

Weslen Schiavon Souza

Literary survey on event processing applied to the Internet of Things

Trabalho Individual apresentado ao Programa de Pós-Graduação em Computação da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Mestre em Ciência da Computação

Orientador: Prof^a. Dr^a. Ana Marilza Pernas
Coorientadora: Prof. Dr. Adenauer Corrêa Yamin

Pelotas, 2018

RESUMO

SOUZA, Weslen Schiavon. **Literary survey on event processing applied to the Internet of Things**. 2018. 40 f. Trabalho Individual (Mestrado em Ciência da Computação) – Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, 2018.

A computação ubíqua apresenta como um de seus objetivos disponibilizar a computação de forma totalmente transparente e integrada ao mundo físico tornando esta imperceptível aos usuários. Um paradigma emergente que tem se mostrado ser uma materialização da computação ubíqua é a Internet das Coisas, porém para que a internet das coisas se torne efetivamente uma realidade para todos os tipos de usuários, é necessário ainda encontrar soluções para superar alguns desafios presentes nesse paradigma: Heterogeneidade dos dispositivos, o grande volume de dados, formas de configurar dispositivos nestas redes e a segurança das informações trafegadas. O objetivo deste trabalho é executar um levantamento literário sobre processamento de eventos aplicado a internet das coisas, onde para atingir este determinado fim, executou-se um mapeamento sistemático baseado na metodologia proposta por Petersen et al. (2008), onde após se seguir uma sequência determinada de passos proposto por esta metodologia, foi descoberto dez trabalhos de interesse, os quais faziam a aplicação de processamento de eventos sobre algum aspecto da internet das coisas. Por fim executou-se uma análise comparativa entre os trabalhos de interesse descobertos onde é levantado uma discussão sobre alguns aspectos destes documentos, dentre estes pode-se citar a reusabilidade dos trabalhos onde apenas um demonstrou uma proposta passível de reimplementação.

Palavras-Chave: internet das coisas; processamento de eventos; computação ubíqua

ABSTRACT

SOUZA, Weslen Schiavon. **Assessment of Adaptive Security Strategies for the Internet of Things**. 2018. 40 f. Trabalho Individual (Mestrado em Ciência da Computação) – Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, 2018.

Ubiquitous computing presents as one of its objectives to make computing available in a totally transparent and integrated way to the physical world, making it imperceptible to users. An emerging paradigm that has proven to be a materialization of ubiquitous computing is the Internet of Things, but in order for the Internet of Things to become a reality for all types of users, it is still necessary to find solutions to overcome some of the challenges present in this paradigm : Device heterogeneity, large data volume, ways of configuring devices in these networks, and the security of information passed. The objective of this work is to carry out a literary survey on event processing applied to the internet of things, where to achieve this end, a systematic mapping was performed based on the methodology proposed by Petersen et al. (2008), where after following a determined sequence of steps proposed by this methodology, ten works of interest were discovered, which made the application of event processing on some aspect of the internet of things. Finally, a comparative analysis was carried out between the works of interest discovered where a discussion about some aspects of these documents was raised, among them we can mention the reusability of the works where only one demonstrated a proposal that can be redeployed.

Keywords: internet of things; event processing; ubiquitous computing

LISTA DE FIGURAS

Figura 1	Conjunto de associação de processamento.	9
Figura 2	Strings de buscas usadas.	16
Figura 3	Percentual de publicações encontradas por base.	17
Figura 4	Número de publicações encontradas por base.	18
Figura 5	Quantidade de publicações de interesse por ano.	19
Figura 6	Fluxo de triagem dos artigos.	20
Figura 7	Comparativo entre os artigos selecionados.	35

LISTA DE ABREVIATURAS E SIGLAS

UbiComp	<i>Ubiquitous Computing</i>
IoT	<i>Internet of Things</i>
MQTT	<i>Message Queuing Telemetry Transport</i>
EP	<i>Event Processing</i>
CEP	<i>Complex Event Processing</i>
ESP	<i>Event Streaming Processing</i>
ML	<i>Machine Learning</i>

SUMÁRIO

1	INTRODUÇÃO	6
2	EMBASAMENTO TEÓRICO	9
2.1	Internet das Coisas	10
2.2	Definição de Evento	11
2.3	Definição de Processamento de Eventos	11
2.3.1	Definição de Processamento de Fluxo de Eventos	12
2.3.2	Definição de Processamento de Eventos Complexos	13
3	ESTADO DA ARTE	15
3.1	Mapeamento Sistemático da Literatura	15
3.1.1	Critérios de Inclusão e Exclusão	17
3.2	Trabalhos Relacionados	21
3.2.1	Towards a Generalized Approach for Deep Neural Network based Event Processing for the Internet of Multimedia Things	21
3.2.2	A Web-based Approach using Reactive Programming for Complex Event Processing in Internet of Things Applications	22
3.2.3	Semantic IoT Middleware-enabled Mobile Complex Event Processing for Integrated Pest Management	23
3.2.4	Predictive Analytics for Complex IoT Data Streams	24
3.2.5	DRESS: A Rule Engine on Spark for Event Stream Processing	25
3.2.6	TrustCEP: Adopting a Trust-Based Approach for Distributed Complex Event Processing	26
3.2.7	Analysis of Controller Based IEEE 802.11 System with Similarity Measure Clustering	28
3.2.8	Parallel big data processing system for security monitoring in Internet of Things networks*	29
3.2.9	An integrated information lifecycle management framework for exploiting social network data to identify dynamic large crowd concentration events in smart cities applications	30
3.2.10	CEML: Mixing and moving complex event processing and machine learning to the edge of the network for IoT	32
3.3	Discussão dos Trabalhos Relacionados	33
3.4	Considerações do Capítulo	36
4	CONSIDERAÇÕES FINAIS	37
	REFERÊNCIAS	38

1 INTRODUÇÃO

A computação ubíqua (*Ubiquitous Computing* - UbiComp) apresenta como um de seus objetivos disponibilizar a computação de forma totalmente transparente e integrada ao mundo físico tornando esta imperceptível aos usuários, proporcionando que estes usuários não necessitem interagir com o gerenciamento da infraestrutura dos dispositivos presentes neste ambiente, sendo necessário apenas que as pessoas façam uso da tecnologia sem quaisquer preocupações. Um paradigma emergente que tem se mostrado ser uma materialização da computação ubíqua é a Internet das Coisas (*Internet of Things* - IOT), este paradigma computacional consiste da integração de redes com dispositivos sem fio, computação em nuvem, dados analíticos, tecnologias interativas, bem como dispositivos inteligentes (WEISER, 1999).

A internet das coisas demonstra ter o potencial de mudar tudo ao redor das pessoas, adicionando "inteligência" aos mais variados itens do nosso dia a dia, visando assim fornecer a seus usuários algum tipo de serviço. Para atingir este objetivo o conceito básico abordada pela IoT consiste em fazer com que diferentes tipos de dispositivos se comuniquem, interagindo e cooperando entre si para que deste modo, estes possam atingir um objetivo comum, como o fornecimento de um serviço qualquer para um usuário final. As áreas de aplicação para a IoT são das mais diversas possíveis, desde o uso em cidades grandes aplicando sensores visando fornecer diferentes tipos de serviços a sua população, como por exemplo fornecer informações sobre o tráfego e eventos públicos; como o uso na agricultura com o intuito de executar um monitoramento preciso de informações sobre as plantas e solo (GONÇALVES, 2017).

Porém para que a internet das coisas se torne efetivamente uma realidade, presente nos mais diferentes lugares e servindo a todos os tipos de usuários e não apenas a especialistas, a comunidade científica precisa encontrar soluções para superar alguns desafios presentes nesse novo paradigma. Heterogeneidade da rede, como a IoT visa que os mais variados dispositivos podem estar presentes em uma rede trabalhando junto, a desafios presentes na forma como os dados gerados por esses diferentes dispositivos devem ser tratados tendo em vista que como estes são disposi-

tivos potencialmente com hardwares e recursos totalmente diferentes as informações geradas por estes são muito discrepantes não havendo nenhum padrão no formato dos dados seguido atualmente (AGRAWAL; VIEIRA, 2013). O grande volume de dados gerados por estas redes de dispositivos, já que potencialmente as redes poderão conter dezenas de milhares de dispositivos interconectados se comunicando constantemente e gerando dados continuamente, onde estes, na maior parte dos casos precisam ser processados e analisados. A configuração desses dispositivos a qual deve ser feita de forma simplificada já que usuários finais sem conhecimentos técnicos precisam ser capazes de adicionar e remover dispositivos e recursos de uma rede sempre que estes desejarem. A segurança aplicada sobre estas redes, tendo em vista que soluções comuns de segurança não podem ser aplicadas sobre estas redes, devido a grande variedade de dispositivos presentes nestas redes, onde muitos destas "coisas" conectados podem potencialmente apresentar hardwares simples de baixo poder computacional que não tem capacidade para executar técnicas de criptográficas modernas (AGRAWAL; VIEIRA, 2013).

Algumas previsões nos mostram que o há um crescimento constante no número de dispositivos conectados, em 2020 serão mais de 50 bilhões de dispositivos conectados a internet, tais perspectivas nos mostram que a IoT é um futuro próximo o que enfatiza a necessidade do desenvolvimento de soluções que sejam capaz de resolver os problemas citados de forma eficiente (XAVIER, 2016). Outro paradigma computacional que vem sendo aplicado sobre a IoT na tentativa de solucionar Tais problemas citados é o Processamento de Eventos (*Event Processing* - EP), o qual pode ser visto como um paradigma onde fluxos de dados (estes chamados de eventos) são analisados continuamente com o objetivo de extrair informações úteis de alto nível destes eventos analisados. O processamento de eventos pode ainda ser visto como uma grande classe, a qual é subdividida em outras duas principais áreas: o processamento de fluxo de eventos (*event stream processing* - ESP), que se caracteriza por ter a capacidade de executar operações contínuas como filtros, agregações, classificações e junções, sob fluxos de dados; e o processamento de eventos complexos (*complex event processing* - CEP) o qual faz uso de padrões pré definidos, aplicando-os sobre sequências de eventos simples, para assim fazer a detecção de eventos compostos (DAYARATHNA; PERERA, 2018).

