5) extração de dados mediante o agrupamento de informações relevantes; e (6) sumarizar e relatar os resultados.

Sendo assim, terminada a primeira etapa, foi elaborada a hipótese inicial: “é possível fazer a validação de um sistema de sensores proprietário (i.e. provar que este é confiável)”. Assim, foram identificadas as seguintes questões de pesquisa:

1. Quais são os métodos para validação de sensores existentes?
2. Os métodos são aplicáveis a sensores biomédicos?

De acordo com a terceira etapa do método de revisão de escopo, foi iniciada a busca por estudos relevantes. Para a seleção destes estudos, foi escolhida a base de dados do Portal CAPES e foram aplicados os seguintes critérios de inclusão/exclusão durante a fase de pré-seleção de artigos: 1) artigos devem ser revisados por pares; 2) devem se referir ao período de publicação (2007-2019); 3) idiomas devem ser inglês, espanhol ou português; 4) título deve estar alinhado com a questão de pesquisa ou com a hipótese; 5) resumo deve estar alinhado com a questão de pesquisa ou com a hipótese.

Ao todo, após a aplicação dos critérios de inclusão/exclusão, foram identificados na base de dados um total de 120.117 artigos, e destes, 42 foram selecionados após a aplicação dos critérios 1, 2 e 3. Após a aplicação dos critérios 4 e 5 restaram 25 artigos.

Estes foram usados durante a fase de extração de dados. Durante a fase de extração de dados, foram selecionados os pontos considerados importantes em cada um dos estudos selecionados, com o objetivo de reunir o máximo de informações possíveis sobre métodos, técnicas e ferramentas utilizadas por autores da área.

A fase final é a de síntese, onde se faz um sumário de cada um dos métodos selecionados. Estes serão abordados na Seção 3, destacando-se os principais e o motivo destes terem sido considerados importantes para o desenvolvimento da pesquisa. Pretende-se apresentar os pontos importantes de cada método selecionado e posteriormente, o método que melhor se encaixa nas características específicas para o desenvolvimento do trabalho.

### **3 MÉTODOS DE VALIDAÇÃO E ANÁLISE DOS DADOS**

Esta Seção apresenta o levantamento referente aos métodos usados na validação de dados provenientes de sensores. A validação dos dados será feita por meio da identificação de anomalias e *outliers* presentes no conjunto de dados a ser validado. Uma anomalia pode ser definida como uma medida que desvia fortemente do padrão normal dos dados coletados (YUAN; ZHOU; CHEN, 2012a). O termo *outlier* em sua definição é similar ao termo anomalia, significando medidas de dados que desviam do restante do conjunto de dados (FAWZY; MOKHTAR; HEGAZY, 2013). A seguir, serão apresentados os métodos de validação selecionados durante a fase de revisão de escopo.

#### **3.1 GREY BOOTSTRAP METHOD (GBM)**

Método proposto por Chen et al. (2015) para fazer a validação online de dados coletados de sensores. É capaz de isolar múltiplas falhas de um sensor e também realizar a recuperação de falhas com uma boa acurácia, além de ter um bom desempenho quando é necessário diferenciar entre uma mudança repentina e uma falha nos dados. A detecção de dados anômalos é feita calculando-se a diferença entre o valor medido e o valor previsto, a qual é comparada com um valor limite. O valor previsto pode ser obtido por meio de um modelo matemático, que usa como base as medições anteriores.

Em testes experimentais, Chen et al. (2015) compararam os resultados obtidos com outros métodos, sendo eles: o Back-Propagation Neural Network (BPNN), o Radial Basis Function Neural Network (RBFNN) e o Relevance Vector Machine (RVM). O método proposto possui vantagem quando comparado a métodos baseados em redes neurais (BPNN e RBFNN), pois estes requerem uma grande quantidade de dados para

treinamento, além de serem mais complexos. Outras vantagens notadas, dizem respeito ao tempo, precisão e complexidade computacional em relação aos demais métodos comparados.

### 3.2 VALIDATED RANDOM FUZZY VARIABLE (VRFV)

Shen e Wang (2013a) relatam em seu trabalho um método de validação online de dados

coletados em uma rede de sensores, usando filtro de predição polinomial, acoplado a VRFV. O método ainda é capaz de decidir entre um dado que sofreu uma mudança repentina ou se o dado em questão se trata de uma falha. Possui baixa complexidade computacional, fornecendo uma estrutura simples, que pode ser muito favorável para aplicações em tempo real.

Em verificação experimental, Shen e Wang (2013a) identificaram que o método foi capaz de distinguir entre um dado que sofreu uma mudança repentina de fato e uma falha que ocorreu, além de conseguir identificar falhas, como por exemplo, falhas de circuito (hardware), produzindo resultados precisos. Para realizar a distinção entre um valor que sofreu mudança repentina de uma falha, o método faz uma comparação do valor coletado com as outras unidades sensíveis equivalentes. Neste caso, se todas as unidades sensoras fizerem uma medição que se assemelha a este valor, o dado é considerado normal. Caso contrário, o dado pode ser considerado uma falha, visto que este só foi coletado em uma das unidades sensoras.

Shen e Wang (2013a) compararam o método criado com o BPNN e um método baseado em Autoregressive Moving Average (ARMA). Como os dois métodos são baseados em redes neurais, o método proposto por Shen e Wang (2013a) mostrou vantagem em relação ao tempo e também a complexidade computacional.

