

Mitigação de Ataques IDFs no Serviço de Agrupamento de Disseminação de Dados em Redes IoT Densas

Carlos Pedroso¹, Fernando Gielow¹, Michele Nogueira¹, Aldri Santos¹

¹Núcleo de Redes Sem-Fio e Redes Avançadas (NR2) – UFPR
Centro de Ciência de Segurança Computacional (CCSC) – UFPR
Caixa Postal 19.081 – 81.531-980 – Curitiba – PR – Brasil

{capjunior, aldri, fhgielow, michele}@inf.ufpr.br

Abstract. *The growth of IoT has made possible the creation of increasingly personalized services, among them, the industrial services stand out. But as IoT grows, so does threat. Among these threats, the false data injection attack stands out as being one of the most aggressive to data networks such as IoT. Although there are several mechanisms that deal with it they do not take into account the validation of the data. This work proposes an intrusion detection mechanism against false data injection attacks on the service of dissemination of dense IoT. The mechanism called CONFINIT uses similarity groupings to organize the network and identify threats. It combines Watchdog strategies and Collaborative Consensus for attacker detection, ensuring the authenticity of data collected by devices. CONFINIT was evaluated in the simulator ns-3, showing its effectiveness in attack mitigation and low interference in the network.*

Resumo. *O crescimento da IoT vem possibilitando a criação de serviços cada vez mais personalizados, entre eles, destaca-se os serviços industriais. Porém a medida que a IoT cresce as ameaças também vem crescendo. Entre essas ameaças está o ataque de injeção de dados falsos que destaca-se por ser um dos mais agressivos às redes de dados como a IoT. Embora existam diversos mecanismos que lidam com a ameaça eles não levam em consideração a validação dos dados. Este trabalho propõe um mecanismo de detecção de intrusão contra ataques de injeção de dados falsos sobre o serviço de disseminação da IoT densa. O mecanismo chamado CONFINIT, utiliza agrupamentos por similaridade para organizar a rede e identificação de ameaças. Ele combina as estratégias de Watchdog e Consenso Colaborativo para detecção de atacantes, garantindo a autenticidade dos dados coletados pelos dispositivos. O CONFINIT foi avaliado no simulador ns-3, mostrando sua eficácia na mitigação do ataque e baixa interferência na rede.*

1. Introdução

A Internet das coisas (IoT) possibilita a conexão de diferentes tipos de objetos físicos, através de tecnologias como redes de sensores sem fio (RSSF), RFID, GPS e NFC, entre outras. Os objetos IoT possuem várias características como identidade, atributos físicos, heterogeneidade, e muitos deles usam interfaces inteligentes para estabelecerem comunicações entre si, além de poder possuir alguma forma de mobilidade [Gubbi et al. 2013]. A IoT faz parte da evolução de domínios densos e complexos como

processos industriais, logística, segurança pública e cidades inteligentes. Nesse contexto, o paradigma da Internet das coisas industriais (IIoT) vem recebendo maior atenção nos últimos anos. A IIoT trata da conexão de diferentes dispositivos dentro de uma fábrica, fazendo com que todos trabalhem de forma sincronizada e organizada. Por tratar-se de uma variação da IoT, a IIoT ainda não está totalmente consolidada, apresentando diversas vulnerabilidades e desafios para sua consolidação [Mumtaz et al. 2017]. A IoT faz-se fundamental para coletar, disseminar e lidar com o volume de dados exigido por diversas aplicações nas suas tomadas de decisões [Minoli et al. 2017, Akpakwu et al. 2018].

Consequentemente, **tratar e disseminar** esse grande volume de dados resultante da interação entre diversos dispositivos expõe a **IoT densa** à diversas vulnerabilidades. Os ambientes normalmente contam com dispositivos móveis e fixos e a infraestrutura varia conforme a interação. Parte dos dispositivos têm recursos limitados, pouca energia, baixa capacidade de processamento e armazenamento, além de sofrerem perdas nos links de conexão [Borgia 2014, Qiu et al. 2018]. Logo, a IoT torna-se alvo de inúmeras ameaças que violam atributos de segurança como, integridade, autenticidade e disponibilidade de serviços, dentre eles a disseminação de dados dentro da rede [Yaqoob et al. 2017, Kouicem et al. 2018], o que prejudica a operação de diversas aplicações [Miorandi et al. 2012, Mendez et al. 2017].

Particularmente, entre as ameaças **internas** ao serviço de disseminação de dados na IoT, destaca-se o ataque de Injeção de Dados Falsos (IDF), considerado um dos ataques **de intrusão** mais nocivo às redes de dados, devido à inconsistência das informações geradas e à imprevisibilidade do seu acontecimento [Sen and Madria 2017]. Em razão da sua complexidade, a detecção do ataque torna-se complexa e trabalhosa, pois normalmente os dispositivos maliciosos estão autenticados na rede e exercem suas funções padrão de coleta e disseminação de dados. Os ataques podem ocorrer em diferentes períodos e de forma contínua, desorientando a rede. O IDF pode ocorrer de duas formas, quando um dispositivo é capturado por outro e tem seus dados fabricados ou alterados, e quando o próprio dispositivo apresenta um comportamento de má conduta. Esse comportamento dificulta a identificação de dispositivos maliciosos, aumenta o tempo de mal funcionamento da rede, gerando inconsistência nos dados. Desta forma, torna-se imprescindível identificar e excluir os dispositivos maliciosos da rede.

Apesar de diferentes técnicas serem empregadas para lidar com os ataques IDF, sejam no contexto de RSSF [Lu et al. 2012, Kumar and Pais 2018], *Smart Grids* [Li et al. 2017] ou IoT [Yang et al. 2017]. Elas têm falhado ou não são adequadas ao contexto de IoT densa, visto que elas geram alto consumo de recursos, não checam os dados e poucas consideram a detecção colaborativa. Dentre as técnicas existentes, as mais empregadas são: esquemas de filtragem em rota, métodos colaborativos e os sistemas de detecção de intrusão (IDS). Os esquemas de filtragem em rotas são constantemente aplicados em RSSF, usam a verificação de relatório em nós intermediários entre origem e destino, descartando pacotes com qualquer tipo de inconsistência. Os relatórios são endossados para garantir sua autenticidade, entretanto eles não consideram alteração dos dados coletados. Os métodos colaborativos são alternativas para lidar com ataques IDF nas *Smart Grids*, onde cada dispositivo desempenha duas funções, as suas funções-padrão e a de agente colaborativo de detecção. Contudo, o consumo de recurso dos dispositivos aumenta consideravelmente. Já os IDS são alternativas robustas para lidar com os ataques

em diferentes contextos, podendo ser baseados em dispositivos ou rede. Entretanto, podem gerar alto consumo de recursos, além de gerar novas vulnerabilidades. Desta forma, torna-se fundamental para o desenvolvimento da IoT mecanismos capazes de detectar e isolar a presença de ameaças de forma distribuída para o serviço de agrupamento e proteger a disseminação de dados.

