Detecção de *Outliers* em Dados Multivariados em Redes de Sensores Sem Fio Disciplina de Projeto e Análise de Algoritmos

Fabrício G. Valadares, André L. L. Aquino, Álvaro R. Pereira Jr.
PPGCC - Programa de Pós-Graduação em Ciência da Computação
UFOP - Universidade Federal de Ouro Preto
Ouro Preto, Minas Gerais, Brasil
email: {fabricio.valadares, alla.lins}@gmail.com, alvarorpj@yahoo.com

***Resumo*—No campo das redes de sensores sem fio, algumas medições podem desviar do padrão de um conjunto de dados monitorado. Esses desvios são conhecidos como *outliers*, e podem ser causados por ruídos ou erros, ataques maliciosos ou um evento. Apresentaremos neste artigo um comparativo entre alguns algoritmos para detecção de *outliers* multivariados, aplicados à detecção de eventos em redes de sensores sem fio. Serão realizadas análises de complexidade de tempo e espaço, considerando-se as restrições de uma rede de sensores.**

***Keywords*-Redes de sensores sem fio, detecção de eventos, *outliers*, algoritmos**

I. Introdução

O mundo ao nosso redor possui uma variedade de fenômenos que podem ser descritos por algumas grandezas como temperatura, pressão e umidade, e podem ser monitorados por dispositivos com poder de sensoriamento, processamento e comunicação [1]. Trabalhando de forma cooperativa, esses dispositivos constituem as redes de sensores sem fio [2], [3], [4], que são um tipo especial de rede *ad hoc* [5], formadas por elementos conhecidos como nós sensores. Outros elementos que também fazem parte das redes de sensores são os nós atuadores, que possuem a capacidade de interferir no ambiente monitorado, e podem ser fisicamente os próprios nós. Além deles, temos também os sorvedouros, que possuem capacidade computacional superior a dos demais nós, e são responsáveis por receber e processar as informações provenientes do sensoriamento. Existem também os *gateways*, dispositivos responsáveis por conectar a rede de sensores a outras redes, como a Internet. Os *gateways* e sorvedouros podem ser o mesmo elemento.

Os nós sensores são dispositivos compactos e autônomos, com capacidade de processamento local e sensoriamento. Devido ao seu tamanho reduzido, possuem uma série de restrições, como, capacidade computacional limitada, largura de banda, memória e principalmente energia, que é fornecida por bateria. Como esses dispositivos são geralmente depositados em áreas remotas e/ou de difícil acesso, a substituição da bateria é normalmente inviável. De acordo com [2], [6], o rádio é o componente de maior consumo energético, muito superior ao processamento. Desta forma, o envio de grandes quantidades de dados pode reduzir a vida útil de uma rede de sensores. Conforme pesquisa citada em [3], o consumo energético para a execução de 3000 instruções é equivalente ao envio de um *bit* por 100 metros via rádio. Sendo assim, devemos favorecer o pré-processamento dos dados diretamente na rede.

O principal objetivo de uma rede de sensores sem fio é o envio de informações a um observador externo. Porém, a transmissão de dados está sujeita a interferências do ambiente, falhas na comunicação ou ruídos, o que pode causar a perda de valores, dados duplicados ou inconsistência nas informações. Para muitas aplicações, como por exemplo, detecção de incêndio em florestas e sensores de inundação, não é possível realizar a tomada de decisões baseadas em dados incorretos. Estas anomalias podem ser encontradas na literatura como *outliers*, termo que tem origem na estatística [7].

Em [7], encontramos algumas definições para os *outliers*. Definições clássicas: Segundo Hawkins [8], "*um outlier é uma observação, que desvia muito de outras observações que despertam suspeitas de que são gerados por um mecanismo diferente.*" Segundo Barnett and Lewis [9], "*um outlier é uma observação (ou um subconjunto de observações) que parece ser inconsistente comparado ao restante do conjunto de dados.*" No campo das redes de sensores, podemos definir os *outliers* como as medições que desviam significativamente do padrão normal dos dados sensoriados.