### 3.3 RELEVANCE VECTOR MACHINE (RVM) AND PRINCIPAL COMPONENT ANALYSIS (PCA)

Shen e Wang (2012) propõem um método baseado em RVM juntamente com PCA para a detecção, isolamento e recuperação de falhas (*failure detection, isolation, and recovery* - FDIR) em tempo real de um sensor multifuncional. Este método é capaz de isolar múltiplas falhas simultaneamente, com um bom tempo de resposta e uma alta taxa de acerto. Além disso, o método ainda é capaz de identificar e diferenciar um dado que sofreu uma mudança repentina de uma falha real. Para verificar se a saída de dados é

válida ou não, é feita uma comparação com o limite estabelecido, de acordo com o cálculo de predição que é obtido usando métodos estatísticos baseados em PCA. Por fim, Shen e Wang (2012) afirmam que o método proposto deve ter uma estrutura simples, para aumentar e garantir o seu desempenho em aplicações online.

Shen e Wang (2012) fizeram uma comparação experimental com os métodos BPNN, RBFNN e Polynomial Finite-impulse-response Filter Predictors (PFPs). Com isso, Shen e Wang (2012) notaram que o modelo proposto mostrava vantagem sobre os demais métodos. O método BPNN é bastante ineficaz, quando se tem poucas amostras para a fase de treinamento, além de consumir maior tempo, por ser baseado em redes neurais. Já o método RBFNN consumiu mais recursos computacionais em relação ao método proposto, também pelo fato de ser baseado em redes neurais. O terceiro método (PFP) tem uma boa capacidade preditiva, mas após uma mudança súbita, o seu tempo de resposta é maior que o método proposto.

#### 3.4 MULTIVARIABLE RELEVANCE VECTOR MACHINE (MVRVM)

O método usado por Shen e Wang (2013b) para a validação de dados online, denominado Multivariable Relevance Vector Machine (MVRVM) trabalha com atualizações online do filtro de predição, com o objetivo de detectar múltiplas falhas, isolamento e recuperação das mesmas. Este método divide o sinal em pequenos segmentos, usando polinômios de baixo grau. A fase de treinamento do filtro de predição polinomial necessita de algumas amostras de dados livre de falhas para treinar a predição. Além disso, o modelo apresentado ainda é capaz de diferenciar mudanças súbitas de falhas obtidas na medição dos dados, comparando os dados obtidos em uma unidade sensora com os vizinhos. Caso a diferença aconteça somente em uma das unidades sensoras, o dado é definido como uma falha. Caso a mesma medição apareça em todas as unidades vizinhas, a medição é caracterizada como uma mudança repentina.

Esse método também foi comparado com os métodos BPNN e RBFNN, e ao realizar esta comparação, Shen e Wang (2013b) identificaram que o método proposto foi mais rápido e eficiente na realização do FDIR, devido a sua complexidade computacional ser mais baixa e levar menos tempo em relação aos outros métodos. Grande parte dessa diferença se deve ao fato dos outros métodos serem baseados em redes neurais, necessitando de mais tempo para treinamento da predição e também por terem uma complexidade computacional maior, quando comparados ao método proposto. Devido a

essas diferenças, Shen e Wang (2013b) classificaram o modelo apresentado como uma melhor solução para aplicações de tempo real.

### 3.5 ALGORITHM FOR DATA VALIDATION

Ravichandran e Arulappan (2013) criaram um método de validação fazendo a junção dos seguintes algoritmos: Heuristic Rule, Temporal Correlation e Modified Z-Score, tornando o algoritmo capaz de identificar diferentes tipos de falhas. Neste algoritmo, Heuristic Rule é usado na detecção falhas de limite. Em seguida, usa-se a parte de Temporal Correlation nos dados considerados normais. Por fim, aplica-se o Modified Z-Score no resultado do passo anterior para detecção de *outliers*. O algoritmo final proposto por Ravichandran e Arulappan (2013) trabalha com validação online. Assim, para cada unidade sensora faz-se a coleta de dados em um certo intervalo de tempo. O algoritmo é aplicado nos dados coletados, enviando aqueles considerados normais para a estação base e descartando os dados anômalos. Quando os dados chegam na estação base, o algoritmo é aplicado novamente para garantir que somente dados válidos serão passados ao usuário, garantindo um melhor desempenho do método.

Experimentalmente, Ravichandran e Arulappan (2013) utilizaram unidades sensoras de temperatura, umidade, um sensor de monóxido de carbono e um de dióxido de carbono. Uma simulação foi realizada, e para testar o algoritmo, alguns dados anômalos foram inseridos propositalmente. Como resultado, concluiu-se que o algoritmo proposto obteve melhores resultados em relação aos métodos Heuristic Rule, Temporal Correlation, Spatial Correlation e Modified Z-Score, quando aplicados isoladamente. Sendo assim, o algoritmo final proposto obteve uma maior porcentagem de falhas encontradas.

### 3.6 LINEAR PROGRAMMING-BASED FUZZY-CONSTRAINT SVDD (LP-FCSVDD)

GhasemiGol et al. (2015) criaram uma versão melhorada do Support Vector Data Description (SVDD), chamada Linear Programming-based Fuzzy-Constraint SVDD (LP-FCSVDD). O SVDD é um tipo de abordagem de classificação em classes, baseada em clusterização de dados por meio de Support Vector Machine. O método apresentado e proposto por GhasemiGol et al. (2015), chamado LP-FCSVDD possui complexidade inferior quando comparado ao método padrão.

Segundo GhasemiGol et al. (2015), cada sensor deverá executar o método no seu próprio conjunto de dados, gerando um raio local, o qual é transmitido para as unidades sensoras vizinhas a esta. Quando o sensor recebe os dados dos raios de todos os vizinhos, ele computa um raio global, o valor de confiança de seus vizinhos e também o seu próprio valor de confiança. Os sensores que possuem um baixo valor de confiança são considerados comprometidos e o raio transmitido por eles é ignorado. Depois disso pode-se calcular os pesos de cada sensor de acordo com seu próprio valor de confiança. Por fim, os sensores executam o método novamente com a nova amostra de peso dos sensores.