O objetivo deste trabalho é executar um levantamento literário sobre processamento de eventos aplicado a internet das coisas, onde para atingir este determinado fim, executou-se um mapeamento sistemático baseado na metodologia proposta por Petersen et al. (2008), onde após se seguir uma sequência determinada de passos proposto por esta metodologia, foi descoberto dez trabalhos de interesse, os quais faziam a aplicação de processamento de eventos sobre algum aspecto da internet das coisas. Por fim executou-se uma análise comparativa entre os trabalhos de in-

teresse descobertos onde é levantado uma discussão sobre alguns aspectos destes documentos, dentre estes pode-se citar a reusabilidade dos trabalhos onde apenas um demonstrou uma proposta passível de reimplementação.

O restante deste trabalho está organizado da seguinte forma. O capítulo 2 onde é apresentado o embasamento teórico necessário para a compreensão da discussão levantada sobre os trabalhos identificados no mapeamento sistemático elaborado. No capítulo 3 é apresentada a execução do mapeamento sistemático elaborado por este documento, ainda é apresentada uma discussão sobre os trabalhos de interesse identificados durante o mapeamento sistemático. Por fim no Capítulo 4 é apresentada as considerações finais obtidas com o desenvolvimento deste trabalho.

2 EMBASAMENTO TEÓRICO

Neste capítulo será apresentado ao leitor o embasamento teórico sobre internet das coisas apresentando seus principais objetivos e desafios a serem superados, ainda neste capítulo será abordado o conceito de processamento de eventos uma grande área de pesquisa científica, a qual pode ser subdividida em outras duas principais áreas: o processamento de fluxo de eventos (*event stream processing* - ESP), que se caracteriza por ter a capacidade de executar operações contínuas como filtros, agregações, classificações e junções, sob fluxos de dados; e o processamento de eventos complexos (*complex event processing* - CEP) o qual faz uso de padrões pré definidos, aplicando-os sobre sequências de eventos simples, para assim fazer a detecção de eventos compostos (DAYARATHNA; PERERA, 2018). Na figura 1 é ilustrada a relação associativa entre estes três paradigmas, onde o processamento de eventos pode ser visto como um conceito mais genérico o qual engloba ESP que por sua vez engloba o CEP.

Nas seções seguintes deste capítulo será apresentado ao leitor o embasamento teórico necessário, sobre internet das coisas e processamento de Eventos, para que assim este possa ter uma melhor compreensão dos trabalhos apresentados no capítulo 3.

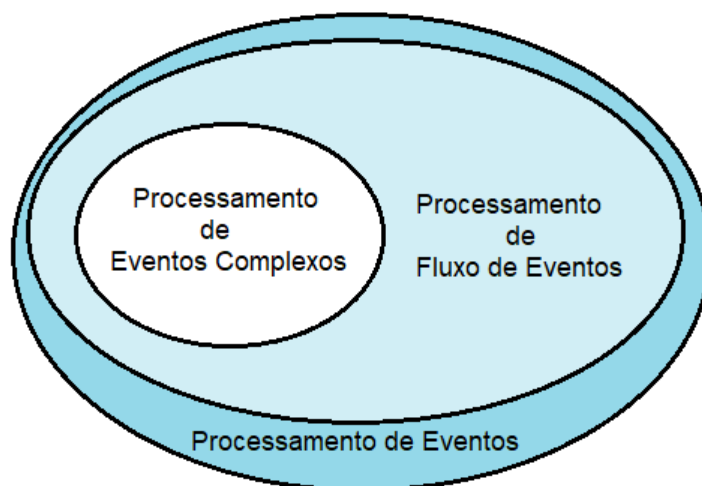


Figura 1 – Conjunto de associação de processamento.

2.1 Internet das Coisas

A tecnologia computacional vem avançando consideravelmente nos últimos anos onde dispositivos cada vez menores com capacidade de se comunicar em rede com alto poder computacional estão se tornando cada vez mais comuns no dia a dia. Esta popularização do uso desses dispositivos "inteligentes" tem sido chamada de internet das coisas, onde este novo paradigma computacional vem mudando a forma como as pessoas interagem com os objetos de seu cotidiano (XAVIER, 2016).

A internet das coisas demonstra ter o potencial de mudar tudo ao nosso redor, adicionando "inteligência" aos mais variados itens do nosso dia a dia, visando assim fornecer a seus usuários algum tipo de serviço. Para atingir este objetivo o conceito básico abordada pela IoT consiste em fazer com que diferentes tipos de dispositivos se comuniquem, interagindo e cooperando entre si para que deste modo, estes possam atingir um objetivo comum, como o fornecimento de um serviço qualquer para um usuário final. As áreas de aplicação para a IoT são das mais diversas possíveis, desde o uso em cidades grandes aplicando sensores visando fornecer diferentes tipos de serviços a sua população, como por exemplo fornecer informações sobre o tráfego e eventos públicos; como o uso na agricultura com o intuito de executar um monitoramento preciso de informações sobre as plantas e solo (GONÇALVES, 2017).

Porém para que a internet das coisas se torne efetivamente uma realidade, presente nos mais diferentes lugares e servindo a todos os tipos de usuários e não apenas a especialistas, a comunidade científica precisa encontrar soluções para superar alguns desafios presentes nesse novo paradigma. Heterogeneidade da rede, como a IoT visa que os mais variados dispositivos podem estar presentes em uma rede trabalhando junto, a desafios presentes na forma como os dados gerados por esses diferentes dispositivos devem ser tratados tendo em vista que como estes são dispositivos potencialmente com hardwares e recursos totalmente diferentes as informações geradas por estes são muito discrepantes não havendo nenhum padrão no formato dos dados seguido atualmente (AGRAWAL; VIEIRA, 2013). O grande volume de dados gerados por estas redes de dispositivos, já que potencialmente as redes conterão dezenas de milhares de dispositivos interconectados se comunicando constantemente e gerando dados continuamente, onde estes, na maior parte dos casos precisam ser processados e analisados. A configuração desses dispositivos a qual deve ser feita de forma simplificada já que usuários finais sem conhecimentos técnicos precisam ser capazes de adicionar e remover dispositivos e recursos de uma rede sempre que estes desejarem. A segurança aplicada sobre estas redes, tendo em vista que soluções comuns de segurança não podem ser aplicadas sobre estas redes, devido a grande variedade de dispositivos presentes nestas redes, onde muitos destas "coisas" conectados podem potencialmente apresentar hardwares simples de baixo poder

computacional que não tem capacidade para executar técnicas de criptográficas modernas (AGRAWAL; VIEIRA, 2013).

Algumas previsões nos mostram que o há um crescimento constante no número de dispositivos conectados, em 2020 serão mais de 50 bilhões de dispositivos conectados à internet, tais dados nos mostram que a IoT é um futuro próximo o que enfatiza a necessidade do desenvolvimento de soluções que sejam capaz de resolver os problemas citados de forma eficiente (XAVIER, 2016).

2.2 Definição de Evento

A definição de evento considerada neste documento é definida como a ocorrência de uma determinada ação dentro de um ambiente, que geralmente envolve uma tentativa de mudança de estado do sistema. Esta mudança inclui, normalmente, a noção de tempo, localidade e detalhes pertencentes ao evento ou a anomalia que desencadeou determinado evento que visam ajudar a compreender as causas ou efeitos desencadeadores do evento (FITZGERALD et al., 2010). Um evento também pode ser separado em campos que descrevem suas propriedades, como por exemplo um evento em uma rede IoT pode incluir quatro atributos: eventID, eventName, eventTime e recordTime. EventID e eventName são normalmente definidos como registros básicos e são armazenados no data center, eventTime e recordTime expressam o conceito de tempo no evento descrevendo sua hora de ocorrência e de captura respectivamente (MINBO; ZHU; GUANGYU, 2013).

Tais eventos podem ser aplicados em diversos sistemas para se atingir determinados fins, como por exemplo o uso em ferramentas de monitoramento onde estes eventos são utilizados para representar mudanças em situações (ETZION; NIBLETT; LUCKHAM, 2011). Estes sistemas monitorados podem ser representados por conjuntos de sensores, onde por exemplo, em aplicações na agricultura de precisão são usados para o monitoramento da umidade e acidez do solo, de forma que tais valores emitidos por estes sensores podem ser vistos como mudanças de estado do ambiente, como uma mudança brusca na acidez do solo ou em sua umidade, as quais podem ser representadas como uma mudança de situação de interesse (SANCHEZ, 2011).

2.3 Definição de Processamento de Eventos

O processamento de eventos pode ser visto como um paradigma onde fluxos de eventos são analisados continuamente com o objetivo de extrair informações úteis de alto nível destes dados analisados. Existem diversas áreas que possuem características com potencial de serem exploradas pelo processamento de eventos, dentre estas pode-se citar os setores da saúde com o monitoramento do status da saúde dos

paciente onde os diversos eventos precisam ser processados e analisados (WEINER et al., 2008); o setor da agricultura de precisão o qual emprega diversos sensores para o monitoramento de plantas, gerando grandes fluxos de eventos que necessitam ser processados (SANCHEZ, 2011) e o setor de energia fazendo uso de eventos para o monitoramento do consumo excessivo de energia visando atingir uma eficiência energética melhor (VIJAYARAGHAVAN; DORNFELD, 2010).