Existem três fontes de *outliers* nas redes de sensores, os ruídos e os erros, os ataques maliciosos e os eventos [7]. Neste trabalho estamos interessados na detecção de eventos, que segundo [2], é a principal tarefa de um nó no campo de monitoramento. Esses eventos são detectados a partir dos dados do ambiente, que são coletados por sensores presentes em cada nó. É importante destacar que os nós nas redes de sensores sem fio podem possuir um ou mais tipos de sensores, o que permite monitorar um ou mais tipos de fenômenos. Desta forma, os dados que trafegam pela rede podem ser classificados como univariados ou multivariados [10].

Dados univariados representam um único conjunto de valores de um mesmo tipo de fenômeno, como por exemplo, um nó com capacidade para monitorar apenas a temperatura

ambiente. Por sua vez, os dados multivariados são representados por um conjunto de valores de diferentes tipos, como os gerados por um nó que possui sensores que monitoram temperatura, pressão e umidade simultaneamente, ou ainda os dados recebidos por um nó, que é o responsável pelo processamento dos dados enviados por um conjunto de nós capazes de monitorar um único tipo de fenômeno, como temperatura, por exemplo [1]. Neste trabalho estamos interessados na utilização de técnicas para detecção de *outliers* multivariados aplicados à identificação de eventos em redes de sensores sem fio.

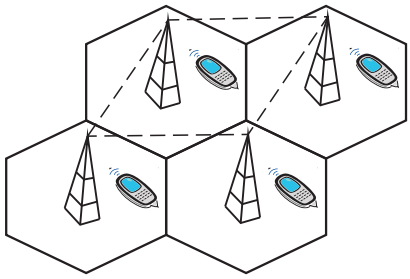
As técnicas para detecção de *outliers* já são utilizadas em diversas áreas de pesquisa, como aprendizagem de máquina, mineração de dados, estatística, teoria da informação, entre outras, onde são aplicadas por exemplo, para detecção de intrusos e identificação de fraudes [11]. Conforme relatado em [7], apesar destas técnicas ainda não serem encontradas na literatura de detecção de eventos, elas podem ser adaptadas para tal fim. Neste trabalho realizaremos a análise de alguns algoritmos para detecção de *outliers* multivariados que podem ser adequados para a detecção de eventos em redes de sensores sem fio. Também avaliaremos sua complexidade de tempo considerando a fragilidade dos nós sensores. Os algoritmos identificados serão implementados no *software* estatístico *R*, para realizar uma comparação entre as técnicas.

O artigo está organizado como segue. A Seção II apresenta as redes de sensores sem fio. Os *outliers* são detalhados na Seção III. Os métodos utilizados são demostrados na Seção IV. Já a Seção V discorre sobre a complexidade dos algoritmos. As Seções VI e VII ilustram respectivamente os testes e as conclusões.

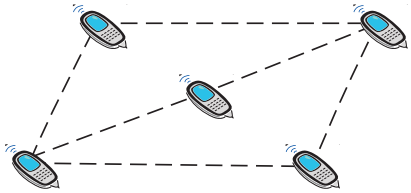
II. Redes de Sensores Sem Fio

Para contextualização das redes de sensores no ambiente sem fio, precisamos considerar as redes estruturadas e as redes *ad-hoc*, ilustradas na Figura 1, adaptada de [1]. Conforme descrito por [6], nas redes estruturadas, existe uma estação base responsável pela comunicação entre os nós da rede (Figura 1(a)). Porém, nas redes *ad-hoc* (Figura 1(b)), a comunicação é realizada através de nós que se encontram entre a origem e o destino e, desta forma, a estação base não é necessária. As redes de sensores sem fio (Figura 1(c)), como as redes *ad-hoc*, não precisam de estação base para que dois nós se comuniquem.