Este trabalho propõe um mecanismo para mitigação do ataque de Injeção de dados falsos sobre o serviço de disseminação de dados de redes IoT. O mecanismo chamado CONFINIT (*CONsensus Based Data FilteriNg for IoT*), busca detectar e isolar da rede IoT dispositivos maliciosos que apresentem comportamento malicioso ao sistema. O mecanismo utiliza agrupamentos por similaridade para lidar com densidade e mobilidade dos dispositivos. Ele combina as estratégias de Watchdog e Consenso colaborativo para tomada de decisão sobre os dispositivos maliciosos. O CONFINIT foi avaliado por meio de simulações e os resultados comprovam sua eficácia na mitigação do ataque.

O restante do artigo está apresentado da seguinte forma: a Seção 2 apresenta os trabalhos relacionados. A Seção 3 apresenta e descreve o funcionamento do CONFINIT. A Seção 4 descreve o método de avaliação e o desempenho do mecanismo juntamente com resultados obtidos. A Seção 5 conclui o artigo e apresenta direcionamentos futuros.

2. Trabalhos Relacionados

A demanda por um serviço de disseminação de dados seguro em redes densas na presença de diversas formas de ataques de intrusão tem sido o foco de diferentes trabalhos [Lu et al. 2012, Kumar et al. 2016, Yang et al. 2017, Cervantes et al. 2018, Bostami et al. 2019]. Em geral, o uso de Sistemas de Detecção de Intrusão (IDS) viabiliza identificar, localizar e mitigar comportamentos de **má conduta de diversas naturezas** [Kumar et al. 2016]. Particularmente, os ataques maliciosos de IDF sobre o serviço de disseminação de dados nas IoT densas tornam-se danosos à rede devido ao grande volume de dados gerados **pelos dispositivos e da sua importância à tomada de decisões dos objetos** pelas aplicações.

Em [Yang et al. 2017] os autores propuseram um IDS baseado em detecção de anomalias através do uso de *Watchdog* para mitigação do ataque de injeção de dados falsos. Eles tentam prever eventos naturais ao usar dados coletados de vigilância ambiental da IoT através de monitoramento Hierárquico Bayesiano Espacial-Temporal (HBT). Logo após, usa-se uma estratégia de decisão estatística em um teste de probabilidade sequencial para identificar atividades maliciosas. Entretanto, o modelo HBT apresenta alto consumo energético e o teste probabilístico, embora efetivo, sobrecarrega a rede, e não se considera dispositivos móveis. Em [Cervantes et al. 2018], os autores apresentam uma IDS para mitigação de ataques *sinkhole* e *selective forwarding* no roteamento de redes IoT densas, onde a rede organizada em *clusters* classifica os nós em três categorias. Uma estratégia *Watchdog* em níveis, monitora a relação entre encaminhamento e dados recebidos, definindo quais nós apresentam comportamento malicioso. Apesar disso, computar a confiança entre participantes com base na quantidade de dados recebidos e transmitidos não exclui a possibilidade do nó alterar os dados coletados.

Um método usado frequentemente contra os ataques IDF em RSSF são os esquemas de filtragem em rota [Lu et al. 2012, Yu and Guan 2010, Wang et al. 2014]. Em [Lu et al. 2012] os autores propõem um esquema de autenticação cooperativa para filtrar

dados falsos em RSSF, onde cada nó requer um número fixo de vizinhos para gerenciar a autenticação, que baseia-se no roteamento de vizinhos. Os nós são autenticados de forma distribuída ao longo do roteamento dos dados até a estação base. O esquema adota a técnica de bit comprimido visando não sobrecarregar o canal, tornando-o adequado para filtrar dados injetado já que a autenticação ocorre ponto-a-ponto. Além disso, o uso de um protocolo de roteamento evita que dispositivos maliciosos comprometam outros dispositivos. Contudo, a proposta gera uma alta sobrecarga na rede e não identifica os dispositivos comprometidos. Em [Yu and Guan 2010], os autores apresentam um esquema de filtragem dinâmica em rota para lidar com dados falsos e ataque DoS em RSSF. Cada nó tem um conjunto de chaves de autenticação usada para endossar relatórios. A autenticação é garantida por um grupo de nós escolhidos antes da rede funcionar. Cada nó disponibiliza sua chave aos nós encaminhadores que devem divulgar suas chaves, o que permite aos encaminhadores verificar todos os relatórios. Porém, esta solução gera sobrecarga na rede devido a constante troca de chaves entre os dispositivos, limitando o seu funcionamento em redes densas.

Uma maneira de alcançar uma detecção colaborativa é através do uso de *Consensus* entre dispositivos de uma rede [Colistra et al. 2014]. Em [Toulouse et al. 2015] os autores propuseram um sistema distribuído para detecção de anomalias, baseado em um protocolo de consenso médio entre participantes. O sistema busca identificar anomalias que possam gerar ataques DDoS. Para isso são realizadas análises em cada ponto de coleta de dados usando um classificador Bayes. Ao final a análise é realizada de forma redundante, paralelo ao nível de cada ponto de coleta de dados, o que evita o ponto único de falhas. A utilização do consenso de forma distribuída facilita a tomada de decisão colaborativa. Contudo, o custo computacional de comunicação entre os participantes pode sobrecarregar a rede e diminuir a efetividade do sistema. Em [Kailkhura et al. 2015], os autores desenvolvem um algoritmo robusto de consenso apoiado em média ponderada distribuída. O algoritmo visa permitir uma adaptação à regras locais estipuladas pela rede, onde se desenvolveu uma técnica de aprendizagem para estimar parâmetros operacionais ou peso de cada nó. Assim, torna-se possível automatizar regras locais de fusão ou atualização para mitigar ataques. Ele atua em uma rede distribuída, mas os autores não consideram a dinamicidade como fator de impacto, o que facilita as ações maliciosas.

Assim, faz-se necessário o desenvolvimento de soluções capazes de atuar de forma distribuída e que sejam capazes de identificar e isolar a presença de dispositivos maliciosos em redes IoT densas. Essas soluções, ao preservar a segurança dos dispositivos, não devem interferir no seu funcionamento padrão ou gerar sobrecargas adicionais. Além disso, deve-se atender normalmente as características de uma IoT densa.