O método proposto trabalha de forma online, porém em seu experimento, GhasemiGol et al. (2015) utilizaram dois cenários para avaliar o modelo. No primeiro, dados anômalos eram gerados aleatoriamente e adicionados juntamente aos dados normais. No segundo, fez-se uma suposição de que algum sensor gerava anomalias intencionais. Com isso, foi possível destacar algumas vantagens da abordagem proposta por GhasemiGol et al. (2015): (1) pode detectar anomalias com uma grande precisão; (2) poderia ser usada em esquemas distribuídos, com uma mínima comunicação; e (3) poderia eliminar anomalias intencionais.

### 3.7 KERNEL PRINCIPAL COMPONENTS ANALYSIS (KPCA)

Ghorbel et al. (2015) usam o método Kernel Principal Components Analysis (KPCA), baseado no Kernel Mahalanobis. Juntamente com o seu método, Ghorbel et al. (2015) usaram a distância de Mahalanobis para calcular implicitamente o mapeamento dos dados, diferenciando os *outliers* dos dados normais. Basicamente, um *outlier* é detectado se o dado está fora do limite permitido. Para fazer o cálculo do limite, Ghorbel et al. (2015) utilizaram os valores coletados pelos sensores em um cluster, calculando um limite global. O método proposto pode ser utilizado em tempo real, porém Ghorbel et al. (2015) conduziram seus experimentos em *datasets*.

Para comprovar a eficiência do método, são usados três conjuntos de dados: (1) Intel Berkeley Research Lab; (2) Grand St. Bernard e (3) Sensorscope Lausanne Urban Canopy Experiment. Dentre os dados presentes nos *datasets*, estavam dados de temperatura e umidade. Para testar o desempenho do método proposto, foram geradas cerca de 120 anomalias aleatoriamente e estas foram inseridas nos conjuntos de dados. Para a fase de treino, foi realizada uma média de 10 execuções, para todos os conjuntos de dados. Por meio dos resultados obtidos ao analisar o experimento, Ghorbel et al. (2015) conseguiram demonstrar que o método é válido e eficiente.

### 3.8 IN-NETWORK KNOWLEDGE DISCOVERY APPROACH

Fawzy, Mokhtar e Hegazy (2013) propõem um método baseado em conhecimento na rede para fazer a detecção de *outliers* em um conjunto de dados. O método proposto é capaz de decidir entre um erro devido a falha do sensor ou um erro por causa de algum evento que tenha ocorrido, o que torna o método adaptativo. O método utiliza correlação espacial e temporal para verificar se tal mudança aconteceu nos dados medidos. Para fins experimentais, este método foi comparado com outro método de detecção chamado Collaborative Event detection and Tracking protocol (COLLECT).

Em ambiente experimental, o método foi testado com dados coletados no Intel Berkeley

Research, possuindo aproximadamente 2,3 milhões de leituras. O método proposto por Fawzy, Mokhtar e Hegazy (2013) foi capaz de obter uma taxa de detecção completa de todas as anomalias, gerando uma taxa baixa de alarme falso. O método foi comparado com o método COLLECT e nessa comparação foi possível perceber que o modelo proposto por Fawzy, Mokhtar e Hegazy (2013) obteve resultados mais satisfatórios, em relação a quantidade de anomalias identificadas e a taxa de alarme falso gerada. Outro experimento realizado foi feito em um conjunto de dados sintético contendo 5.000 objetos, com 53 *outliers* incluídos. Novamente o método proposto foi capaz de detectar todos os *outliers*. Com os resultados obtidos, Fawzy, Mokhtar e Hegazy (2013) concluíram que o método desenvolvido é eficiente e preciso em identificar *outliers* em conjuntos de dados.

### 3.9 PIECEWISE AGGREGATE APPROXIMATION (PAA) AND ARTIFICIAL IMMUNE SYSTEM (AIS)

Guo, Wang e Chen (2015) apresentam um algoritmo para detecção de anomalias baseado em um algoritmo de fusão de dados, que pode ser usado em aplicações de tempo real. Primeiramente, o algoritmo emprega Piecewise Aggregate Approximation (PAA) para comprimir os dados coletados, fazendo assim uma economia de energia dos sensores e diminuindo a quantidade de comunicação entre eles. Posteriormente, é aplicado um algoritmo de detecção baseado em K-Means e Artificial Immune System (AIS) para classificar os dados comprimidos na fase anterior em dados normais ou dados anômalos. O algoritmo usado trabalha com *clusters* de dados e distância euclidiana. Dependendo da distância do dado ao *cluster*, esse dado é considerado anômalo ou não.

Apesar do método trabalhar em tempo real, Guo, Wang e Chen (2015) conduziram os experimentos a partir de dois conjuntos de dados sintéticos que continham dados medicinais com a presença de anomalias. A taxa de detecção para o algoritmo proposto foi de aproximadamente 95% e a taxa de alarmes falsos foi de aproximadamente 5%. Ao realizar uma comparação desse método com outros propostos por outros autores, foi possível chegar a conclusão de que o método é eficiente quando aplicado em dados coletados por sensores, com a finalidade de detectar anomalias neles.

### 3.10 TWO-SIDED MEDIAN METHOD AND ONE-SIDED MEDIAN METHOD

Basu e Meckesheimer (2007) propõem duas variações de um modelo para detecção de

*outliers*. No primeiro modelo, denominado Two-sided Median Method, o objetivo principal é

usar a mediana da vizinhança de dados para verificar se determinado dado é um *outlier* ou não. No segundo modelo é feito um sistema de predição usando também a mediana. Se a diferença do valor medido atual for maior que um limite permitido estabelecido, o dado é considerado *outlier* e substituído pelo valor predito anteriormente. Caso a medição esteja dentro do limite estabelecido, a medição é considerada normal e é mantida. Este último modelo foi denominado por Basu e Meckesheimer (2007) como One-sided Median Method. Ambos os métodos, após detectarem os valores considerados anômalos, deixam a cargo do profissional classificar aquela medida entre uma anomalia de fato, ou algo válido que aconteceu durante a coleta de dados. Sendo assim, não é um modelo totalmente automático, porém pode ter melhores resultados, uma vez que o profissional sabe ao certo se o dado é válido ou não.