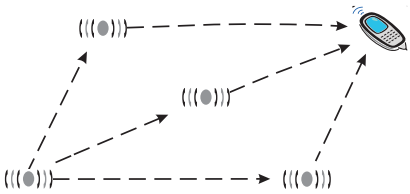
Conforme [2], as redes de sensores sem fio se diferem das redes *ad-hoc* por possuírem um número elevado de nós, na ordem de centenas ou milhares, que geralmente se mantém estacionários (imóveis) após a deposição. Também se diferem por uma frequente mudança na topologia da rede, que pode ser causada pelo desligamento do rádio para economia de energia ou danos físicos nos nós durante o sensoriamento. Além destas características, podemos ainda citar a alta densidade destas redes, como por exemplo, 21



(a) Redes estruturadas



(b) Redes *ad-hoc*



(c) Redes de sensores

Figura 1. Tipos de redes sem fio

vizinhos por nó [12], e a comunicação centrada em dados, podendo os nós não possuírem identificadores únicos, como o IPv6, uma vez que estes exigem um espaço (*bits*) para identificar cada nó, aumentando o tamanho da mensagem e também o custo de comunicação [13].

Os nós sensores são dispositivos compactos e autônomos, tem arquitetura simples e são compostos por quatro componentes básicos, que estão ilustrados na Figura 2, adaptada de [2]. Estes componentes são: 1) a unidade perceptiva, com um ou mais tipos de sensores e um conversor de sinais analógicos para digitais (*Analog-to-Digital Converter* - ADC); 2) uma unidade de processamento, com processador e memória; 3) um transceptor, responsável por conectar o nó à rede e 4) a fonte de energia, geralmente uma bateria não renovável. Além dos componentes citados, podem existir outros, dependendo da aplicação, como sistema de localização, mecanismo para mobilidade e um gerador de energia [2].

A deposição dos nós pode ser feita muito próxima ou dentro do fenômeno a ser observado, e sua posição não precisa ser previamente planejada, o que permite monitorar ambientes hostis e de difícil acesso. As redes de sensores sem fio devem, portanto, ser autoconfiguráveis, adaptáveis e com gerenciamento escalável [2]. Essas características possibilitam sua utilização em uma grande variedade de aplicações, como médicas, militares, domésticas, industriais,

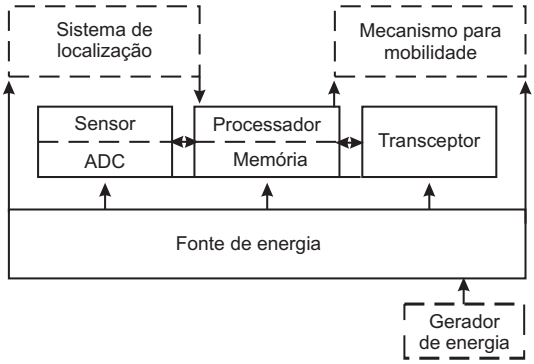


Figura 2. Componentes de um nó sensor

meio ambiente, agropecuária, robótica, dentre outras [2], [4].

De acordo com [3], [1], as redes de sensores sem fio podem receber diferentes classificações, a partir da observação de seus elementos básicos, da disposição dos nós na área de sensoriamento e também da maneira como os fenômenos são monitorados. A rede é considerada hierárquica se existirem agrupamentos de nós e um nó líder os representar, caso contrário, será definida como plana. É homogênea se os nós possuírem a mesma configuração de *hardware*, senão, é heterogênea. A rede é simétrica se todos os nós tiverem o mesmo raio de comunicação, caso contrário, é assimétrica. Quando os dados são enviados obedecendo a uma programação pré-estabelecida, a rede é chamada programada, por outro lado, se os dados forem enviados continuamente a rede é contínua. Por último, as redes de sensores podem ser classificadas como dirigida a eventos, se o envio de dados acontecer somente na ocorrência de algum evento, ou sob demanda, quando a rede permite a consulta total ou parcial dos dados a qualquer momento.