3. Um Mecanismo para Mitigação de Ataques IDFs

Esta seção apresenta o sistema CONFINIT (*CONsensus Based Data FIlteriNg for IoT*) para mitigação de ataques de injeção de dados falsos sobre o serviço de disseminação de dados em redes IoT densas voltadas ao ambiente da indústria 4.0. Ele baseia-se no monitoramento entre participantes para detectar anomalias na rede e utiliza a técnica de Consenso Colaborativo para tomada de decisão. O objetivo é garantir a autenticidade dos dados disseminados na rede IoT para apoiar a tomada de decisões das aplicações. O mecanismo CONFINIT atua sobre o serviço de agrupamento executando numa estrutura de rede IoT densa no contexto industrial, conforme ilustrado na Figura 1 e descrito a

seguir, composta por nós heterogêneos, que podem ou não apresentar mobilidade. Além disso, assume-se que a comunicação entre os agrupamentos ocorre através dos nós líderes.

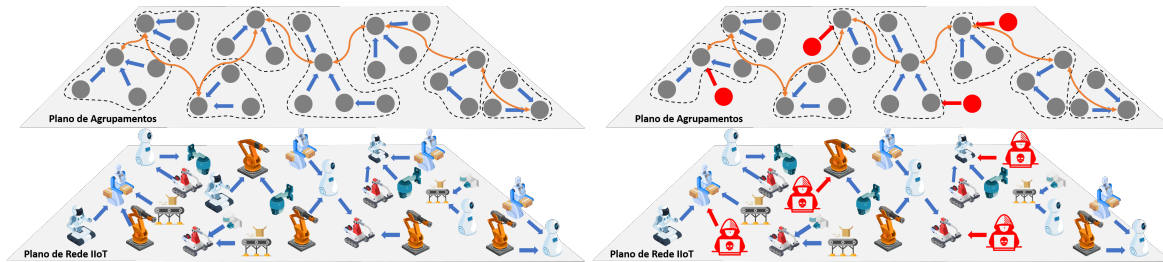


Figura 1. Modelos de Agrupamentos de Dados em rede IoT e de Ataque IDF

Considera-se uma rede IoT densa, como de um ambiente industrial, formada pelo conjunto N de n nós denotados por $N = \{n_1, n_2, n_3, \dots, n_n\}$, onde $n_i \in N$. Cada nó n_i tem um endereço físico exclusivo (ID) que o identifica na rede e a comunicação ocorre através do meio sem fio mediante um canal assíncrono com perda de pacotes devido a ruídos e posição dos nós. Os nós diferenciam-se pelo nível de recursos que possuem, como energia, processamento e memória. Além disso, todos atuam na mesma faixa de transmissão para formarem os agrupamentos. Todos os dispositivos começam como nós isolados buscando vizinhos para formarem os agrupamentos, podendo ou não participar de um agrupamento. Os nós que não fazem parte de nenhum agrupamento podem ser detectados como atacantes ou suspeitos. Os nós podem assumir dois papéis, o de nó comum e o de nó líder. Os nós comuns integram os agrupamentos enviando seus dados em direção aos líderes. Os líderes, por sua vez, são responsáveis por receber os dados e disponibiliza-los à aplicação.

A ameaça à rede se caracteriza como um ataque de injeção de dados falsos (IDF) onde o atacante, uma vez intruso na rede, apresenta má-conduta com características de adulteração, falsificação e fabricação de dados coletados por dispositivos IoT. O ataque IDF ocorre de duas formas. Na primeira forma, o atacante captura o dispositivo e manipula seus dados, seja alterando ou falsificando-os. Já na segunda forma, o próprio dispositivo é o atacante e altera, fabrica ou manipula seus próprios dados. Considera-se que o ataque inicia-se por meio da exploração de vulnerabilidades decorrentes de outros ataques, além do atacante ter total conhecimento sobre a rede [Yang et al. 2017, Yaqoob et al. 2017]. Assim, o atacante explora essas vulnerabilidades, dificultando a identificação dos dispositivos, como ilustrado na Figura 1.

3.1. Arquitetura

A arquitetura do CONFINIT é composta por dois módulos **Controle de Agrupamentos** e **Controle de Falhas** como ilustrado na Figura 2. Ambos trabalham de maneira conjunta e em paralelo para garantir a disseminação de dados segura na rede. O módulo controle agrupamentos organiza a rede em *cluster* e o módulo controle de falha trata do monitoramento dos dispositivos da rede, detecção e isolamento de ações maliciosas de injeção de dados falsos.

O **Módulo Controle de Agrupamentos (MCA)** é responsável pela formação e manutenção dos agrupamentos dentro da rede. Ele avalia com base em um limiar de similaridade de leituras dos dispositivos (nós) que estão próximos e determina quando estão

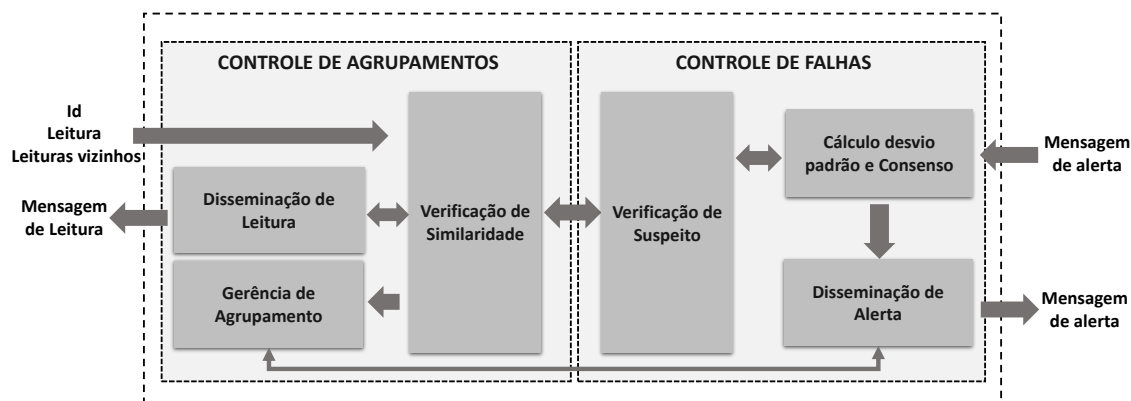


Figura 2. Arquitetura CONFINIT

aptos para formar um agrupamento. Assim, ao receber uma mensagem ele verifica a identificação, a quantidade de vizinhos e as leituras desses vizinhos. O módulo é composto pelos componentes, *Verificação de Similaridade (VA)*, *Gerência de Agrupamento (GA)* e *Disseminação de Leitura (DL)*. O componente VA é responsável por receber e interpretar as mensagens trocadas entre os nós da rede. O componente GA trata de formar e manter os agrupamentos a partir da similaridade entre os nós, ele também é responsável pela eleição dos líderes. Já o componente DL é responsável por divulgar sua leitura, quantidade de vizinho e leitura dos vizinhos, com isso todos os nós que receberem a mensagem saberão quando fazem ou não parte de um agrupamento.