Ao realizar os experimentos, foi possível notar nos resultados que os dois métodos foram eficientes em detectar anomalias e substituí-las quando necessário. Basu e Meckesheimer (2007) afirmaram que dependendo da aplicação, cada método pode trazer melhores resultados. Caso a aplicação seja em tempo real, Basu e Meckesheimer (2007) recomendam o método One-sided Median Method. Para a aplicação offline, o método Two-sided Median Method é mais apropriado. Os resultados do teste feito por Basu e Meckesheimer (2007) mostraram que os métodos podem ser eficientes, se aplicados na situação correta.

### 3.11 PCA MODEL AND STATISTICAL TECHNIQUES

Nisha et al. (2014) apresentaram uma técnica para análise de dados multivariados que trabalha de forma online, com um algoritmo de duas etapas. Na primeira etapa é usado um modelo baseado em Principal Component Analysis (PCA) para detecção de falhas. Esse modelo PCA particiona o espaço de medição em dois espaços ortogonais: um espaço com dados normais e outro com variações de dados. Para essa divisão de espaços é usada a distância de Mahalanobis. Sendo assim, dado é considerado como uma anomalia se essa distância exceder um limite permitido. O limite é definido de acordo com a média dos valores das distâncias. Na segunda etapa, algumas técnicas estatísticas são aplicadas para determinar alguma semelhança entre o dado medido contra o conjunto de dados reais. Quando um nó sensor identifica um comportamento diferente ele repassa essa informação aos nós próximos a ele, fazendo o uso também de correlação espacial para isso. Quando o algoritmo identifica algum tipo de anormalidade em um nó sensor ele deve reportar esse fato para a estação base.

Os dados obtidos através da fase experimental indicam que o método foi capaz de identificar grande parte de dados anômalos, chegando em cerca de 90%, produzindo apenas cerca de 10% de taxa de alarmes falsos. Através destes experimentos, Nisha et al. (2014) puderam chegar à conclusão de que o método proposto pode ser considerado eficiente para realizar tal tarefa.

### 3.12 OUTLIER-ROBUST KALMAN FILTER (ORKF) AND LOCALLYWEIGHTED PARTIAL LEAST SQUARES (LW-PLS)

Feng et al. (2017) apresentam um sistema que faz o uso de um sistema híbrido para predição de valores e detecção de anomalias em tempo real. Para isso, utilizaram duas técnicas principais: ORKF e LW-PLS. O primeiro método (ORKF) é utilizado para fazer a detecção e eliminação da medição errada, já o LW-PLS é usado para realizar as predições de valores. O sistema também inclui Nominal Angle Analysis (NAA) para distinguir entre falhas de sinal e mudanças repentinas. Feng et al. (2017) ainda afirmaram que o sistema híbrido aumenta a probabilidade da detecção rápida de falhas no sensor e também melhora a precisão de detecção.

No experimento conduzido para realizar a análise do método, Feng et al. (2017) notaram que o modelo proposto é eficiente, pois o mesmo conseguiu detectar com sucesso a maioria dos erros do sensor e reconciliar suas leituras para quase todos os erros

detectados com valores mais próximos do valor real. Além disso, o número de falsos alarmes foi relativamente baixo.

### 3.13 TIME-SERIES ANALYSIS AND GEOSTATISTICS

Zhang et al. (2012) apresentaram uma metodologia que faz o uso de correlação espacial e temporal nos dados coletados para identificar o comportamento normal dos dados e posteriormente, identificar os dados anômalos, distinguindo-os entre erros ou eventos que poderiam ter acontecido no ambiente de monitoramento, de maneira distribuída e online. Neste contexto, Zhang et al. (2012) definem que esse evento que pode ter acontecido no ambiente, pode se tratar tanto de uma mudança gradual quanto repentina. Isso permite que o método seja capaz de distinguir um erro de uma mudança repentina. Além disso, são empregadas nos modelos técnicas estatísticas para a detecção de falhas.

Após experimentos realizados em (ZHANG et al., 2012), foi possível observar que o modelo Temporal Outlier Detection (TOD) possuía baixa complexidade em relação às comunicações, porém a precisão não era tão boa quanto a de outros modelos que utilizavam um grande número de comunicações (correlação espacial). O segundo modelo Spatial Outlier Detection (SOD) apresentava resultados precisos, porém realizava um grande número de comunicações.

Os modelos Spatial Predicted-data-based Outlier Detection (POD) e Spatial and Temporal Integrated Outlier Detection (STIOD) foram desenvolvidos para diminuir a quantidade de comunicações realizadas, porém, percebeu-se uma queda na precisão do modelo. O modelo STIOD permitiu que cada nó sensor detectasse mudanças no comportamento normal dos dados, utilizando uma técnica de predição. Por meio da comparação do valor medido com a previsão feita, o modelo foi capaz de classificar a medição entre normal ou anômala. Neste método, o número de comunicações permaneceu elevado, mas não tanto quanto no SOD. Zhang et al. (2012) afirmam que é inevitável manter uma alta precisão e um baixo número de comunicações quando é necessário a incorporação da dimensão espacial.

### 3.14 ADAPTIVE ELLIPTICAL ANOMALY DETECTION

Moshtaghi et al. (2014) apresentaram uma abordagem adaptativa utilizada para criar limites de decisão elíptica, que pode ser utilizada em aplicações de tempo real. Nesse modelo, cada sensor faz o cálculo dos limites locais do elipsoide e o envia para uma

estação base. Logo em seguida, a estação base que recebeu os dados é capaz de agrupar todos os limites recebidos dos nós sensores e calcula um limite de elipsoide global utilizando uma abordagem estatística denominada Chebyshev *inequality*. Com esse limite global calculado, a estação base envia este de volta para os sensores. Assim, os dados que se localizam fora dos limites estabelecidos do elipsoide são considerados anômalos e enviados para a estação base. Através destes dados recebidos pela estação base, é possível atualizar o limite global que é utilizado pelos sensores, tornando o modelo adaptativo.