Loureiro *et al.* [6] reúne as funcionalidades das redes de sensores em cinco grupos de atividades: (I) o estabelecimento, no qual acontecem as configurações iniciais da rede; (II) a manutenção, na qual são realizadas adaptações às mudanças de configuração ocorridas durante todo o tempo de vida da rede; (III) o sensoriamento, responsável pela coleta de dados do campo monitorado; (IV) o processamento dos dados coletados e que serão enviados ao sorvedouro, e, por fim (V) a comunicação, responsável pela transmissão dos dados processados.

III. *Outliers*

Conjuntos de dados, sejam eles grandes ou pequenos, podem conter elementos que não condizem com a distribuição do restante do conjunto, ou seja, podem existir pontos que se destacam quanto à uma ou mais variáveis, prejudicando a modelagem estatística dos valores analisados [14]. Essas anomalias no conjunto de dados podem receber o nome de *outliers*, termo derivado da estatística.

Como citado na Seção I, existem duas definições clássicas para os *outliers*. Segundo Hawkins [8], "*um outlier é uma observação, que desvia muito de outras observações que despertam suspeitas que são gerados por um mecanismo diferente.*" Segundo Barnett & Lewis [9], "*um outlier é uma observação (ou um subconjunto de observações) que parece ser inconsistente comparado ao restante do conjunto de dados.*" Ainda existe uma variedade de definições, dependendo de algum método particular de detecção de *outliers*, e também do tipo específico de dados do conjunto. No campo das redes de sensores sem fio, podemos definir os *outliers* como as medições que desviam significativamente do padrão normal dos dados sensoriados.

Figura 3. Fontes de *outliers* em Redes de Sensores Sem Fio

A detecção de *outliers* tem sido extensivamente pesquisada em várias disciplinas, como a estatística, mineração de dados, aprendizagem de máquina e teoria da informação. Nessas áreas, a detecção de anomalias é utilizada para identificar fraudes, intrusão de rede, análise de performance, previsão do tempo entre outras [7]. Porém apenas recentemente ele tem atraído a atenção em pesquisas relacionadas às redes de sensores sem fio.

Conforme citado por [7], e ilustrado na Figura 3, podemos considerar três fontes de *outliers* nas redes de sensores sem fio, que são os ruídos ou erros, eventos e ataques maliciosos. A detecção destas anomalias pode melhorar a confiança da rede, reportar eventos, além de melhorar a qualidade dos dados monitorados, tornando a análise de dados robusta, mesmo sob a presença de erros ou ruídos na rede. Nesse trabalho, estamos interessados na detecção de eventos, que segundo [2], é a principal tarefa de um nó no campo de monitoramento.

De acordo com [7], existem algumas diferenças entre as técnicas tradicionais para a detecção de eventos em redes de sensores e as técnicas para detecção de *outliers*. A primeira delas é referente a uma condição gatilho. Nas técnicas para detecção de anomalias não existe conhecimento *a priori* sobre estas condições, já as técnicas para detecção de eventos sim. A segunda diferença também está relacionada à condição gatilho, uma vez que as técnicas para detecção de eventos comparam os dados monitorados com estas condições pré-definidas. Por outro lado, as técnicas para detecção de *outliers* comparam as medições entre os sensores em si. Por último, as técnicas para detecção de *outliers* precisam evitar que dados normais sejam classificados como anomalias, mantendo a taxa de detecção elevada, e a taxa de falsos alarmes baixa. Já as técnicas para detecção de eventos precisam evitar que dados incorretos

sejam classificados como eventos, ou que padrões possam influenciar a confiança da detecção.

Em [7], são definidos critérios de classificação de técnicas para detecção de *outliers* em redes de sensores sem fio. Estes critérios são o tipo de entrada de dado, a correlação entre os dados de entrada, os tipos de *outliers*, que podem ser locais ou globais. Além disso, temos ainda as fontes de *outliers*, que podem ser erros, eventos ou ataques maliciosos, e a disponibilidade de dados pré definidos como normais. Essa última característica define a categoria em que as técnicas para a detecção de anomalias se enquadram, e a partir delas, foi criado uma taxonomia para as técnicas de detecção de *outliers* em redes de sensores sem fio, disponível em [7].