O **Módulo Controle de Falhas (MDF)** garante a segurança da disseminação de dados entre os nós da rede IoT de modo que apenas leituras autênticas sejam disseminadas. Ele consiste dos componentes *Verificação de Suspeito (VS)*, *Detecção (DE)* e *Disseminação de Alerta (DA)*. O componente verificação monitora os nós verificando aqueles que não respeitam o limiar de similaridade. O DE emprega a técnica de consenso colaborativo e desvio padrão para detectar os nós IDf. O consenso é a concordância e uniformidade de opiniões que os nós estabelecem por meio de troca de informações entre eles. O desvio padrão visa determinar quão discrepantes estão as leituras. O componente DA atua para isolar os nós atacantes e informar aos demais membros sobre esse ameaça. Logo, quando um ataque é detectado, os nós participantes da detecção propagam um alerta para que os líderes do agrupamento disseminem-o pela rede. Os dados propagados na mensagem de alerta consistem no Id e leitura individual do atacante.

3.1.1. Formação dos agrupamentos

Visto que o nosso modelo de redes

Uma vez que o modelo de rede IoT é composto por uma grande diversidade de dispositivos com diferentes características, logo torna-se um desafio gerenciar e controlar esses dispositivos. Desta forma, o módulo **Controle de Agrupamentos** organiza a rede em grupos com base nos líderes para criar uma topologia de rede. Inicialmente, os nós iniciam suas operações de forma isolada, transmitindo e coletando mensagens de controle dos nós vizinhos para composição dos agrupamentos. O Algoritmo 1 apresenta o funcionamento do controle de agrupamentos. Periodicamente cada nó envia em *broadcast*

uma mensagem de controle, informando seu identificados (ID), sua leitura atual $L_{ind}()$, a média de leitura agregada dos nós com leitura similares em sua vizinhança $L_{agr}()$, e a quantidade de vizinhos $N_{viz}()$ (l.5). O envio das mensagens leva em consideração um intervalo definido aleatoriamente a fim de evitar transmissões simultâneas entre todos os nós. Ao receber uma mensagem o nó saberá a origem org , a leitura individual iR do nó emissor, a leitura agregada aR de sua vizinhança e a quantidade de nós vizinhos considerados nR . Inicialmente o nó atualiza as informações recebidas (l.7). As leituras agregadas são consideradas para verificar a leitura do nó atual satisfazem o limiar de similaridade (l.8). A verificação de similaridade é executada em (l.9). A vizinhança N_{viz} é atualizada definindo incluindo org caso o limiar de similaridade seja satisfeito (l.10) ou removendo o nó caso não seja similar (l.12). Esse processo ocorre dinamicamente em cada nó da rede, assim cada nó pode manter sua estrutura de vizinhança atualizada.

Verificar com o Agnaldo/gustavo a possibilidade de fazer uma minipage cuja na parte da esquerda fique o algoritmo 1 na parte da direita o algoritmo 2. Não apague os seus algoritmos abaixo, será apenas um teste primeiro.

Algoritmo 1: Estabelecimento dos Agrupamentos

```

1 procedure SENDCONTROLMESSAGE
2    $Send(ID, L_{ind}(), L_{agr}(), |N_{viz}|)$ 
3    $WaitInterval()$ 
4    $MControlTimerExpire()$ 
5 end procedure

6 procedure RECEIVECONTROLMESSAGE( $org, iR, aR, nR$ )
7    $NeighR[org] \leftarrow \{iR, aR, nR\}$ 
8    $localRead \leftarrow L_{agr}()$ 
9   if  $(|iR - localRead| < Threshold)$  and  $(|L_{ind}() - aR| < Threshold)$  then
10     $N_{viz} \leftarrow SNeigh \cup \{org\}$ 
11  else if
12     $N_{viz} \leftarrow SNeigh - \{org\}$ 
13  and if
14 end procedure

```

Colocar a primeira sentença na voz ativa do parágrafo abaixo.

A Equação 1 verifica a existência de similaridade entre duas leituras, ela foi desenvolvida por [Gielow et al. 2015]. Em sua composição são atribuídos, as leituras dos nós, quantidade de vizinhos e leituras agregadas desses vizinhos para determinar quando existe similaridades entre dois nós. Devido a sua característica de leituras dinâmicas, a melhor forma de de se avaliar a similaridade entre dois nós é através da comparação entre elas. Na equação, X representa a leitura atual do nó e Y a leitura com a qual ele está sendo comparada, a fim de determinar com base em um limiar quando as duas leituras são similares.

$$\left| Y - \frac{X + \sum_{v \in SNeigh} (NeighR[v].aR * NeighR[v].nR)}{1 + \sum_{v \in SNeigh} (NeighR[v].nR)} \right| < CThresh \quad (1)$$

3.1.2. Detecção de Falhas

Na formação dos agrupamentos os nós que não atendem ao limiar de similaridade de leituras em um primeiro momento são considerados suspeitos e passam entregar uma lista de suspeitos. A utilização da lista de suspeitos se dá pelo fato de apresentar melhor avaliação sobre a detecção de falhas, já que o dispositivo às vezes apresenta apenas uma falha não o comportamento de atacante. O Algoritmo 2 detalha o funcionamento da detecção de falhas dentro da rede diante de uma ameaça IDF. A detecção passa a atuar efetivamente após a primeira troca de mensagens, visto que necessita de outras mensagens para fazer comparação. Assim, em um primeiro momento verifica-se se o nó em questão consta na lista de suspeito (l.3). Caso ele não esteja mas suas leituras sejam consideradas suspeita ele é inserido na lista de suspeitos, (l.12). Caso seja verificado que ele consta na lista de suspeitos mas suas leituras respeitem o limiar aceitável ele é removido da lista de suspeitos (l.14).