Apesar da abordagem de Moshtaghi et al. (2014) trabalhar em tempo real, os experimentos foram conduzidos offline, utilizando um conjunto de dados sintético. Por meio dos experimentos foi possível concluir que o modelo é capaz de ser eficiente em ambientes não homogêneos, porém, o método depende fortemente da clusterização dos dados e da similaridade entre os elipsoides.

### 3.15 SUBJECTIVE LOGIC BASED ANOMALY DETECTION (SLAD)

Yuan, Zhou e Chen (2012b) apresentam uma ferramenta chamada de Subjective Logic Based Anomaly Detection (SLAD). Neste modelo, primeiramente cada nó sensor coleta uma quantidade de dados e envia para a estação base, permitindo que esta calcule um modelo auto-regressivo. Posteriormente, cada nó recebe um coeficiente do modelo auto-regressivo construído para fazer uma predição de dados, utilizando a coleta mais recente. Ao realizar essa predição, o nó sensor é capaz de compará-lo com a próxima coleta de dados e classificar a coleta entre uma coleta normal ou uma coleta com dados anômalos. Caso a diferença da predição com os dados coletados seja maior que o limite estipulado, o dado é considerado anômalo. Por fim, aquele sensor que possuir seu dado considerado anômalo, envia o mesmo para os sensores vizinhos com o objetivo de obter um placar geral. Caso o placar resultante seja maior que o limite, o dado é considerado anômalo. Esta última fase é chamada de Extended Subjective Logic-Based (ESLB) e evita o impacto de vizinhos com dados suspeitos, efetivamente distinguindo os dados com anormalidade dos dados válidos. Isso leva em conta as correlações espaciais históricas do nó e de seus vizinhos.

Foi realizado um experimento com um conjunto de dados do Intel Berkeley Research Lab, contendo dados de temperatura. Com este experimento, Yuan, Zhou e Chen (2012b) chegaram a conclusão de que o modelo SLAD sempre manteve a taxa de detecção acima de 90%, mesmo quando a taxa de falhas aumentava.

### 3.16 COMPREHENSIVE OUTLIER DETECTION (COD) AND FAST COMPREHENSIVE OUTLIER DETECTION (FCOD)

Yao, Cao e Li (2015) apresentaram dois métodos que podem trabalhar com aplicações em tempo real, com o objetivo de fazer a detecção de anomalias nas leituras realizadas por sensores, chamados de Comprehensive Outlier Detection (COD) e Fast Comprehensive Outlier Detection (FCOD). Ambos os métodos são baseados em um método chamado Quarter-Sphere Support Vector Machine (QSSVM) e utilizam como base esferas para identificar se um dado é anômalo ou não, caracterizando um método baseado em clusterização de dados.

Nos métodos apresentados por Yao, Cao e Li (2015), o modelo inicial é construído com a primeira coleta de dados realizada por cada nó sensor. Nesta situação, um dado anômalo é identificado caso ele se situe fora da esfera criada no modelo e, no momento que uma anomalia é identificada, um novo modelo é construído baseado no conjunto de treinamento anterior, caracterizando o uso de correlação temporal.

A diferença entre os dois modelos é a forma pela qual um determinado parâmetro utilizado na construção do modelo é obtido. No COD, tal parâmetro é obtido com um algoritmo chamado Golden Section Search, através do desvio padrão; já o segundo é obtido através do método chamado Optimization Algorithm of QSSVM model base on the Mean of Euclidean Distance (OQMED), que faz o uso da distância Euclidiana. A avaliação dos métodos foi realizada observando valores de taxa de detecção, taxa de falsos positivos e complexidade computacional. Após a realização de todos os experimentos em cada um dos métodos, Yao, Cao e Li (2015) identificaram que o algoritmo FCOD teve uma melhor precisão em relação ao algoritmo COD.

### 3.17 TWO-LEVEL SENSOR FUSION-BASED OUTLIER DETECTION TECHNIQUE

Titouna, Aliouat e Gueroui (2015) apresentaram uma técnica de detecção de *outliers* para aplicações em tempo real que utiliza como base a fusão de dois níveis. A primeira fase é realizada nos nós sensores, onde estes assumem a operação de detecção. A segunda fase ocorre em um *Cluster-Head*, com os sensores subordinados a uma estação base central, na qual ocorre o processamento de dados.

Para a primeira etapa do modelo apresentado, cada sensor classifica determinada medida em *outlier* ou não, usando o classificador baseado em aprendizado de máquina chamado *Naive-Bayes*. Os recursos para o método proposto por Titouna, Aliouat e Gueroui (2015) são compostos pelos dados atuais e pelos últimos dados coletados pelo

sensor e por sua vizinhança. Isso caracteriza o uso de correlação temporal e espacial entre os dados medidos e os sensores, respectivamente. Na segunda etapa, os resultados da decisão fornecidos pela fase anterior são usados para se chegar a um resultado final ao se fundir vários resultados.

Os recursos são compostos pelos dados atuais e pelos últimos dados detectados pelo sensor, os quais permitem explorar a correlação temporal (CT) entre diferentes leituras do sensor e os valores atuais e os últimos detectados e que são gerados por um vizinho. Estes representam a correlação espacial (CE) existente entre diferentes sensores.