Apesar do processamento de eventos ser aplicado para resolver problemas em áreas totalmente distintas, existem alguns requisitos exigidos por estes que são normalmente semelhantes, tais como a necessidade de processar em tempo de execução grandes volumes de dados. Um setor em constante crescimento que tem gerado grandes interesses no processamento de eventos é a Internet das Coisas onde processamento de eventos tem se aplicado visando solucionar problemas de tomadas de decisão a partir da análise de grandes volumes de dados gerados por estas redes, diversas ferramentas foram desenvolvidas para o processamento de eventos, visando auxiliar na análise desse grande volume de dados, dentre estas pode-se citar o Apache Storm¹, Apache Spark² e Apache Flink³.

2.3.1 Definição de Processamento de Fluxo de Eventos

Para uma melhor e simples compreensão do significado de processamento de fluxo de eventos pode-se separar esta classe em outras três subclasses menores:

1. **Evento** - pode-se definir evento neste contexto como qualquer ação que aconteça com um tempo claramente definido, onde o mesmo pode ser mensurado.
2. **Fluxo** - é definido como um sequência constante e contínuo de eventos disparados por dispositivos, onde esta corrente de eventos são claramente ordenados pelo tempo.
3. **Processamento** - é basicamente a ação final de executar a análise sobre o conjunto de informações capturadas.

Desta forma com a combinação destes três sub-termos pode-se dizer que o processamento de fluxo de eventos nada mais é que processo de analisar em tempo de execução fluxo de eventos disparados por dispositivos assim que estes são criados (DAYARATHNA; PERERA, 2018).

O processamento de fluxo de dados trata da identificação de padrões ou relacionamentos significativos entre os fluxos de dados analisados a fim de detectar determinados padrões como a correlação de eventos, causalidade ou tempo. Características

¹<https://storm.apache.org/>

²<https://spark.apache.org/>

³<https://flink.apache.org/>

presentes em sistemas com aplicabilidade de processamento de fluxo de eventos são a necessidade de analisar grandes fluxos de dados correlacionando estas informações, aplicando filtros em tempo de execução e dando uma resposta de forma imediata, (APPEL et al., 2013).

2.3.2 Definição de Processamento de Eventos Complexos

O processamento de eventos complexos é uma paradigma da computação onde este é aplicada para o processamento e análise de conjunto de fluxos de informações em sistemas baseados em eventos, visando analisar a interação destes eventos entre si, sistemas que empregam este paradigma normalmente apresentam as seguintes características: a necessidade de se verificar e informar a ocorrência de uma ação composta, isto é, a necessidade de identificar que dado a ocorrência de uma ação A e B em um determinado intervalo de tempo, o sistema deve ser capaz de informar que tais ações em conjunto formam uma nova ação C, a qual tem um valor semântico distinto se comparado com A e B individualmente. Uma área com grande aplicabilidade para o processamento de eventos complexos que tem se destacado nos últimos anos é a Internet das Coisas que apresentam grandes fluxos de dados de fontes heterogêneas e que normalmente necessitam ser analisados em tempo de execução (JUN; CHI, 2014).

Assim para definirmos o conceito de processamento de eventos complexos tomado no desenvolvimento deste documento, será executado uma separação deste termo mais complexo em outros dois sub-conceitos menores que o compõe com o objetivo de facilitar a sua compreensão, Processamento e eventos complexos onde pode-se definir processamento como o ato final de analisar as informações já coletadas pelo sistema e eventos complexos como um forma evento nova mais abstrata de alto nível inferida a partir de eventos simples. Mais especificamente pode-se definir um evento complexo como sendo a combinação de dois ou mais eventos simples com o objetivo de criar um novo evento de mais alto nível, por exemplo (DAYARATHNA; PERERA, 2018), em um data center onde sensores monitoram o uso do disco rígido do sistema e o uso de rede, os eventos de alto uso do disco repentinamente disparam, após a análise destes eventos o sistema pode "decidir"disparar um novo evento "possível ataque Hacker"onde a partir deste novo evento os administradores podem decidir tomar alguma decisão baseada neste novo evento como por exemplo desligar da rede do data center (WU; DIAO; RIZVI, 2006).

Assim quando combinamos esses dois sub-conceitos citados, tem-se que o processamento de eventos complexos é o ato de analisar conjuntos de fluxos de eventos simples, visando assim inferir a partir destes, um novo conjunto de eventos semanticamente distintos dos anteriores. Ou seja, pode-se dizer que a partir de uma análise combinatória de eventos simples o CEP é capaz de gerar um novo conjunto de in-

formações, semanticamente de mais alto nível que as informações combinadas na análise.

3 ESTADO DA ARTE

Neste capítulo é apresentado o estado da arte das pesquisas que tem como tema processamento de eventos e internet das coisas. Na seção seguinte é apresentado o protocolo seguido para a execução do mapeamento sistemático assim como todos os passos executados que levaram a escolha dos trabalhos de interesse. Por fim será apresentado uma discussão sobre as soluções abordadas nos trabalhos de interesse selecionados.

3.1 Mapeamento Sistemático da Literatura

O mapeamento sistemático abordado neste capítulo é baseada na metodologia proposta por Petersen et al. (2008), onde seguindo a série de passos proposto, torna o estudo realizado, possível de ser replicado por outros pesquisadores (PETERSEN et al., 2008). A partir desta metodologia, pode ser citado cinco etapas das quais serão seguidas por este mapeamento:

1. Definição das questões de pesquisa;
2. Execução da pesquisa para identificação de estudos primários realizados;
3. Triagem, inicial empregando critérios de inclusão e exclusão considerando o resumo dos artigos;
4. Triagem final, considerando as seções de introdução, concepção do projeto e conclusão;
5. Extração dos dados e mapeamento.

Para a consulta dos trabalhos relacionados primeiramente foi definido um conjunto de palavras como candidatas a palavras chave para a *string* de busca, dentre estas pode-se citar: *internet of things*, *distributed* e *complex event processing*. A Partir da definição destas como palavras chave, foi possível elaborar a *string* de busca usada para executar as consultas sobre as bases da: ACM Digital Library, IEEE Explore,

ScienceDirect, Springer, Web of Science e Scopus; e assim obter-se os trabalhos relacionados com o tema de pesquisa, as strings de consulta podem ser vistas na figura 2 incluindo em qual respectiva base estas foram executadas.

Base de Dados	String de Busca
<i>ACM Digital Library</i>	recordAbstract:(distributed AND ("internet of things" OR iot) AND ("event stream processing" OR "event processing" OR "complex event processing"))
<i>Demais Bases</i>	distributed AND ("internet of things" OR iot) AND ("event stream processing" OR "event processing" OR "complex event processing"))

Figura 2 – Strings de buscas usadas.

Após a execução desta consulta preliminar, que entende-se como a etapa de levantamento dos estudos primários relevantes, foram identificado 647 trabalhos de interesse onde este valor compreende-se da soma dos resultados obtidos em todas as bases de consulta.

Todas as buscas foram realizadas sobre os metadados dos artigos(título, resumo e palavras chave), porém, como a base de dados Springer não oferecia suporte a este tipo de consulta, este problema foi contornado da seguinte forma: primeiramente foi feita a exportação do resultado preliminar da busca na base para o formato CSV(o único suportado) resultando em 472 artigos. Após isto fez-se uso da ferramenta CSV2Bib¹ para converter o arquivo CSV para .bib com o intuito de importar o resultado, para a ferramenta Zotero², oque permitiu a execução da String de busca sobre os metadados dos 472 artigos encontrados preliminarmente, resultando em 6 documentos de interesse. A figura 4 apresenta um gráfico de barras contendo o número de artigos encontrados pela *string* de busca em cada uma das bases, já a figura 3 apresenta o percentual de publicações que cada uma das bases contribuiu para o montante final.

O gráfico 5 apresenta o número de publicações de interesse encontradas e cada uma das bases. O eixo X apresenta o ano do qual os artigos foram publicados e o eixo Y apresenta o número total de publicações em relação ao ano. Para a representação do gráfico foram removidas todas as publicações duplicadas. Podem-se perceber pela figura 5 que a partir do ano de 2015 há um considerável aumento no número de publicações, e ainda um grande pico no ano de 2017, demonstrando assim pontos de interesse neste período de publicações.

¹<https://github.com/jacksonpradolima/csv2bib>

²<https://www.zotero.org/>

Percentual de Publicações por Base

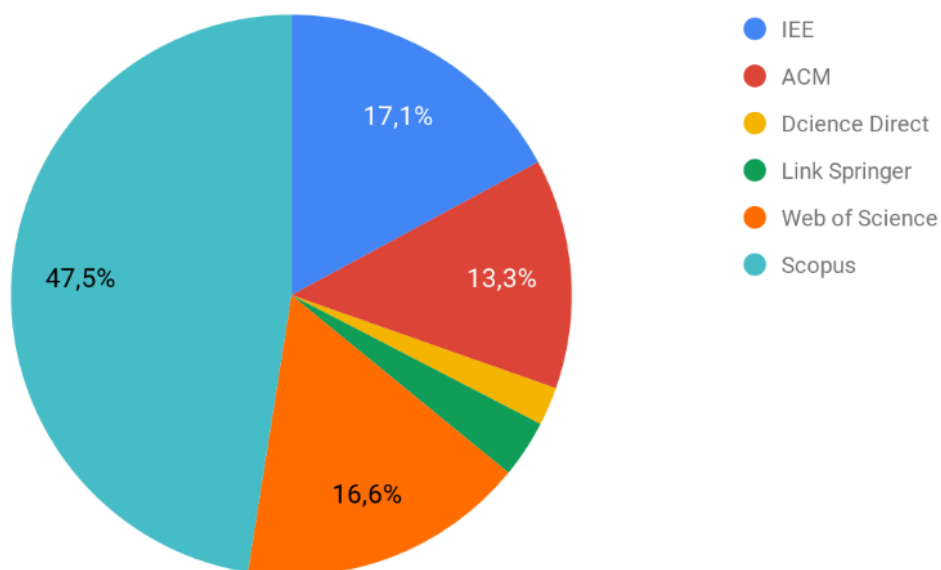


Figura 3 – Percentual de publicações encontradas por base.