Algoritmo 2: Detecção de nós atacantes

```

1 procedure CONSENSUS(ID, Read)
2   Valid  $\leftarrow$  checkSuspicious(ID, Out)
3   checkSuspicious(ID, out)
4   if (ID  $\in$  SuspectList & Out == False)
5     return 1
6   if (ID  $\in$  SuspectList & Out == True)
7     return 2
8   end if
9   if (Read  $\leq$  ThresholdConsensus)
10    Switch  $\leftarrow$  Valid
11    Case 1
12      Descarta e classifica como Atacante
13    Case 2
14      SuspectList - RemoveSuspeito (ID)
15  else
16    AddSuspect  $\leftarrow$  (ID, SuspectList)
17 end procedure

```

A Equação 2 descreve o cálculo do consenso para verificar qual o desvio dos valores comparados. Para isso, são usados valores observados, os valores de entrada são representados os valores de comparação entre os nós. Para isso, utiliza-se um conjunto $D = (x_i, x_{i+1}, \dots, x_n)$ que representa as amostras a serem verificadas. O cômputo do consenso é indicado por $\sum_{i=1}^n$ que soma todos os valor do conjunto D , desde a primeira posição ($i=1$) até a posição n . X_i representa o valor na posição i no conjunto de dados. M_A representa a média aritmética dos dados. N quantidade de dados a serem avaliados. O *Thresholdconsensus* representa o valor predefinido que pode ser variado de acordo com o tipo de dados a ser validado.

$$DP = \sqrt{\frac{\sum_{i=1}^n (X_i - M_A)^2}{N}} \leq \text{Thresholdconsensus} \quad (2)$$

A Figura 3 ilustra uma formação de consenso colaborativo entre os participantes para a detecção de falhas em razão de um atacante IDF. As setas pontilhadas representa a comunicação entre os nós (**A**, **B**, **C**, **D**) no instante $T1$, garantindo assim a troca de mensagem de controle entre eles. Já no instante $T2$, apenas os nós (**A**, **B**, **C**) agrupam-se visto que eles respeitam o limiar de similaridade, entretanto o nó **D** não respeita o

limiar, então em um primeiro momento é classificado como suspeito. No instante $T3$, o nó **D** envia novamente mensagens de controle para tentar fazer parte do agrupamentos, assim o conjunto formado pelos nós (**A**, **B** e **C**) que executam o cálculo da Equação 2, e classificam o nó **D** como atacante, já que ele novamente apresenta leituras distintas dos demais nós. Por fim, no instante $T4$ apenas os nós honestos estão fazendo parte da rede. Desta forma, a segurança da rede é mantida pelos próprios participantes sem a necessidade de entidade externas.

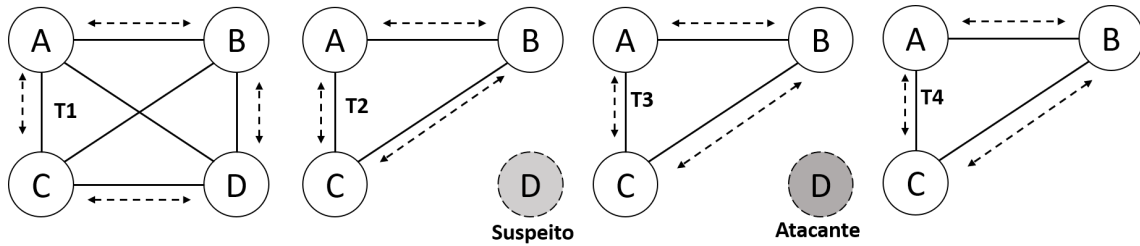


Figura 3. Formação dos Consenso entre os nós

3.2. Funcionamento

A formação de agrupamento ocorre de maneira dinâmica em cada nó da rede. As interações entre os nós realizam-se sob grandezas de espaço e tempo, assim as mensagens de dados são enviadas e recebidas pelos nós que estão dentro do raio de transmissão do emissor. Cada nó envia em *broadcast*, sua leitura, as leituras dos seus vizinhos, e a quantidade de vizinhos. Quando a mensagem é recebida, o nó receptor a interpreta, verificando seus campos e realizando o cálculo da similaridade. Se o limiar de similaridade é respeitado o nó em questão passa a compor o agrupamento, entretanto se o limiar não é respeitado ele passa a integrar a lista de suspeito.

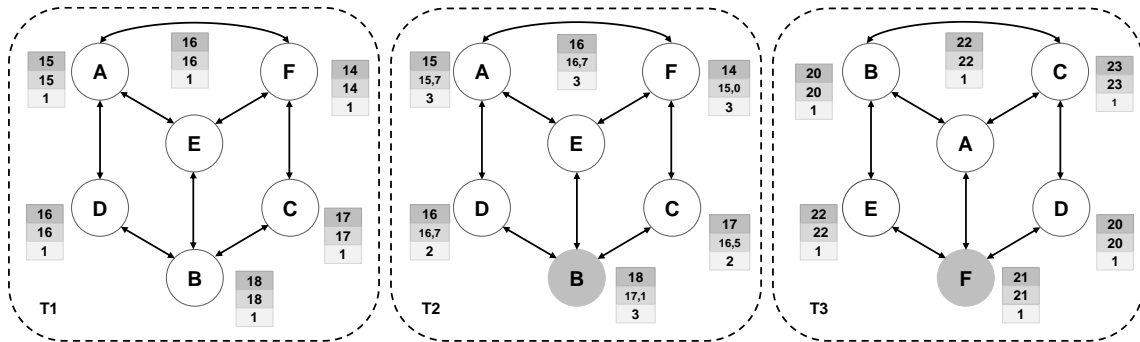


Figura 4. Formação dos agrupamentos

A Figura 4 ilustra um exemplo do funcionamento do mecanismo na formação dos agrupamentos e eleição dos líderes. As arestas sólidas indicam os nós que estão dentro do raio de transmissão um do outro e podem trocar mensagens. As caixas ao lado de cada nó correspondem à estrutura que indica, de cima para baixo, a leitura individual do nó, a leitura agregada sua e de seus vizinhos, a quantidade de leituras agregadas, e o nível de energia do nó. Assim, considera-se um limiar de similaridade = 3 para formação dos agrupamentos. Cada instante T corresponde a uma troca de mensagem entre os nós para formar os agrupamentos e eleger os líderes. Com o controle de agrupamentos

operando desta maneira, cada nó mantém atualizada suas informações através da troca de mensagens. Essa estrutura determina quais nós da vizinhança são vistos como membros do mesmo agrupamento. Essa estrutura simples garante uma melhor escalabilidade à rede, além de ajudar a classificar nós que estejam com leituras divergentes, facilitando a identificação de atacantes pelo módulo de controle de falhas.