Titouna, Aliouat e Gueroui (2015) testaram sua abordagem usando simulações de dados obtidos do Intel Berkeley Research Lab. Foram usados dados de umidade, temperatura, luz e valores de tensão no experimento. Primeiramente, foram feitas medições do sensor de temperatura, nomeando-se esse conjunto de *dataset01*, adicionando cerca de 100 *outliers*. Depois foram utilizados outros dois conjuntos de dados, obtidos a partir do *dataset01*, onde foram adicionados 500 valores de *outliers*. Foi realizado o treinamento e teste de cada conjunto de dados e a simulação foi repetida por 30 vezes. Assim, Titouna, Aliouat e Gueroui (2015) perceberam que em todos os conjuntos de dados foi alcançado cerca de 90% de precisão na detecção de *outliers*.

## 4 DISCUSSÃO

Os métodos apresentados na Seção 3 foram utilizados em suas pesquisas de origem para fazer a validação e também análise de dados coletados por sensores que monitoram um ambiente de forma remota, verificando assim, se os dados são confiáveis. As características de cada método levantado são apresentadas na Tabela 1. Fazendo uma análise generalista dos métodos identificados, foi possível notar que aqueles que são baseados em Artificial Neural Network (ANN) ou aprendizado de máquina são mais complexos e podem consumir mais tempo. Isso ocorre devido a fase de treinamento. Outro ponto notado na maioria dos métodos identificados, foi que grande parte usa como base um limiar para comparação. As medições que desviam fortemente deste limiar são consideradas anômalas ou *outliers*. Um ponto negativo foi que na maioria dos métodos identificados não existe menção a aplicação dos mesmos em sensores biomédicos.

### 4.1 CARACTERÍSTICAS DOS MÉTODOS IDENTIFICADOS

Para realizar a escolha por um método de validação, alguns pontos foram levados em consideração na etapa de análise. As características avaliadas segundo a Tabela 1 são

a possibilidade de uso do método em aplicações de tempo real (ON) ou alguma aplicação offline (OFF), o uso de correlação temporal (CT) ou correlação espacial (CE), significando que o método utiliza medidas anteriores e/ou comunicação com a vizinhança, respectivamente, se o método foi/poderia ser aplicado em sensores biomédicos (SB), se existe a possibilidade de adaptação do método frente a mudanças repentinas nas medições (A) e por fim, se a abordagem apresentada pelo(s) autor(es) foi baseada em algum tipo de técnica estatística (E) ou em clusterização de dados (CD).

Tabela 1 – Tabela para comparação dos métodos

Métodos (índice da Seção)	ON	OFF	CT	CE	SB	A	E	CD
Seção 3.1	X		X			X		
Seção 3.2	X			X		X		
Seção 3.3	X					X	X	
Seção 3.4	X			X		X		
Seção 3.5	X		X				X	
Seção 3.6	X			X				X
Seção 3.7	X							X
Seção 3.8		X	X	X		X		X
Seção 3.9	X				X			X
Seção 3.10	X	X		X			X	
Seção 3.11	X			X			X	X
Seção 3.12	X					X		
Seção 3.13	X		X	X		X	X	
Seção 3.14	X					X	X	X
Seção 3.15	X	X		X				
Seção 3.16	X		X				X	X
Seção 3.17	X		X	X				X

A possibilidade do método ser aplicado online foi descartada devido ao fato de a plataforma pretendida como continuidade para este trabalho não ser compatível. Sendo assim, a escolha de um método offline, ou de tempo real que possa trabalhar offline (i.e. embarcado), é mais adequada para este caso.

Outra característica levantada na Tabela 1 foi o uso de correlação temporal ou espacial. Esse fato é importante, pois um sistema de monitoramento remoto que busca validar dados coletados por sensores, pode levar em consideração medidas capturadas anteriormente, permitindo a identificação de alguma mudança que possa acontecer. Por exemplo, a correlação temporal pode identificar se batimentos cardíacos por minuto aumentam ou diminuem. Assim, a escolha de um método que faça o uso de correlação temporal pode ser mais adequada. Já a correlação espacial não seria relevante, visto que será usado somente um equipamento para monitorar o paciente. Sendo assim, não existiria uma vizinhança para que os sensores pudessem trocar informações.

Uma característica relevante seria a SB, a qual denota se a abordagem proposta é capaz de trabalhar na área de saúde, com a possibilidade de validar dados provenientes de sensores biomédicos. Essa característica não foi encontrada com frequência na literatura, o que dificultou a escolha do método a ser usado neste projeto.

A característica E levantada na Tabela 1 se refere ao fato do método proposto pela literatura fazer o uso de alguma técnica estatística no processo de validação dos dados, como por exemplo, uso de desvio padrão, variância, distância Euclidiana etc. A última característica se refere a clusterização de dados (CD). Nesta técnica, o algoritmo de validação dos dados separa anomalias de dados normais. Ambas as características não possuem grande relevância para o escopo desta pesquisa. Sendo assim, nenhuma delas foi considerada como um fator decisivo na escolha do método.

#### 4.2 CRITÉRIO DE SELEÇÃO

Levando em consideração as características destacadas na Seção 4.1, foi possível selecionar o método mais adequado para o cenário pretendido. Tentando buscar um método que permita atender aos pontos identificados (OFF, tempo real embarcado, CT, SB) e que, ao mesmo tempo, não possua complexidade elevada, optou-se pelo uso do método “*Algorithm for data validation*” apresentado na Seção 3.5.

Vale observar que o modelo apresentado na seção 3.5 pode não se adaptar bem a mudanças repentinas no sinal medido. Para que as características sejam atendidas, algumas alterações foram feitas no modelo selecionado, para que o mesmo trabalhe adequadamente e consiga suprir as necessidades da aplicação considerando os dados provenientes de um sensor. Com isso, foi incorporado o método TOD (apresentado na Seção 3.13) como adaptação ao modelo a ser desenvolvido. Essa técnica de adaptação funciona juntamente com a correlação temporal presente no modelo escolhido. A adaptação é reconhecida e classificada como falha quando somente uma medida se distancia fortemente do padrão. Sendo assim, caso uma medida desvie do padrão, o algoritmo analisará se as próximas medidas também desviarão. Caso o desvio continue acontecendo, essa mudança não será classificada como falha, mas como algum evento ocorrido no ambiente de monitoramento. Caso o desvio não ocorra nas próximas medidas, a mudança acontecida será classificada como uma falha.