3.1.1 Critérios de Inclusão e Exclusão

Após a seleção inicial realizada sobre as bases de dados, executou-se a triagem inicial sobre o resumo dos artigos, aplicando os seguintes critérios de inclusão e exclusão conforme a ordem apresentada abaixo:

- (E) Foi publicado antes de 2015;
- (E) Não é um artigo *full paper*;
- (E) Não estar em Inglês ou Português;
- (E) Indisponibilidade de acesso ao artigo completo;
- (E) Artigos que não apresentam avaliação da proposta;
- (I) Explora conceitos de segurança;
- (I) Explora conceitos de Computação Ubíqua;
- (E) O artigo não possui nenhum dos critérios de inclusão.

Para auxiliar na aplicação dos critérios de inclusão e exclusão foi feita a importação dos resultados preliminares das buscas na ferramenta Start³, para isso usou-se os arquivos .bib exportados pelas ferramentas das bases de busca, com exceção apenas da Spriger, onde usou-se o arquivo .bib exportado pelo Zootero, que foi gerado após

³http://lapes.dc.ufscar.br/tools/start_tool

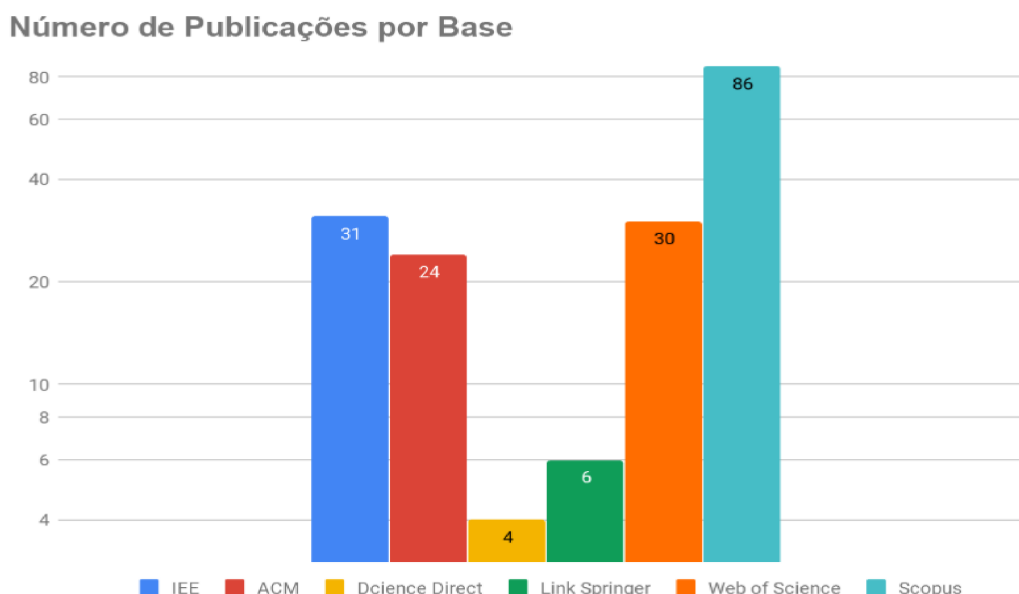


Figura 4 – Número de publicações encontradas por base.

a execução da consulta sobre os metadados, aplicada sobre o resultado preliminar da base.

Os critérios de exclusão foram aplicados seguindo a seguinte ordem e etapas:

- **Remoção de Trabalhos Duplicados** - Muitos dos trabalhos retornados pela *string* de busca estavam indexados em ambas as bases de consulta, tornando necessário a execução de uma etapa de remoção dos mesmos, resultando em 74 trabalhos duplicados removidos.
- **Filtro por Data** - O intervalo de interesse para a aplicação do filtro foi adotado com base no número de publicações por ano. Após o levantamento dos trabalhos de interesse, identificou-se o ano de 2015 como sendo o ano em que o número de publicações aumenta consideravelmente, continuando a ascender até o pico máximo no ano de 2017, como pode ser visto na figura 5. Assim optou-se por eliminar todas as publicações que fossem anteriores ao ano de 2015 eliminando desta forma 26 artigos.
- **Artigos Full Paper** - Com o intuito de remover artigos que apresentassem apenas resumos superficiais sobre os trabalhos, ou que não tivesse apelo científico, optou-se por remover artigos que não sejam Full Paper (livro ou capítulo de livro, introdução de anais, entre outros), sendo removidos 9 trabalhos.
- **Filtro por Idioma** - Como as pesquisas foram realizadas sobre várias bases de dados onde muitas destas indexam trabalhos em vários idiomas, optou-se por usar um filtro por idioma para remover qualquer trabalho que não esteja em

Número de Publicações por Ano

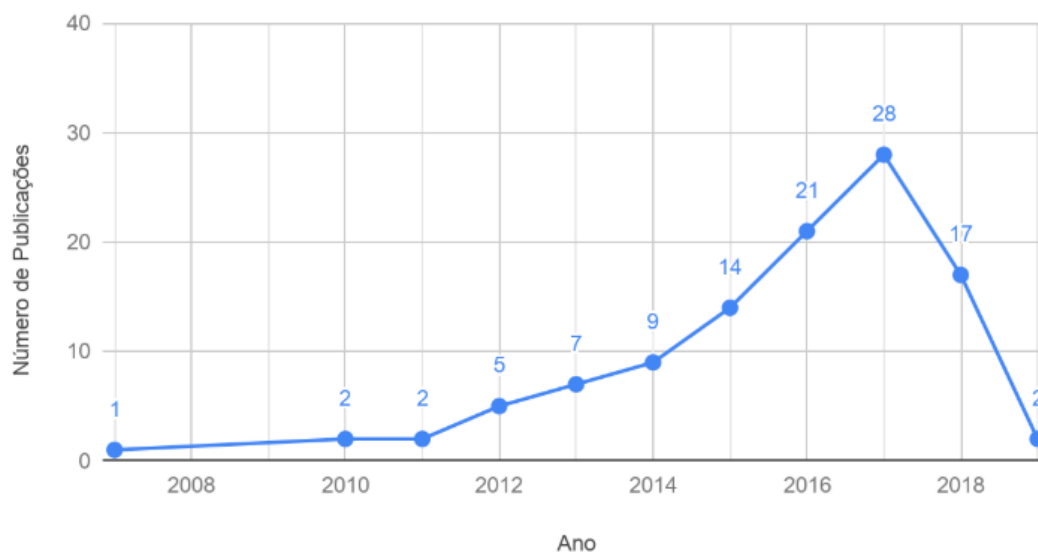


Figura 5 – Quantidade de publicações de interesse por ano.

Português ou Inglês (idiomas de total domínio do autor) removendo desta forma 1 artigo.

- **Indisponibilidade do Artigo completo** - Dado que alguns dos estudos de interesse selecionados apresentaram apenas seus resumos e introdução disponíveis não oferecendo a opção de obter-se o trabalho completo, optou-se por remover estes da pesquisa, sendo excluído 3 trabalhos.
- **Avaliação da Proposta** - Foram removidos todos os artigos que não executaram algum tipo de teste ou estudo de caso das soluções propostas por seus trabalhos, excluindo assim 17 artigos.
- **Sem Nenhum Critério de Inclusão** - Todos trabalhos que não se enquadraram em nenhum dos critérios de inclusão foram removidos, excluindo desta forma 28 trabalhos da pesquisa.

Após execução da triagem inicial dos trabalhos, aplicando os critérios de inclusão e exclusão sobre o resumo dos artigos, selecionou-se 24 documentos de interesse, o fluxo da aplicação dos critérios de exclusão pode ser visto na figura 6 assim como o número total de trabalhos removidos por cada um dos critérios de aplicação.

A execução da 4ª etapa do mapeamento, que consiste da triagem final dos trabalhos de interesse, selecionou dez dos 24 artigos para análise completa de seu conteúdo e da extração das informações destes. O critério considerado para a seleção dos dez trabalhos foi se estes exploravam conceitos de segurança da informação ou ainda se estes apresentavam o uso de conceitos de Computação Ubíqua.