A detecção de falhas atua considerando a formação dos agrupamentos. Assim, os nós que não respeitam o limiar de similaridade passam a integrar uma lista de suspeitos em um primeiro momento. A classificação dos nós que não fizeram parte do agrupamento tem início na mensagem de controle, que sem os campos devidamente preenchidos são descartadas. Na formação dos agrupamentos um nó que está próximo fisicamente de seus vizinhos em determinado instante e apresenta leituras distintas não respeitando o limiar de similaridade, pode ser considerado um nó suspeito em um primeiro momento. Assim, ele passa a integrar a lista de suspeitos, a qual contém nós que apresentaram um comportamento anômalo, mas não necessariamente são atacantes. Quando o nó em questão entra para lista de suspeitos e tenta participar do agrupamento, e, novamente, não consegue devido à leituras distintas, ele é classificado como atacante. Logo, seu (*Id*) é inserido em uma lista que contém todos os (*Ids*) das ameaças. Em seguida, o líder o agrupamento envia uma mensagem aos outros líderes, avisando sobre a ameaça, tal que caso ele tente se agrupar em outro momento, não consiga.

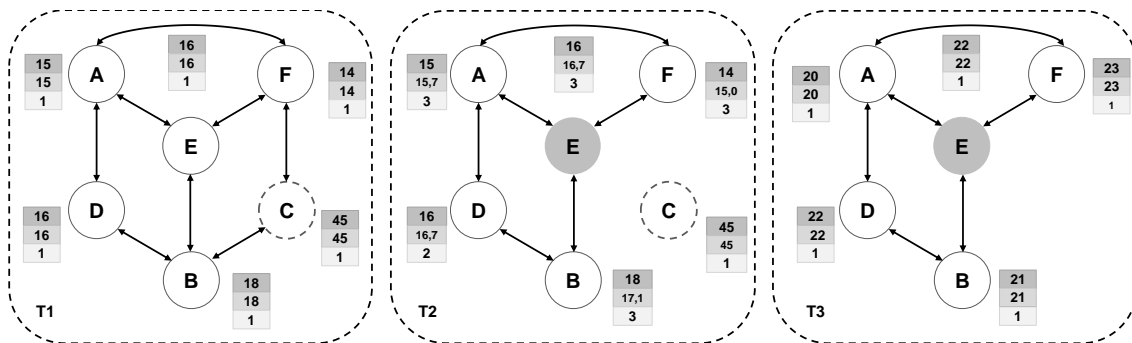


Figura 5. Detecção de falhas

A Figura 5 ilustra um exemplo da detecção de um ataque. Cada instante T corresponde a um processo de agrupamento, assim, os nós que satisfazem o limiar de similaridade passam a formar um agrupamento. No instante $T1$, os nós (A, B, D, E, F) satisfazem esse limiar, mas o nó C está com sua leitura divergente comparada aos dos seus vizinhos, isto é, não satisfazendo a similaridade de dados. Assim, o nó C não pode fazer parte do agrupamento nesse momento. No instante $T2$, o nó C tenta novamente agrupar-se, mas como seu (*Id*) já consta na lista de suspeitos, é feito um consenso entre os participantes com base em suas leituras e comparadas às do nó C. Logo, constata-se que C é um atacante, não podendo integrar nenhum agrupamento. No instante $T3$, retirou-se o nó C da rede e é disseminada uma mensagem de alarme direcionada aos líderes com o (*Id*) do atacante em questão.

4. Avaliação

Esta seção descreve a avaliação de desempenho do mecanismo CONFINIT. Ele foi implementado no simulador NS-3, versão 3.28. O simulador foi escolhido por permitir

representar um ambiente de massiva escala de objetos IoT num cenário industrial com as características da rede. O cenário desenvolvido visa criar um ambiente mais próximo a de uma fábrica, onde os dispositivos IoT estão sobre objetos, representando o modelo de uma rede IIoT. Esses objetos podem ser variados, entre o objetivos aqui é usar leituras realística, assim, o cenário foi baseado nas leituras de sensores pressão de gás. Os dados utilizados foram coletadas e disponibilizados pelo laboratório UCI Machine Learning Repository [UCI 2013] e os *datasets* estão disponíveis para análises. Assim, tenta-se chegar mais próximo de um ambiente real.

O cenário avaliado é composto por 100 nós distribuídos aleatoriamente em uma área retangular de $400m \times 200m$ operando por $1200s$ com um raio de transmissão de $100m$. A área definida tenta representar um ambiente IoT massivo, isso devido ao fato de que os dados foram obtidos de um ambiente real, porém sem descrições sobre o tamanho da área. Assim, ela foi aferida com base nos trabalhos [Cervantes et al. 2018, Kumar and Pais 2018]. A definição do raio de transmissão levou em consideração a proporção da área de cobertura, além de não interferir diretamente nos resultados. A comunicação entre os dispositivos ocorre através do protocolo IPv6, sendo estabelecida uma rede Ad-Hoc no padrão IEEE 802.15.4. Além disso, o protocolo de agrupamento DDFC foi implementado como parte do mecanismo para formação dos agrupamentos. Entretanto foram necessárias algumas modificações para melhor adaptá-lo ao cenário de um rede IoT massiva no contexto Industrial.

Tabela 1. Métricas de Avaliação

Descrição	Equações
Taxa de detecção (T_{det}), calcula os ataques IDF identificados corretamente pelo mecanismo, a partir da interação entre os participantes. O cálculo da (T_{det}) entende a razão entre o total de detecção (det_{ni}), e a quantidade de ataques inseridos, (A_{ins}).	$(T_{det}) = \frac{\sum det_{ni}}{A_{ins}}$
Taxa de Falsos positivos (T_{fp}) define a quantidade de vezes que o mecanismo identificou um ataque IDF quando o mesmo não existia. Seu cálculo se dá através da divisão entre quantidade de detecção, (det_{ni}), é pelo total de interações feitas à rede. (T_{int})	$T_{fp} = \frac{\sum det_{ni}}{T_{int}}$
Taxa de Falsos Negativos (T_{fn}) calcula o percentual de nós identificados de forma errada como não intrusos. O cálculo é representado por X que é o total de interação, i.e. envio de mensagens de controle e T_{det} a taxa de detecção do ataque.	$T_{fn} = X - T_{det}$

O modelo de ataque empregado foi implementado baseia-se no modelo desenvolvido em [Deng et al. 2016], neste modelo o atacante altera dados da matriz de leituras de energia, para isso, são inseridos dois tipo de dados, um que visa uma pequena alteração e outro que altera completamente o que foi coletado. Para melhor se adequar ao cenário desenvolvido optou-se pela alteração completa dos dados. Desta maneira, são selecionados nós específicos para atuarem com atacantes. Os atacantes tem total conhecimento sobre os dados trafegados na rede, facilitando as alteração nos valores coletados. Para alterar os valores são considerados operação que podem zerar os valores, duplicar os dados, ou mesmo substituí-los por valores superiores aos coletados. Além disso, considera-se a atuação de nós suspeitos dentro da rede, que em determinados momentos podem apre-

sentar ou não comportamento de atacante. Desta maneira, chegou-se mais próximo ao comportamento real do ataque IDF. Os resultados apresentados abaixo foram obtidos a partir da realização de 35 simulações. Os gráficos apresentam um intervalo de confiança de 95%. As seguintes métricas foram escolhidas para avaliar o desempenho do CONFINIT na detecção do ataque IDF, elas são apresentadas na Tabela 1.