Como estamos interessados em dados multivariados, optamos por trabalhar com técnicas baseadas em estatística, e que utilizam uma distribuição normal dos dados. O motivo desta escolha ficará claro na próxima seção.

IV. Métodos Utilizados

Conforme citado na Seção III, neste trabalho estamos interessados na detecção de *outliers* multivariados, cujas técnicas serão utilizadas na identificação de eventos em redes de sensores sem fio. Porém, antes de apresentarmos as técnicas aqui utilizadas, precisamos fazer algumas considerações.

De acordo com [14], em um cenário onde o conjunto de dados é univariado, a identificação de anomalias pode ser facilmente realizada através de um gráfico de dispersão, que indica a distância de uma amostra aos demais elementos do grupo. Além desse gráfico, pode-se utilizar também a distância Euclidiana do $i-$ésimo elemento da amostra à média do conjunto de dados. Essa distância é dada pela equação

$$ED_i = \sqrt{(x_i - \overline{x})^2}, \quad (1)$$

onde x_i é o $i-$ésimo elemento do conjunto de entrada X, $\overline{x}$ é a média amostral de X e ED_i é a distância Euclidiana de cada elemento à média do conjunto.

Ao se considerar um conjunto de dados com mais de uma variável, podemos utilizar a distância de Mahalanobis, que identificará possíveis anomalias no conjunto [14], *i.e.*,

$$MD_i = \sqrt{(x - \overline{x})^T S_x^{-1} (x_i - \overline{x})}, \quad (2)$$

onde, $\overline{x}$, de dimensão p, é o vetor de médias do conjunto X e $S_x = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})(x_i - \overline{x})^T$ é a matriz $p \times p$ de covariância amostral de X. Os dados de entrada são representados pela variável X, uma matriz de $p \times n$, onde p indica o número de variáveis e n o número de amostras.

Observe que tanto a distância Euclidiana quanto a distância de Mahalanobis são afetadas pelos *outliers*, pois utilizam a média amostral do conjunto. Quando estamos trabalhando entradas que contém poucas variáveis, é possível realizar o cálculo da média, removendo-se o elemento que é atualmente avaliado, porém quando o número de variáveis cresce, este procedimento torna-se inviável. Desta forma, precisamos de métodos robustos para realizar a detecção dos *outliers* em conjuntos de dados multivariados.

Em seu trabalho, Giroldo [14] apresenta quatro técnicas para detecção de *outliers* multivariados, os métodos baseados em estimadores robustos MVE - *Minimum Volume Elipsoid* e MCD - *Minimum Covariance Determinant* e os métodos iterativos MED - *Max-eigen Difference* e *Fast Forward*. Com exceção do MED, todas as outras técnicas utilizam a distância de Mahalanobis, para a identificação dos *outliers* em um conjunto de dados, porém, a média e a matriz de covariância são substituídos por estimadores robustos.

Neste trabalho, implementamos as técnicas MVE, MCD e MED, utilizando os algoritmos disponíveis em [14].] O *Fast Forward* não foi implementado, pois a biblioteca para o *software* R não está mais disponível. A descrição dos métodos será realizada nas próximas sub-seções. Vale ressaltar, que as técnicas baseadas no MCD e MVE identificam os *outliers* através da distância de Mahalanobis Robusta, o que requer uma distribuição aproximadamente normal dos dados. Isso é necessário, pois se a amostra de dados tiver uma distribuição aproximadamente normal, então, a distância de Mahalanobis ao quadrado (MD_i^2) possui uma distribuição aproximadamente qui-quadrado, com p graus de liberdade ($\mathcal{X}_p^2$), onde p indica o número de variáveis da amostra. Podemos então classificar um elemento i como atípico, comparando-se o quadrado de sua distância de Mahalanobis com um quantil da distribuição qui-quadrado escolhido.