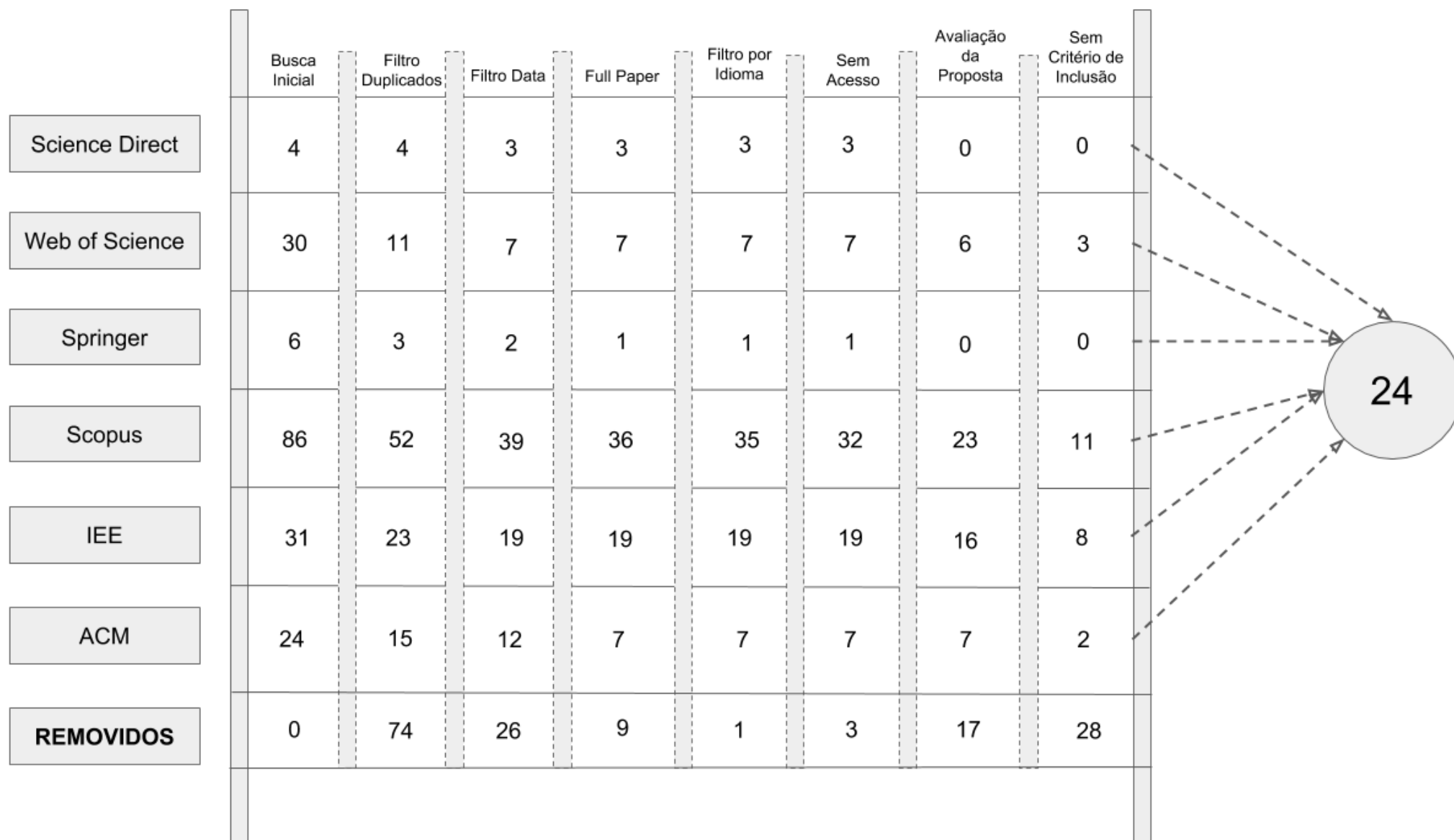


Figura 6 – Fluxo de triagem dos artigos.

3.2 Trabalhos Relacionados

Com a execução do mapeamento sistemático da literatura foram selecionados dez trabalhos de interesse, os quais serão apresentados nas subseções a seguir, onde serão abordados os seguintes tópicos de interesse: motivação e justificativa do trabalho, solução apresentada, avaliação e Resultados da Proposta.

3.2.1 Towards a Generalized Approach for Deep Neural Network based Event Processing for the Internet of Multimedia Things

- **Motivação e Justificativa do Trabalho:** O uso de dispositivos IoT multimídia tem aumentado consideravelmente (uso de câmeras para monitorar o tráfego de uma cidade por exemplo). Desta forma os tipos dos eventos criados na IoT estão também mudando, onde estes eventos multimídia geram dados não estruturados, gerando uma procura crescente na utilização eficiente do processamento de fluxos de eventos multimídia. No entanto, os mecanismos de processamento de eventos atuais têm suporte limitado ou inexistente para tipos de evento não estruturados (ASLAM; CURRY, 2018).
- **Solução Apresentada:** É proposto um sistema genérico para manipular eventos da Internet das Coisas Multimídia (IoMT) como um tipo de evento nativo em ferramentas de processamento de eventos com alta eficiência. O sistema proposto estende as linguagens de processamento de eventos com a introdução de operadores para análise multimídia de eventos não estruturados (eventos multimídia). A ferramenta proposta faz uso de um combinador de eventos baseado em redes neurais convolucionais profundas para executar o processamento de eventos de imagem e assim extrair recursos destas (ASLAM; CURRY, 2018).
- **Avaliação e Resultados da Proposta:** O sistema desenvolvido foi otimizado, com o uso de uma abordagem de seleção de classificadores baseado em restrições de assinatura. Os resultados obtidos pelos autores mostram que a ferramenta atinge uma taxa de transferência média de 110 quadros/segundo com uma precisão aproximada de 66,34% em eventos do mundo real de várias aplicações de cidades inteligentes. Por fim foi apresentado um teste de desempenho com o aumento gradativo do número de classes por classificador, onde os resultados obtidos mostram uma taxa de transferência estável para um classificador de uma classe, porém com o aumento do número destas classes a taxa de transferência cai continuamente (ASLAM; CURRY, 2018).

3.2.2 A Web-based Approach using Reactive Programming for Complex Event Processing in Internet of Things Applications

- **Motivação e Justificativa do Trabalho:** Nos últimos anos a Internet das coisas (IoT) tem crescido substancialmente, aumentando progressivamente o número de dispositivos conectados a rede, estima-se que em poucos anos cada um dos objetos de uso comum irá conter sensores para coletar e/ou fornecer algum tipo informação ou serviço para seus usuários, se conectando na Internet e gerando cada vez mais uma enorme quantidade de dados para serem trafegados pela rede. Esse crescente aumento no número de dispositivos e consequentemente a grande expansão no volume de dados, gera a necessidade que seja desenvolvida uma abordagem simples para que se possa lidar com esta nova grande avalanche de dados (ZIMMERLE; GAMA, 2018).
- **Solução Apresentada:** É proposto pelos autores a combinação de duas abordagens distintas para solucionar o problema citado por estes da nova grande avalanche de dados gerados pelas redes IoT: o CEP (*Complex Event Processing*) e WoT (*Web of Things*) com o uso de Ferramentas gráficas que exploram programação de fluxo (Mashups) o que irá permitir o uso de operadores CEP em eventos de alto níveis visualmente criados na plataforma WoT. Para o desenvolvimento dos operadores CEP utilizou-se Programação Reativa, os quais foram fornecidos como uma extensão da plataforma WoT Node-RED⁴, onde foram implementados três dos programas Coral8 amplamente referenciados para sistemas CEP, gerando a primeira extensão do CEP para a plataforma Node-RED, esta sendo uma ferramenta de programação construída em cima do Node.js, como protocolo de comunicação da troca de mensagens, os autores optaram por usar o MQTT⁵ (*Message Queuing Telemetry Transport*) (ZIMMERLE; GAMA, 2018).
- **Avaliação e Resultados da Proposta:** Para a avaliação da proposta do trabalho, um cenário de caso de uso é apresentado onde para este fim usou-se o simulador node-red-node-pi-sense-hatsimulator, desenvolvido pela equipe Node-RED que tem como objetivo reproduzir uma placa real Raspberry PI a qual incorpora alguns sensores e atuadores como LEDs, temperatura, pressão barométrica e sensores de umidade, entre outros, os quais foram usados para a simulação dos CEP da proposta. Com o desenvolvimento deste trabalho os autores citaram duas como as principais contribuições da pesquisa: (1) uma abordagem visual para construir consultas CEP para aplicativos de Internet e (2) o uso de Programação Reativa para detectar e acionar o CEP (ZIMMERLE;

⁴<https://nodered.org/>

⁵<https://mqtt.org/>

GAMA, 2018).

3.2.3 Semantic IoT Middleware-enabled Mobile Complex Event Processing for Integrated Pest Management

- Motivação e Justificativa do Trabalho:** Existem diversos desafios na agricultura moderna que são tipicamente encontrados no domínio dos Sistemas Ciber-Físicos (CPSs), dentre estes pode-se citar: conhecimento e deficiência de infraestrutura, informações incompletas, fontes limitadas de informação, perturbações externas (clima), autoridade de controle limitada (fertilizantes não podem fazer uma planta crescer arbitrariamente rápido). O gerenciamento agrícola moderno depende de muitas metodologias diferentes de sensoriamento para fornecer informações precisas sobre a cultura, o clima e as condições ambientais. Graças a miniaturização, da grande evolução e da difusão de sensores e recursos computacionais de baixo custo, tornou-se possível o desenvolvimento de dispositivos que produzem dados e que interagem entre si, produzindo assim uma rede de "coisas", gerando dados, processos e serviços interconectados. Desta forma os CPSs estão transformando a indústria agrícola (NOCERA et al., 2017).
- Solução Apresentada:** O trabalho propõe uma infraestrutura inteligente projetada para processar fontes de dados heterogêneas, como dados de sensores, dados meteorológicos e conhecimento agrícola coletado em uma ontologia, possibilitando uma comunicação mais suave e homogênea entre os dispositivos de uma infraestrutura dinâmica, configurável e extensível. A solução desenvolvida é baseada em processamento de eventos complexos (CEP) com o uso da ferramenta Esper⁶, onde um módulo é executado parcialmente em dispositivos móveis através da introdução do DeviceHive, um *middleware* da Internet das Coisas, ainda fez-se uso de uma linguagem de programação com suporte a mecanismos de reflexão, os quais servem como uma interface entre os componentes da IoT e o conhecimento ontológico. Para o sistema de comunicação entre o servidor e os dispositivos moveis é usado um *middleware* de publicação/assinatura composto por um DSB (*Distributed Service Bus*), um serviço do *Google Cloud Messaging* e um serviço da Web de assinatura. Segundo os autores a solução visa tornar-se um instrumento utilizado para conscientizar sobre o uso dos tratamentos agrícolas, onde o agricultor pode (i) ter acesso a todas as informações relacionadas ao domínio de interesse no momento necessário, (ii) criar um plano de defesa personalizado, (iii) receber alertas de mudanças nas condições climáticas e (iv) receber notificações e recomendações sobre seus planos de tratamento (NOCERA et al., 2017).