4.1. Resultados

O gráfico da Figura 6 apresenta a disponibilidade da formação dos agrupamentos ao longo do tempo de três maneiras diferentes, com ataque e sem o mecanismo, com ataque e com o mecanismo e apenas a formação de agrupamento atuando. Para a avaliação o cenário foi configurado com 100 nós e 10% de nós atacantes. Observa-se como a atuação do ataque afeta a formação dos agrupamentos, já que o número de cai pela metade, demonstrando o quanto nocivo é o ataque para rede. Também, observa-se que o mecanismo não tem uma grande interferência na na formação dos agrupamentos, demonstrando assim sua eficiência perante o ataque IDF.

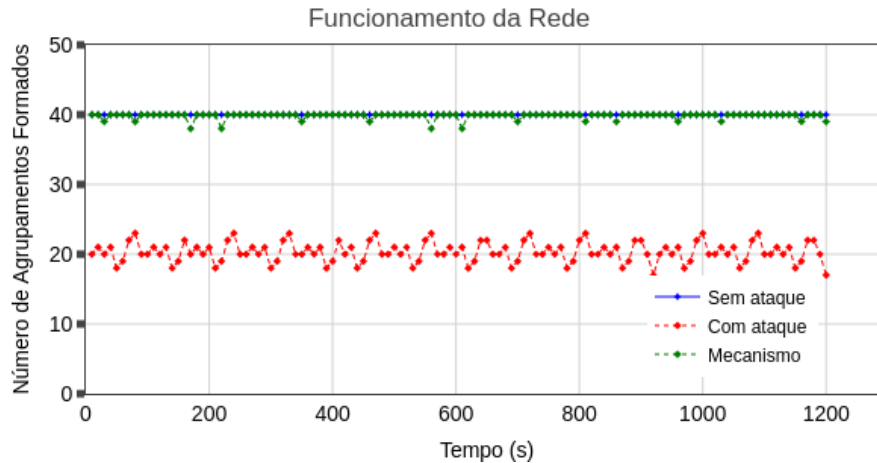


Figura 6. Número de agrupamentos formados ao longo do tempo

O gráficos da Figura 7 apresentam a capacidade de detecção e mitigação obtida pelo mecanismo CONFINIT sobre o ataque. O CONFINIT obteve uma taxa de detecção de (T_{det}) 95%, inclusive alcançando algumas vezes um taxa de 97%. Esta efetividade na detecção alcançada pelo mecanismo deve-se ao uso de monitoramento entre participantes empregado pelo *watchdog* e o consenso colaborativo pelos nós da rede. **quando ocorreu a mitigação? quem eram os atacantes? Qual tipo de comportamento malicioso ele(s) tinha(m)? como eles chegaram ao consenso? quem eram os nós participantes do consenso? eles eram nós líderes? Qual(is) estratégia(s) será(ão) comparada(s) com o CONFINIT?. É preciso mostrar o funcionamento ao longo do tempo porque sem esses gráficos do tempo, os gráficos abaixo não têm credibilidade. É preciso fazer para cada comportamento malicioso e também situação mista.** Além disso, é possível observar que a taxa detecção apresenta pouca varia com relação a quantidade de nós da rede e a porcentagem de ataques inseridos, demonstrando a capacidade do mecanismo em lidar com um ambiente massivo.

A Figura 8 apresenta o desempenho do CONFINIT quanto a taxas de falsos negativos e positivos na presença do ataque IDF. O mecanismo obteve uma taxa falsos negativos menor que 5% para o ataque. Mostrando a baixa taxa de nós que não são detectados. A

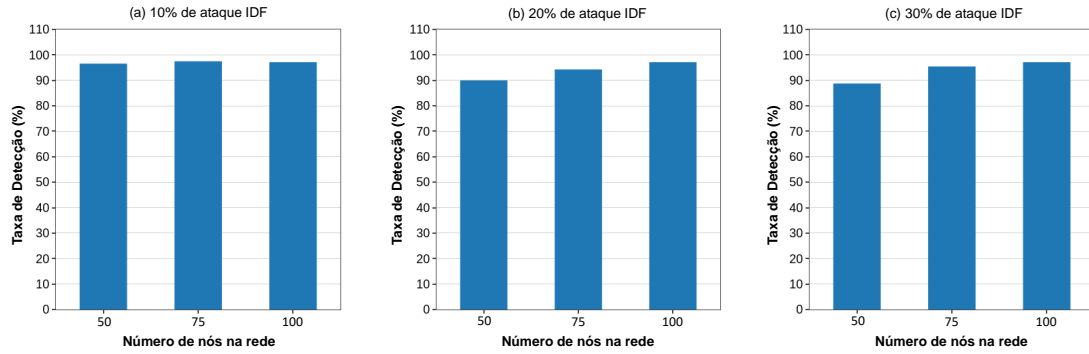


Figura 7. Taxa de detecção (T_{det}) para o ataque IDF

falha na detecção de um atacante pode acontecer quando há erros no cálculo da similaridade, e o nó é indicado como suspeitos, consequentemente o consenso identifica de maneira errada o nó. Desta forma, um nó pode demorar a ser identificado como atacante.

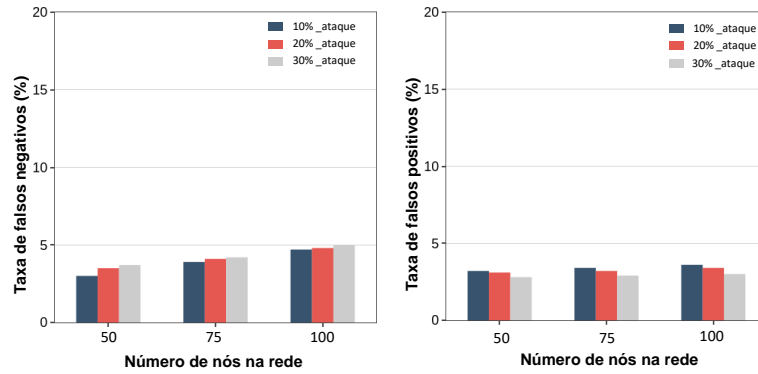


Figura 8. Taxa de falsos negativos (T_{fn}) e positivos (T_{fp}) para o ataque IDF