A. Minimum Volume Elipsoid

O método baseado em MVE (*Minimum Volume Elipsoid*), procura tornar a distância de Mahalanobis robusta, substituindo a média e a matriz de covariância por estimadores que sofram menor interferência das anomalias. Estes estimadores são $T(X)$ e $C(X)$, respectivamente um vetor de tamanho p e uma matriz de tamanho $p \times p$, onde $T(X)$ é o centro do elipsoide e $C(X)$ é a estrutura de covariância do mesmo elipsoide [14].

B. Minimum Covariance Determinant

O método baseado em MCD (*Minimum Covariance Determinant*) também torna a distância de Mahalanobis um método robusto, e assim como no MVE, são calculados estimadores $T(X)$ e $C(X)$, que permitem traçar um elipsoide, e um limiar, a partir de onde as amostras serão consideradas *outlires* [14].

C. Max-Eigen Distance

O MED, ao contrário das outras técnicas, não utiliza a distância de Mahalanobis para identificar os *outliers*. Esta detecção é realizada através dos autovetores e autovalores da matriz de covariância, aos quais são aplicados a norma euclidiana [14].

V. Análise de Complexidade

Conforme citado em [15], a técnica MCD é NP-Completo, reduzido a partir do problema do clique em grafos, porém se a quantidade de variáveis for mantida fixa, pode-se conseguir um algoritmo com complexidade exponencial. A mesma complexidade pode ser atribuída ao MVE, conforme [15]. Estimamos que a complexidade do MED seja $O(n^3)$, onde k representa a quantidade de elementos de uma matriz $n \times p$.

VI. Testes

O teste realizado visa comparar comparar a taxa de detecção de *outliers* entre as técnicas aplicadas. A entrada destes algoritmos é uma matriz X, de dimensões $p \times n$ que representam respectivamente a quantidade de variáveis e o número de amostras. Os dados foram gerados a partir da média 25 e desvio padrão de 20%, ou seja 5. Essa média e desvio padrão são aplicadas às 5 colunas de X, e foram consideradas 64 amostras. Todos estes valores foram escolhidos de forma arbitrária.

As linhas $62, 63$ e 64 foram alteradas, acrescentando-se dois desvios padrões aos seus valores originais, desta forma pretendemos simular um evento nos dados. Durante o teste, todas as técnicas conseguiram detectar as amostas alteradas, porém apenas o MCD utilizando a função *dd.plot*, que aplica o algoritmo *fast-MCD*, exato para pequenas instâncias [14], localizou apenas os dados alterados. Quando utilizamos as abordagens MVE e MED, além das anomalias inseridas, foram encontradas respectivamente outros 21 e 13 *outliers*. O MCD, através da função *aq.plot*, também localizou duas amostras a mais. Isso já era esperado, pois estes algoritmos são aproximados, mesmo para pequenas instâncias. Vale ressaltar que este teste serve apenas para ilustrar um comparativo entre as técnicas, e que testes exaustivos precisam ser realizados antes da tomada de qualquer decisão.

VII. Conclusão

Neste trabalho foram identificadas algumas técnicas para a detecção de *outliers* multivariados. Os algoritmos analisados são em sua maioria aproximados, devido à complexidade dos cálculos. Os testes realizados não são conclusivos, e serviram como demostração do funcionamento das técnicas. Realizaremos mais estudos para identificar quais algoritmos serão adequados à aplicação nas redes de sensores sem fio, um ambiente bastante restritivo. Acreditamos, que através da utilização das ferramentas para detecção de *outliers*, possamos identificar eventos de forma eficiente e dinâmica. Os próximos passos da pesquisa são um estudo mais profundo das técnicas aqui discriminadas e um estudo de complexidade mais avançado.