⁶<http://www.espertech.com/>

- **Avaliação e Resultados da Proposta:** Para a validação da proposta desenvolvida foi apresentado uma instanciiação de um cenário real, projetando também uma Ontologia OWL (*Ontology Web Language*) que codifica o conhecimento sobre aspectos relacionados à prática de Manejo Integrado de Pragas. Os autores executaram um primeiro conjunto de experimentos para validar a abordagem de fornecer a ferramenta para uma empresa e testar componentes individuais. Por fim os autores citaram como principais contribuições do trabalho: (i) a possibilidade de usar o motor CEP em tempo de execução do sistema, o que permite o monitoramento orientado a eventos e notificações de atualização, e (ii) sistema de modelagem com ontologias compreensíveis homem-máquina, que garante uma reconfiguração mais fácil a ferramenta (NOCERA et al., 2017).

3.2.4 Predictive Analytics for Complex IoT Data Streams

- **Motivação e Justificativa do Trabalho:** Os CEP são capazes de fornecer soluções escaláveis e distribuídas para lidar com fluxos de dados complexos em tempo real, no entanto, os CEPs não possuem a capacidade de realizar previsões assim como muitas das técnicas de aprendizado de máquina e análise estatística de dados. A grande parte dos aplicativos CEP disponíveis na literatura destina-se apenas a fornecer soluções reativas ao correlacionar fluxos de dados usando regras predefinidas, não explorando dados históricos devido a sua memória limitada. Existem diversos casos em que a previsão de um evento futuro é muito mais útil que apenas a detecção do mesmo, por exemplo, seria muito mais útil a previsão de um congestionamento em uma auto estrada do que sua detecção, já que com a previsão deste evento, com alguma antecedência, torna-se possível notificar aos administradores de tráfego, para que estes possam tomar as medidas preventivas de modo a evitar o congestionamento. Pode-se citar ainda diversos outros casos onde a previsão de eventos futuros podem trazer diversos ganhos como a previsão de desastres naturais e doenças epidêmicas (AKBAR et al., 2017).
- **Solução Apresentada:** Neste trabalho foi proposto uma arquitetura pró-ativa capaz de explorar dados históricos usando técnicas de aprendizado de máquina em conjunto com processamento de eventos complexos, de forma a combinar o poder do processamento de dados em tempo real do CEP com a capacidade de previsão de eventos das técnicas de aprendizado de máquina (*Machine Learning* - ML). Foi apresentado um algoritmo de previsão adaptativo chamado de AMWR (*Adaptive Moving Window Regression*) para dados dinâmicos de IoT, capaz de realizar previsões precisas quase que em tempo real, e ainda sendo capaz de trabalhar em conjunto com o CEP. Para a execução da proposta os autores fizeram uso das seguintes ferramentas: Node-RED para fornecer o Front-End da

arquitetura, o Apache Kafka⁷ como Broker de mensagens e por ultimo, a implementação foi elaborada em Python com o uso do módulo de aprendizagem de máquina scikit-learn (AKBAR et al., 2017).

- **Avaliação e Resultados da Proposta:** Para avaliação da proposta foi elaborado um caso de uso do mundo real onde dados de trafego de sistemas de transportes inteligentes foram usados para os testes, onde o algoritmo de predição foi capaz de atingir uma precisão de 96%, demonstrando assim a sua viabilidade de uso, já que com predições corretas sobre o trafego, como as que foram apresentadas nos testes, permitem que os administradores do sistema gerenciem o tráfego de uma maneira melhor, tomando decisões para evitar situações indesejadas, como congestionamentos por exemplo. Os autores do artigo citam como principais contribuições do trabalho: a implementação de uma arquitetura genérica baseada em componentes de código aberto para combinar ML com CEP, a fim de prever eventos complexos para aplicativos proativos de IoT; o desenvolvimento de um algoritmo de predição adaptativo para fluxos de dados dinâmicos de IoT que foi implementado em um caso de uso real do ITS atingindo uma precisão de até 96%. Também foi proposto um novo método para encontrar tamanho ótimo para janela de treinamento, explorando componentes espectrais de dados de séries temporais; A modelagem do erro introduzido pelo algoritmo de previsão usando uma distribuição paramétrica e a derivação em expressões para o erro global do sistema, à medida que o erro se propaga através do CEP (AKBAR et al., 2017).

3.2.5 DRESS: A Rule Engine on Spark for Event Stream Processing

- **Motivação e Justificativa do Trabalho:** Nos últimos anos o número de dispositivo conectados a rede vem aumentando, com esse crescimento, a quantidade de fluxos de dados aumenta simultaneamente, gerando a necessidade de sistemas capazes de reagir automaticamente a determinados eventos desencadeados por fluxos de dados. Tais sistemas se baseiam em um conjunto de regras pre definidas, onde através da análise dos fluxos de informações, executam determinadas ações que satisfaçam a alguma das regras deste conjunto. Nas últimas três décadas, sistemas como estes têm sido amplamente empregados em empresas, governos e organizações. Porém o aumento crescente no tamanho dos fluxos de dados, como o grande número de fluxos produzidos por eventos de dispositivos da Internet of Things (IoT), faz com que os atuais sistemas baseados em regras enfrentem sérios desafios em termos de velocidade, escalabilidade e tolerância a falhas (CHEN; BORDBAR, 2016).

⁷<https://kafka.apache.org/>

- **Solução Apresentada:** O artigo apresenta a proposta de adaptar sistemas baseados em regras para trabalhar em conjunto com o Spark Streaming visando melhorar seu desempenho. Os autores apresentaram uma Transformação do algoritmo Rete, que está por trás de muitos dos *Rule based systems* (RBSs) atuais, esta transformação faz com que o algoritmo funcione como um mecanismo de regras no ambiente do Spark. Também foi introduzido juntamente um novo sistema de mensagens baseado em Kafka o DRESS (*Distributed Rule Engine no Spark Streaming*) onde demonstrou-se uma forma automatizada de transformar regras escritas no estilo do Apache Drools para serem executadas no DRESS, tornando fácil para os atuais usuários do Drools mover seus sistemas para o DRESS sem esforço, este método de transformação de regras é baseada em técnicas MDA e na biblioteca de usuários SiTra (*Simple Transformer*) (CHEN; BORDBAR, 2016).
- **Avaliação e Resultados da Proposta:** O sistema proposto foi avaliado com a ajuda de um estudo de caso, onde simulou-se um sistema bancário para a execução dos testes. Os autores usaram o DRESS para transformar as regras CEP definidas para o ambiente de estudo, em código Scala e assim executá-lo no Spark Streaming. Um gerador de dados foi criado com o intuito de produzir informação aleatoriamente para a simulação do ambiente bancário, incluindo fluxos de caixa, contas e períodos contábeis com um parâmetro de escala. Durante os testes o DRESS demonstrou uma melhora significativa de desempenho e escalabilidade em comparação ao Drools, demonstrando ser capaz de lidar com grandes volumes de dados, em contra partida, o Drools não demonstrou esta mesma capacidade. Além da alta capacidade de processamento, o DRESS se demonstrou mais flexível em termos de gerenciamento de memória, mesmo nos testes executados em uma única máquina, este pode processar um conjuntos de dados maiores que o Drools e em menos tempo. Assim os autores destacam que com os dados coletados pelo estudo de caso, se pode demonstrar que o DRESS tem potencial para resolver muitos dos problemas de processamento de grande fluxo de dados presentes nos RBSs (CHEN; BORDBAR, 2016).

3.2.6 TrustCEP: Adopting a Trust-Based Approach for Distributed Complex Event Processing

- **Motivação e Justificativa do Trabalho:** O avanço da Internet das Coisas(IoT), com o uso de sensores modernos e dispositivos moveis capazes de capturar grandes quantidades de informações, estimulou o desenvolvimento de aplicativos aptos a trabalhar com essa nova grande avalanche de informações. Uma técnica eficaz que surgiu com o objetivo de extrair informações contextuais de alto nível deste grande fluxo de dados foi o CEP(*Complex Event Processing*), fa-

cilitando a análise de dados em tempo real provenientes de fontes heterogêneas e distribuídas. Considerando que o contexto dos usuários pode ser de informações sensíveis, a preservação da privacidade destes dados é crítica, tendo em vista que o processamento do contexto do usuário pode ocorrer em vários dispositivos (possivelmente maliciosos), especialmente em cenários colaborativos. Os trabalhos atuais sobre processamento de eventos complexos geralmente negligenciam o nível de privacidade dos dados do contexto de seus usuários onde estes são processados e diferentes dispositivos, muitas vezes com níveis de segurança desconhecidos (DWARAKANATH et al., 2017).