Em relação às taxas de falsos positivos, o CONFINIT obteve uma taxa inferior a 4% para o ataque IDF com diferentes porcentagens. As detecções erradas são decorrentes de erros no cálculo de consenso entre os nós monitores sobre um ataque. Logo, em um primeiro momento eles são considerados suspeitos, porém conforme as novas trocas de mensagens vão acontecendo e os cálculos são efetuados, eles são identificados como nós corretos

5. Conclusão

Este trabalho apresentou o mecanismo CONFINIT para mitigação de ataques de injeção de dados falsos em rede IoT densa no contexto Industrial. O mecanismo organiza a rede em agrupamentos para lidar com a densidade da rede, levando em consideração a similaridade de leituras entre os dispositivos. Além disso, ele monitora o comportamento dos nós com relação às suas leituras, a fim de determinar quais nós apresentam um comportamento malicioso em relação aos outros nós. Para isso, são empregadas técnicas de monitoramento Watchdog e Consenso colaborativo. Os resultados obtidos por simulação demonstraram a eficácia do CONFINIT na detecção dos nós atacante IDF e a garantia da disseminação de dados segura. Como trabalhos futuros pretende-se comparar o CONFINIT com outros mecanismos propostos, além de avaliá-lo em outros contextos de rede IoT densa e diferentes leituras.

Referências

- Akpakwu, G. A., Silva, B. J., Hancke, G. P., and Abu-Mahfouz, A. M. (2018). A survey on 5g networks for the internet of things: Communication technologies and challenges. *IEEE Access*, 6:3619–3647.
- Borgia, E. (2014). The internet of things vision: Key features, applications and open issues. *Computer Communications*, 54:1–31.
- Bostami, B., Ahmed, M., and Choudhury, S. (2019). False data injection attacks in internet of things. In *Performability in Internet of Things*, pages 47–58. Springer.
- Cervantes, C., Nogueira, M., and Santos, A. (2018). Mitigação de ataques no roteamento em iot densa e móvel baseada em agrupamento e confiabilidade dos dispositivos. In *Anais do XXXVI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*. SBC.
- Colistra, G., Pilloni, V., and Atzori, L. (2014). Task allocation in group of nodes in the iot: A consensus approach. In *2014 IEEE International Conference on Communications (ICC)*, pages 3848–3853. IEEE.
- Deng, R., Xiao, G., Lu, R., Liang, H., and Vasilakos, A. V. (2016). False data injection on state estimation in power systems—attacks, impacts, and defense: A survey. *IEEE Transactions on Industrial Informatics*, 13(2):411–423.
- Gielow, F., Jakllari, G., Nogueira, M., and Santos, A. (2015). Data similarity aware dynamic node clustering in wireless sensor networks. *Ad Hoc Networks*, 24:29–45.
- Gubbi, J., Buyya, R., Marusic, S., and Palaniswami, M. (2013). Internet of things (iot): A vision, architectural elements, and future directions. *Future generation computer systems*, 29(7):1645–1660.
- Kailkhura, B., Brahma, S., and Varshney, P. K. (2015). Consensus based detection in the presence of data falsification attacks. *arXiv preprint arXiv:1504.03413*.
- Kouicem, D. E., Bouabdallah, A., and Lakhlef, H. (2018). Internet of things security: A top-down survey. *Computer Networks*.
- Kumar, A. and Pais, A. R. (2018). Deterministic en-route filtering of false reports: A combinatorial design based approach. *IEEE Access*, 6:74494–74505.
- Kumar, S. A., Vealey, T., and Srivastava, H. (2016). Security in internet of things: Challenges, solutions and future directions. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 5772–5781. IEEE.
- Li, B., Lu, R., Wang, W., and Choo, K.-K. R. (2017). Distributed host-based collaborative detection for false data injection attacks in smart grid cyber-physical system. *Journal of Parallel and Distributed Computing*, 103:32–41.
- Lu, R., Lin, X., Zhu, H., Liang, X., and Shen, X. (2012). Becan: A bandwidth-efficient cooperative authentication scheme for filtering injected false data in wireless sensor networks. *IEEE transactions on parallel and distributed systems*, 23(1):32–43.
- Mendez, D. M., Papapanagiotou, I., and Yang, B. (2017). Internet of things: Survey on security and privacy. *arXiv preprint arXiv:1707.01879*.
- Minoli, D., Sohraby, K., and Occhiogrosso, B. (2017). Iot considerations, requirements, and architectures for smart buildings—energy optimization and next-generation building management systems. *IEEE Internet of Things Journal*, 4(1):269–283.
- Miorandi, D., Sicari, S., De Pellegrini, F., and Chlamtac, I. (2012). Internet of things: Vision, applications and research challenges. *Ad hoc networks*, 10(7):1497–1516.
- Mumtaz, S., Alsohaily, A., Pang, Z., Rayes, A., Tsang, K. F., and Rodriguez, J. (2017). Massive internet of things for industrial applications: Addressing wireless iiot connectivity challenges and ecosystem fragmentation. *IEEE Industrial Electronics Magazine*, 11(1):28–33.

- Qiu, T., Chen, N., Li, K., Atiquzzaman, M., and Zhao, W. (2018). How can heterogeneous internet of things build our future: A survey. *IEEE Communications Surveys & Tutorials*, 20(3):2011–2027.
- Sen, A. and Madria, S. (2017). Risk assessment in a sensor cloud framework using attack graphs. *IEEE Transactions on Services Computing*, 10(6):942–955.
- Toulouse, M., Minh, B. Q., and Curtis, P. (2015). A consensus based network intrusion detection system. In *2015 5th International Conference on IT Convergence and Security (ICITCS)*, pages 1–6. IEEE.
- UCI, C. (2013). Estatísticas de acesso web. <https://archive.ics.uci.edu/ml/datasets/Gas+Sensor+Array+Drift+Dataset>. Acessado em 21/05/2018.
- Wang, J., Liu, Z., Zhang, S., and Zhang, X. (2014). Defending collaborative false data injection attacks in wireless sensor networks. *Information Sciences*, 254:39–53.
- Yang, L., Ding, C., Wu, M., and Wang, K. (2017). Robust detection of false data injection attacks for data aggregation in an internet of things-based environmental surveillance. *Computer Networks*, 129:410–428.
- Yaqoob, I., Ahmed, E., ur Rehman, M. H., Ahmed, A. I. A., Al-garadi, M. A., Imran, M., and Guizani, M. (2017). The rise of ransomware and emerging security challenges in the internet of things. *Computer Networks*, 129:444–458.
- Yu, Z. and Guan, Y. (2010). A dynamic en-route filtering scheme for data reporting in wireless sensor networks. *IEEE/ACM Transactions on Networking (ToN)*, 18(1):150–163.