Uma observação importante, é que todas as técnicas implementadas neste trabalho geram gráficos como saída, logo, será necessário executar uma adaptação das mesmas antes de aplicá-las às redes de sensores. Não podemos deixar de agradecer à Professora Lúcia Pereira Barroso, do Instituto de Matemática e Estatística da USP, que gentilmente nos forneceu o documento [14], de onde foram extraídos os algoritmos.

Referências

[1] A. L. L. Aquino, "Redução de dados em redes de sensores sem fio baseada em stream de dados," Ph.D. dissertation, Universidade Federal de Minas Gerais, February 2008.

[2] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *Communications Magazine, IEEE*, vol. 40, no. 8, pp. 102 – 114, aug 2002.

[3] S. Tilak, N. B. Abu-Ghazaleh, and W. Heinzelman, "A taxonomy of wireless micro-sensor network models," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 6, pp. 28–36, April 2002. [Online]. Available: http://doi.acm.org/10.1145/565702.565708

[4] T. Arampatzis, J. Lygeros, and S. Manesis, "A survey of applications of wireless sensors and wireless sensor networks," in *Intelligent Control, 2005. Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation*, Hawaii, USA, June 2005, pp. 719–724.

[5] E. Royer and C.-K. Toh, "A review of current routing protocols for ad hoc mobile wireless networks," *Personal Communications, IEEE*, vol. 6, no. 2, pp. 46 –55, apr 1999.

[6] A. A. F. Loureiro, J. M. S. Nogueira, L. B. Ruiz, R. A. de Freitas Mini, E. F. Nakamura, and C. M. S. Figueiredo, "Redes de sensores sem fio (tutorial)," in *XXI Simpósio Brasileiro de Redes de Computadores (SBRC'03)*, 2003, pp. 179–226.

[7] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier detection techniques for wireless sensor networks: A survey," *Communications Surveys Tutorials, IEEE*, vol. 12, no. 2, pp. 159 –170, quarter 2010.

[8] D. Hawkins, "Identification of outiliers," *Capman and Hall*, 1980.

[9] V.Barnett and T. Lewis, "Outliers in statistica data," *John Wiley Sons*, 1994.

[10] O. S. Junior, A. L. L. de Aquino, C. Figueiredo, and R. A. F. Mini, "Multivariate reduction in wireless sensor networks," in *14th IEEE Symposium on Computers and Communications (ISCC'09)*, 2009, pp. 726–729.

[11] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surv.*, vol. 41, pp. 15:1–15:58, July 2009. [Online]. Available: http://doi.acm.org/10.1145/1541880.1541882

[12] L. A. Villas, H. S. Ramos, and A. A. F. Loureiro, "Um algoritmo de roteamento ciente da agregação de dados para redes de sensores sem fio," in *XXVII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC'09)*, 2009, pp. 233–246.

[13] L. B. Ruiz, L. H. A. Correia, L. F. M. Vieira, D. F. Macedo, E. F. Nakamura, C. M. S. Figueiredo, M. A. M. Vieira, E. H. Bechelane, D. Câmara, A. A. F. Loureiro, J. M. S. Nogueira, D. C. da Silva Jr., and A. O. Fernandes, "Arquiteturas para redes de sensores sem fio," in *XXII Simpósio Brasileiro de Redes de Computadores*, 2004.

[14] F. R. de Santana Giroldo and L. P. Barroso, "Alguns métodos robustos para detectar outliers multivariados," Master's thesis, Instituto de Matemática e Estatística - Universidade de São Paulo - IME-USP, 2008.

[15] T. Bernholt, L. Informatik, and P. Fischer, "The complexity of computing the mcd-estimator," *Theoretical Computer Science*, vol. 326, pp. 383–398, 2004.

APÊNDICE

Todos os arquivos de testes, assim como os algoritmos estão disponíveis no endereço https://sites.google.com/site/fabricioppgccufop/outliers.

Agradecimentos à Polly, minha esposa, pelo auxílio com a correção do texto.