- **Solução Apresentada:** Para solucionar o problema de controle de privacidade, os autores propõem uma abordagem baseada em confiança, onde usam esta métrica de confiança definida para o posicionamento e a execução de operadores CEP em ambientes distribuídos, atribuindo o processamento de dados sensíveis para dispositivos que tenham um nível de confiança mais alto. Para a definição deste valor, a ferramenta pode usar o histórico de interação entre os dispositivos, ou ainda usar uma funcionalidade de recomendações de confiança, a qual faz uma verificação de similaridade baseada em cosseno, evitando assim ataques de collusion e on-off. Como as fontes de informações em um ambiente IoT são descentralizadas o modelo do sistema construído escolhido pelos autores foi entorno de uma rede device to device(D2D) (DWARAKANATH et al., 2017).
- **Avaliação e Resultados da Proposta:** Para a validação da proposta, um sistema CEP foi desenvolvido de forma distribuída baseado em *SmartPhones*, o qual possibilita aos usuários se comunicarem com uso de *Bluetooth* e processar gráficos de maneira distribuída. Esta ferramenta foi chamada de TrustCEP a qual foi usada para avaliar a abordagem. Para medir as relações de confiança, foi gerado um histórico de interações entre usuários em canais síncronos e assíncronos, os quais representam aspectos comportamentais da confiança dos usuários. Como métricas de comparação foram usados o consumo médio de energia e a troca de dados na rede, onde os autores observaram que com a implementação da proposta os *SmartPhones* usados para os testes apresentaram um leve aumento de 2-6% no consumo de energia, se comparado a abordagens quem não levam em consideração a privacidade dos dados, em contra partida o modelo proposto se mostra robusto contra ataques collusion e on-off. Os autores citam como principais contribuições de seu trabalho: o desenvolvimento de um modelo de gestão de confiança (descentralizada) para adaptar a disseminação de eventos e a colocação de operadores para o CEP distribuído; Foi introduzido um modelo de gestão de confiança baseado nas relações do usuário e no histórico de interação de comunicação; Foi apresentado um esquema de reco-

mendação de confiança robusto usando a medida de similaridade de cosseno (DWARAKANATH et al., 2017).

3.2.7 Anaysis of Controller Based IEEE 802.11 System with Similarity Measure Clustering

- Motivação e Justificativa do Trabalho:** A eficiência de um sistema WiFi que contenha dezenas de estações em uma área física pequena, em suma, é dada pela capacidade do sistema de alocar de forma ótima canais de rede para estas estações de forma a evitar conflitos de frequências nessas redes ao máximo. Com a evolução dos dispositivos moveis, com alta capacidade de tráfego de dados, sua grade popularização e seu uso acentuado em locais densamente povoados como por exemplo o uso de *smartphones* em grandes edifícios, se faz necessário o desenvolvimento ferramentas inteligentes capazes de oferecer bons níveis de QoS aos usuários. Para a configuração dos canais em um modo de operação normal uma rede WiFi usa o algoritmo de gerenciamento de recursos de rádio (RRM) onde este é executado periodicamente. Sendo definido vários valores de configurações nos pontos de acesso para iniciar tarefas de gerenciamento adicionais no controlador entre os períodos, as quais podem ser enxergadas como eventos complexos. Devido à complexidade, esses parâmetros para os valores do algoritmo RRM são normalmente fornecidos como valores padrões. Existindo uma falta significativa de experiências práticas sobre os operadores de serviço WiFi em busca dos valores ideais (GAL; TERDIK, 2017).
- Solução Apresentada:** Como proposta do trabalho os autores levantam as seguintes questões: Qual é o nível real de desempenho de um determinado sistema WiFi configurado como controlador? Quão sensato é este algoritmo RRM para futuros ataques de inundação nos canais de rádio em um AP ou terminal móvel? Com o intuito de responder a estas questões levantadas, os autores propõe a execução de uma análise profunda deste sistema onde estes usam uma abordagem de análise estatística baseada em mecanismos de agrupamento de comportamento e detecção de mudanças, fazendo a analise das informações coletadas para assim chegar a uma conclusão final. Alem disso os autores também propõem um novo método de clustering baseado em medidas de similaridade e aplicado nas redes Wifi IEEE 802.11 (GAL; TERDIK, 2017).
- Avaliação e Resultados da Proposta:** Para a validação da proposta do trabalho, foram executadas as medições do sistema WiFi na rede de informática da Universidade de Debrecen, os autores também ressaltam que as medições das duas tecnologias WiFi relativas as bandas de 2,4 GHz e 5 GHz foram analisadas separadamente. Após a execução dos testes sobre as redes wifi da universi-

dade, os autores concluem que o método de clustering baseado em comportamento proposto, é capaz de avaliar o desempenho do algoritmo de gerenciamento de recursos de rádio do controlador WiFi de forma eficiente, sendo capaz de informar os valores reais de desempenho que os AP produzem com a configuração do algoritmo padrão e gerar os valores ótimos mais adequados para aquela rede (GAL; TERDIK, 2017).

3.2.8 Parallel big data processing system for security monitoring in Internet of Things networks*

- Motivação e Justificativa do Trabalho:** Atualmente as redes de Internet das Coisas (IoT) tem se popularizado em diversas áreas. Assim como a sua popularização, a preocupação com a segurança dessas redes tem aumentado, levando ao interesse de desenvolvimento de sistemas de segurança sofisticados para a proteção destas redes, os quais são necessários já que o uso de sistemas de proteção tradicionais são de difícil ou impossível aplicação devido às peculiaridades para a construção e operação de redes IoT. Pode-se citar como fatores complicantes na implementação de sistemas de segurança nas redes IoT como: a necessidade de analisar grandes quantidades de dados em tempo real com o menor custo computacional possível, grande número de fontes de dados heterogêneas, computação limitada e recursos de energia limitado. Outro fator que destaca a importância de sistemas de segurança para redes IoT é a grande variedade de ataques cibernéticos existentes e a gravidade de suas consequências. Sistemas de informações de segurança e gerenciamento de eventos (SIEM) tem a capacidade de monitorar a segurança de redes por meio da coleta de dados sobre: eventos de interesse de dispositivos remotos, sensores de informação e seu processamento preliminar. Porém redes IoT possuem um grande número de tipos de fontes de dados, o que pode tornar extremamente complexo o monitoramento da segurança de rede devido a alta intensidade de fluxos de eventos, levando a necessidade do desenvolvimento de sistemas de segurança com capacidade de processamento de Big Data (KOTENKO; SAENKO; KUSHNEREVICH, 2017).
- Solução Apresentada:** Levando em consideração as limitações citadas para o desenvolvimento de sistemas de segurança para redes IoT, o trabalho propõe uma nova arquitetura de segurança para redes IoT baseada em um sistema de processamento paralelo distribuído de Big Data. A ferramenta de processamento de dados paralelo desenvolvido tem as seguintes características: devido ao uso do CEP, o sistema implementa funções básicas de pré-processamento em tempo real, as quais são: normalização de dados, filtragem de dados, agregação de dados e correlação de dados; os resultados do processamento preliminar dos

dados são fornecidos pela representação visual do sistema; a ferramenta é configurada para operar sob condições de limitações computacionais, inerentes aos elementos de rede da IoT. Para o desenvolvimento do sistema de processamento paralelo de dados de segurança, foi usado como base a ferramenta de código aberto Hadoop⁸ em conjunto com o ambiente de processamento de dados distribuído Apache Spark. A arquitetura do sistema também inclui componentes responsáveis pela coleta, armazenamento, agregação, normalização, análise e visualização de dados onde: a Agregação dos dados, normalização, análise e visualização são realizadas "on-the-fly"; Os dados são armazenados em um sistema de arquivos distribuídos do HDFS (*Hadoop Distributed File System*), proporcionando um aumento da confiabilidade do armazenamento e da velocidade com que as solicitações de dados são processadas (KOTENKO; SAENKO; KUSHNEREVICH, 2017).

- **Avaliação e Resultados da Proposta:** Para avaliação da proposta do artigo os fluxos de dados usados para os testes foram obtidos combinando fluxos de eventos de segurança em um fragmento de rede IoT com fluxos representados em um banco de dados externo de tráfego em uma rede real de computadores. A avaliação aplicada mostrou que, mesmo em um ambiente IoT com recursos computacionais limitados, quando o sistema é executado com o Hadoop, a ferramenta desenvolvida apresenta um desempenho razoavelmente alto, excedendo significativamente as implementações conhecidas, porém quando este é executado no Apache Spark a ferramenta mostrou um aumento de desempenho de cerca de dez vezes, caso o ambiente apresentar uma quantidade de memória RAM suficiente. Como principais contribuições do trabalho os autores citam: a execução de um comparativo de desempenho das plataformas Hadoop e Spark implementadas em um sistema de segurança de redes aplicado a IoT; o desenvolvimento de uma arquitetura destinada ao processamento paralelo e ao monitoramento de redes IoT (KOTENKO; SAENKO; KUSHNEREVICH, 2017).

3.2.9 An integrated information lifecycle management framework for exploiting social network data to identify dynamic large crowd concentration events in smart cities applications

- **Motivação e Justificativa do Trabalho:** Com o advento de novas tecnologias, como a Internet das Coisas e o processamento em nuvem, há uma grande diversidade de fontes de dados e serviços distintos disponíveis. O que nos leva a questão de como pode-se identificar maneiras inteligentes, abstratas e adaptativas de correlacionar e combinar os vários níveis de informações disponíveis

⁸<https://hadoop.apache.org/>

gerados por estas novas tecnologias. Com o cenário tecnológico atual pode-se destacar que o maior desafio dentre as questões levantadas é o de combinar diferentes fontes de dados heterogêneas de maneira inteligente, integrando e raciocinando sobre este fluxo de informação para inferir a consciência situacional, transformando desta forma o conjunto de dados brutos e sem significado para o nível de conhecimento e sabedoria dando um sentido semântico para estas informações. Para lidar com estes desafios, há necessidade do desenvolvimento de uma estrutura de *big data* de análise concisa capaz de analisar e lidar com o grande fluxo de dados gerado por ambientes IoT (KOUSIOURIS et al., 2018).