Para que o método trabalhe de forma offline, foi necessário também ajustar a entrada de dados. Ao contrário da versão original, onde os dados coletados por sensores são enviados diretamente ao algoritmo, no modelo implementado, a entrada de dados se

deu por meio de um arquivo de texto, onde estavam presentes os dados obtidos na coleta de dados dos voluntários.

Por fim, a escolha do método apresentado na seção 3.5 se deve ao fato deste trabalhar com correlação temporal e também por ser possível adicionar a funcionalidade de trabalhar com adaptação a mudanças repentinas. Com essas funcionalidades adicionadas, o método atende a maioria das características identificadas: correlação temporal, adaptação a mudanças repentinas, que é um fator importante para a análise de alguns sensores como o ECG, além de possuir uma baixa complexidade de implementação. Um ponto negativo é que Ravichandran e Arulappan (2013) não mencionam se este método é aplicável em sensores biomédicos.

De acordo com Guo, Wang e Chen (2015), o método apresentado na Seção 3.9 pode ser aplicado em sensores biomédicos. O motivo deste método não ter sido cogitado para a aplicação no contexto deste trabalho se deve ao fato de o mesmo ser baseado em aprendizado de máquina. Tendo isso em vista, a aplicação do método seria mais complexa para possíveis aplicações em tempo real.

## 5 CONCLUSÃO

Levando em consideração o levantamento realizado para este trabalho, notou-se que o monitoramento remoto de pacientes tem relevância e mostra avanços significativos. Com o avanço da computação ubíqua, esse tipo de monitoramento pode ser cada vez mais utilizado e se tornar mais natural ao paciente.

As questões de pesquisa foram respondidas graças a revisão bibliográfica realizada e os resultados pretendidos foram alcançados. Com isso, foi possível responder a questão “Quais são os métodos para validação de sensores existentes?” na Seção 3, onde são apresentados alguns métodos de validação encontrados na literatura. Já a questão “Os métodos são aplicáveis a sensores biomédicos?” foi brevemente discutida na análise apresentada na Seção 4.

Este trabalho teve como objetivo identificar na literatura métodos capazes de fazer análise e validação de dados provenientes de sensores. Nos métodos encontrados na literatura (Seção 3 e discussão na Seção 4) foi possível perceber que eles são aplicados a diversas áreas, como no setor industrial, experimentos químicos, etc.

Com relação à hipótese destacada, pode-se considerá-la como válida. Fazendo uma análise de cada método identificado (Seção 4) foi possível levantar as características de cada modelo apresentado pela literatura (Tabela 1) e também identificar um modelo a

ser aplicado no contexto deste trabalho. O modelo desenvolvido foi uma adaptação do método apresentado na seção 3.5 (*algorithm for data validation*). Como trabalho futuro pretende-se realizar a aplicação do método resultante deste trabalho a uma plataforma experimental, com o intuito de se validar os dados coletados a partir de sensores (ex. ECG, EMG, GSR etc). Para se verificar a validade da plataforma experimental, cada sensor validado será comparado a uma contraparte certificada para uso clínico.

## REFERÊNCIAS

BASU, S.; MECKESHEIMER, M. Automatic outlier detection for time series: an application to sensor data. *Knowledge and Information Systems*, Springer, v. 11, n. 2, p. 137–154, 2007.

Disponível em: <<https://link.springer.com/article/10.1007/s10115-006-0026-6>>. Acesso em: 15 set. 2021.

CARVALHO, S. T.; ERTHAL, M.; MARELI, D.; SZTAJNBERG, A.; COPETTI, A.; LOQUES, O.; JANEIRO-RJ-BRASIL, R. de. Monitoramento remoto de pacientes em ambiente domiciliar. XXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos-Salao de Ferramentas, Gramado, RS, Brasil, p. 1005–1012, 2010. Disponível em: <[http://sbrc2010.inf.ufrgs.br/anais/data/pdf/salao/st01\\_03\\_salao.pdf](http://sbrc2010.inf.ufrgs.br/anais/data/pdf/salao/st01_03_salao.pdf)>. Acesso em: 15 set. 2021.

CHEN, Y.; JIANG, S.; YANG, J.; SONG, K.; WANG, Q. Grey bootstrap method for data validation and dynamic uncertainty estimation of self-validating multifunctional sensors. *Chemometrics and Intelligent Laboratory Systems*, Elsevier, v. 146, p. 63–76, 2015. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0169743915001161>>. Acesso em: 15 set. 2021.

FAWZY, A.; MOKHTAR, H. M.; HEGAZY, O. Outliers detection and classification in wireless sensor networks. *Egyptian Informatics Journal*, Elsevier, v. 14, n. 2, p. 157–164, 2013. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1110866513000224>>. Acesso em: 15 set. 2021.

FENG, J.; TURKSOY, K.; SAMADI, S.; HAJIZADEH, I.; LITTLEJOHN, E.; CINAR, A. Hybrid online sensor error detection and functional redundancy for systems with time-varying parameters. *Journal of Process Control*, Elsevier BV, v. 60, p. 115–127, dez. 2017. Disponível em: <<https://doi.org/10.1016/j.jprocont.2017.04.004>>. Acesso em: 15 set. 2021.

GHADEMIGOL, M.; GHAEMI-BAFGHI, A.; YAGHMAEE-MOGHADDAM, M. H.; GHORBEL, O.; AYEDI, W.; SNOUSSI, H.; ABID, M. Fast and efficient outlier detection method in wireless sensor networks. *IEEE sensors journal*, IEEE, v. 15, n. 6, p. 3403–3411, 2015. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/7004035/>>. Acesso em: 15 set. 2021.

GUO, X.; WANG, D.; CHEN, F. An anomaly detection based on data fusion algorithm in wireless sensor networks. *International Journal of Distributed Sensor Networks*, SAGE Publications Sage UK: London, England, v. 11, n. 5, p. 943532, 2015. Disponível em: <<https://journals.sagepub.com/doi/full/10.1155/2015/943532>>. Acesso em: 15 set. 2021.

LEVAC, D.; COLQUHOUN, H.; O'BRIEN, K. K. Scoping studies: advancing the methodology. *Implementation science*, BioMed Central, v. 5, n. 1, p. 69, 2010. Disponível em: <<https://implementationscience.biomedcentral.com/track/pdf/10.1186/1748-5908-5-69>>. Acesso em: 15 set. 2021.

MORRISSEY, J. Remote patient monitoring: how mobile devices will curb chronic conditions. *Med Econ*, 2014. Disponível em: <<https://lifecaresolutions.resideo.com/wp-content/uploads/2014/08/Medical-Economics-Remote-Patient-Monitoring.pdf>>. Acesso em: 15 set. 2021.

MOSHTAGHI, M.; LECKIE, C.; KARUNASEKERA, S.; RAJASEGARAR, S. An adaptive elliptical anomaly detection model for wireless sensor networks. *Computer Networks*, Elsevier BV, v. 64, p. 195–207, maio 2014. Disponível em: <<https://doi.org/10.1016/j.comnet.2014.02.004>>. Acesso em: 15 set. 2021.

NISHA, B. U.; UMAMAHESWARI, N.; VENKATESH, R.; ABDULLAH, R. Statistical based outlier detection in data aggregation for wireless sensor networks. *Journal of Theoretical and Applied Information Technology*, v. 59, p. 770–780, 01 2014. Disponível em: <<http://www.jatit.org/volumes/Vol59No3/27Vol59No3.pdf>>. Acesso em: 15 set. 2021.

RAVICHANDRAN, J.; ARULAPPAN, A. I. Data validation algorithm for wireless sensor networks. *International Journal of Distributed Sensor Networks*, SAGE Publications Sage UK: London, England, v. 9, n. 12, p. 634278, 2013. Disponível em: <<https://journals.sagepub.com/doi/full/10.1155/2013/634278>>. Acesso em: 15 set. 2021.

SHEN, Z.; WANG, Q. Failure detection, isolation, and recovery of multifunctional self-validating sensor. *IEEE Transactions on Instrumentation and Measurement*, IEEE, v. 61, n. 12, p. 3351–3362, 2012. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/6236157>>.

Acesso em: 15 set. 2021.

SHEN, Z.; WANG, Q. Data validation and validated uncertainty estimation of multifunctional self-validating sensors. *IEEE Transactions on Instrumentation and Measurement*, IEEE, v. 62, n. 7, p. 2082–2092, 2013. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/6506939>>. Acesso em: 15 set. 2021.

SHEN, Z.; WANG, Q. Status self-validation of a multifunctional sensor using a multivariate relevance vector machine and predictive filters. *Measurement Science and Technology*, IOP Publishing, v. 24, n. 3, p. 035103, 2013. Disponível em: <<https://iopscience.iop.org/article/10.1088/0957-0233/24/3/035103/meta>>. Acesso em: 15 set. 2021.

1088/0957-0233/24/3/035103/meta>. Acesso em: 15 set. 2021.

TITOUNA, C.; ALIOUAT, M.; GUEROUI, M. Outlier detection approach using bayes classifiers in wireless sensor networks. *Wireless Personal Communications*, Springer Nature,

v. 85, n. 3, p. 1009–1023, jun. 2015. Disponível em: <<https://doi.org/10.1007/s11277-015-2822-3>>. Acesso em: 15 set. 2021.

YANG, P.; STANKEVICIUS, D.; MAROZAS, V.; DENG, Z.; LIU, E.; LUKOSEVICIUS, A.;  
DONG, F.; XU, L.; MIN, G. Lifelogging data validation model for internet of things enabled personalized healthcare. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, IEEE, v. 48, n. 1, p. 50–64, 2018. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7516690>>. Acesso em: 15 set. 2021.

YAO, H.; CAO, H.; LI, J. Comprehensive outlier detection in wireless sensor network with fast

optimization algorithm of classification model. *International Journal of Distributed Sensor Networks*, SAGE Publications, v. 11, n. 7, p. 398761, jan. 2015. Disponível em: <<https://doi.org/10.1155/2015/398761>>. Acesso em: 15 set. 2021.

YUAN, J.; ZHOU, H.; CHEN, H. Subjective logic-based anomaly detection framework in wireless sensor networks. *International Journal of Distributed Sensor Networks*, SAGE Publications Sage UK: London, England, v. 8, n. 1, p. 482191, 2012. Disponível em: <<https://journals.sagepub.com/doi/pdf/10.1155/2012/482191>>. Acesso em: 15 set. 2021.

YUAN, J.; ZHOU, H.; CHEN, H. Subjective logic-based anomaly detection framework in wireless sensor networks. *International Journal of Distributed Sensor Networks*, SAGE Publications, v. 8, n. 1, p. 482191, jan. 2012. Disponível em: <<https://doi.org/10.1155/2012/482191>>. Acesso em: 15 set. 2021.

ZHANG, Y.; HAMM, N.; MERATNIA, N.; STEIN, A.; VOORT, M. van de; HAVINGA,

P. Statistics-based outlier detection for wireless sensor networks. *International Journal of Geographical Information Science*, Informa UK Limited, v. 26, n. 8, p. 1373–1392, ago. 2012.

Disponível em: <<https://doi.org/10.1080/13658816.2012.654493>>. Acesso em: 15 set. 2021.