- **Solução Apresentada:** O objetivo do trabalho é o desenvolvimento de uma plataforma *Smart Transportation* capaz de identificar eventos de interesse do usuário em uma determinada área (como grandes eventos públicos com alta concentração de pessoas que possam afetar a jornada do usuário), a fim de enriquecer as informações no nível de aplicação com a identificação de eventos relacionados que podem permitir ações mais sofisticadas em nome deste usuário, a identificação é baseada na observação dos picos de atividade do Twitter em comparação com os dados históricos em um tempo dinâmico e local de interesse. O sistema proposto inclui os seguintes aspectos: integração entre dados de cidades inteligentes, proveniente de um sistema específico de monitoramento de passageiros da cidade de Madrid e dados de redes sociais provenientes do Twitter, para alertar passageiros com necessidades especiais e seus cuidadores sobre Grandes Concentrações de Multidões ao longo de sua jornada; Ingestão de dados integrados em soluções escaláveis baseadas em nuvem (OpenStack Swift) e integradas com ferramentas de análise de dados (Apache Spark) para trabalhar diretamente nos conjuntos de dados adquiridos e atuando com o processamento de eventos complexos, a fim de monitorar e emitir alertas de dados de fluxo. Ainda a solução deverá ser capaz de se adaptar a dados heterogêneos ajustando ao modelo Apache AVRO; Conexão de todos esses sistemas com uma lógica baseada em aplicações, obtida através de uma camada de *middleware* baseada no NodeRED, a fim de orquestrar as ações necessárias no fluxo de dados previsto e fornecer as adaptações necessárias em termos de protocolos e formatos de dados (KOUSIOURIS et al., 2018).
- **Avaliação e Resultados da Proposta:** Para validação do sistema proposto, a ferramenta foi aplicada em um experimento de dois meses na cidade de Madri, em torno de dois locais de eventos públicos, atuando na identificação de eventos de grandes concentrações de pessoas como eventos esportivos e analisando várias abordagens com relação à definição de limiares necessários como: estrutura integrada e automatizada, incluindo serviços da plataforma *Smart City* e

Twitter. Com este estudo de caso os autores obtiveram 49 erros de 2042 casos de avaliação, onde destes erros existiram apenas 4 falsos negativos. Porém como o sistema é flexível, este pode ser ajustado para especificações mais rígidas objetivando diminuir falsos positivos ou falsos negativos dependendo da preferencia do usuário. Ainda os autores citam como principais contribuições do trabalho: o desenvolvimento de um *middleware* para adaptação e sincronização necessária do sistema, a execução de um experimento de 2 meses na cidade de Madri, o desenvolvimento de um sistema de identificação detalhado da concentração de multidões por meio da localização dada com o uso de dados de redes sociais (Twitter), a identificação de limiares ideais para ignorar intervalos de tempo baseando-se em limites gerais de funções de distribuição cumulativas (CDF) (KOUSIOURIS et al., 2018).

3.2.10 CEML: Mixing and moving complex event processing and machine learning to the edge of the network for IoT

- Motivação e Justificativa do Trabalho:** A internet das coisas é um campo em constante crescimento. As previsões mostram que em 2020 mais de 50 bilhões de dispositivos conectados, levando a uma avalanche de informações gerada constantemente por estes dispositivos, onde muito deste grande aumento na quantidade de informação gerada se dá graças ao desenvolvimento de hardwares de baixa potência e protocolos de rede mais eficientes. Atualmente já existem muitos setores da indústria que acumulam um grande volume de informação (por exemplo serviços investimento, mídia, bancos) as quais poderiam ser usadas por aplicações de exploração para gerar algum tipo de benefício para estas entidades. Sistemas de análise em nuvem altamente escalonáveis estão sendo usados para lidar com esta explosão de dados gerada pelas redes IoT. Porém devido a natureza onipresente desses dados, á novos requisitos técnicos e não técnicos que são de difícil resolução com uma implantação em nuvem. Precisamos de um novo conjunto de tecnologias para resolver estes problemas, como mineração de dados distribuída e mineração de dados ubíqua desenvolvidas e otimizadas especialmente para a aplicação IoT (SOTO et al., 2016).
- Solução Apresentada:** Os autores propõem um novo *framework* denominado *Complex Event Machine Learning* (CEML) o qual apresenta um conjunto de ferramentas para aprendizado de máquina, distribuído automaticamente em ferramentas de avaliação contínuas automáticas em tempo real e gerenciamento automático de regras para a implementação de regras, onde estes recursos são desenvolvidos para uma implantação de borda de rede ou em nuvem. A ferramenta desenvolvida pelos autores apresenta os seguintes aspectos: Capaz de lidar com fluxos contínuos de dados através das instruções de fluxo do mecanismo

CEP; não há requisitos de memória o qual é permitido pelo uso de mecanismos CEP em janelas de fluxo permitindo o uso inteligente da memória conforme necessário; o sistema é capaz de transferir resultados de mineração por uma rede sem fio com largura de banda limitada tratado pelo protocolo MQTT com o uso do Mosquitto Broker⁹; a ferramenta é capaz de modelar as mudanças dos dados ao longo da execução já que o CEML é um modelo de aprendizagem de máquina os modelos usados por este se ajustam conforme aprendem; Um ambiente de mineração iterativo fornecido por uma API REST; Integração entre sistemas de gerenciamento de fluxo de dados e abordagens de mineração através da API REST. A ferramenta proposta foi desenvolvida a partir do Data-Fusion Manager, um micro-serviço LinkSmart® baseado em Java para *Smart Cities*, o desenvolvimento do sistema de coleta de dados foi feito usando o broker Mosquitto e o cliente Paho Java, por fim para a implementação dos algoritmos de inteligência artificial foi usado a API do Weka (SOTO et al., 2016).

- **Avaliação e Resultados da Proposta:** Para a avaliação do sistema proposto, foi executada a monitoração do desempenho da ferramenta em uma implantação incorporada em um problema de classificação de tempo real aplicada para melhorar o sistema de detecção de presença na Universidade Federal de Pernambuco (UFPE) o qual apresentou um desempenho satisfatório durante os testes. Os autores validaram a proposta também em um ambiente de mais fácil reprodução, assim estes usaram um conjunto de dados da íris onde cada um destes dados foi enviado pela rede, assim como se fosse capturado por algum sensor, onde estas informações são recebidas por um Broker implantado em uma Raspberry PI2 e processadas no mesmo pela ferramenta proposta de aprendizado, as medidas foram enviadas a uma média de 19,17 por segundo, sendo usado o algoritmo NaiveBayesUpdateable que foi capaz de apresentar bons resultados durante as medições (SOTO et al., 2016).

3.3 Discussão dos Trabalhos Relacionados

A figura 7 apresenta uma análise comparativa entre os trabalhos selecionados pelo mapeamento sistemático realizado por este trabalho, onde os campos selecionados para a comparação foram:

- **Paradigma de Processamento de Eventos:** apresenta qual tipo de técnica de processamento de eventos o artigo emprega em sua proposta, esta coluna pode apresentar dois valores CEP ou ESP.

⁹<https://mosquitto.org/>

- **Uso de Técnicas em Aprendizado de Máquina:** Esta coluna pode apresentar somente dois valores: Sim que se refere ao uso de alguma técnica qualquer de aprendizado de máquina e - onde o autor não fez uso de nenhuma técnica de inteligência artificial.
- **Ferramentas:** nesta coluna é mostrado na figura 7 quais foram as principais ferramentas citadas nos documentos de cada uma das propostas analisadas.
- **Cenário:** os artigos selecionados são separados pelas áreas que estes propuseram uma solução das quais foram identificadas: IoT-Multimídia soluções para fluxos de dados de eventos não estruturados; IoT-Big Data soluções para lidar com o grande fluxo de dados gerado pelas redes IoT; Agricultura de Precisão soluções que são aplicadas de alguma forma a algum meio da agronomia; IoT-ML soluções que aplicam aprendizado de máquina para solucionar e oferecer ferramentas novas a redes IoT; *Smart Cities* soluções que são aplicadas em cidades inteligentes; Segurança soluções que usam processamento de evento para resolver algum problema de segurança genérico.
- **Reusabilidade:** demonstra se o trabalho apresenta detalhes suficientes que permitem a reimplementação da proposta, sendo passível de replicação dos testes realizados. Esta coluna pode ter três valores: - não é possível fazer a reimplimentação da proposta; Sim onde é possível fazer a reimplimentação da proposta e dos testes executados pelo autor; Parcial onde é possível fazer a reimplimentação da proposta porém os testes não são passíveis de reimplimentação.

A partir da análise da figura 7 identificou-se que dentre os artigos selecionados durante o mapeamento apenas dois destes, (ASLAM; CURRY, 2018) e (CHEN; BORDBAR, 2016) trabalhavam com ESP, todos os demais trabalhos aplicam CEP de alguma forma em suas soluções. As técnicas de aprendizado de máquina também não foram usadas por muitos dos artigos selecionados, apenas quatro documentos fizeram uso destas técnicas para solucionar os problemas abordados, os quais foram (ASLAM; CURRY, 2018), (AKBAR et al., 2017), (KOUSIOURIS et al., 2018) e (SOTO et al., 2016).

Um grande problema identificado dentre os dez documentos selecionados foi a questão de reusabilidade, apenas (ZIMMERLE; GAMA, 2018) apresentou uma solução passível de reimplimentação e reutilização completa, as demais não demonstraram detalhamento suficiente em suas propostas para que terceiros pudessem fazer a reimplimentação de suas abordagens e a remontagem das validações executadas, demonstrando que as soluções não podem ser verificável e observável algo imprecisável em um documento científico.

	Paradigma em EP	Uso de Técnicas em ML	Ferramentas	Cenário	Reusabilidade
ASLAM;CURRY, 2018	ESP	Sim	-	IOT-Multimídia	-
ZIMMERLE; GAMA, 2018	CEP	-	SQL, javascript, Node-RED	IOT-Big Data	Sim
NOCERA et al., 2017	CEP	-	NODE.js, Redis, Apache Esper	Agricultura de Precisão	Parcial
AKBAR et al., 2017	CEP	Sim	Node-RED, Apache Kafka, Apache Esper, Python	IOT-ML	Parcial
CHEN; BORDBAR, 2016	ESP	-	Apache Kafka, Apache Spark	IOT-Big Data	Parcial
DWARAKANATH et al., 2017	CEP	-	-	IOT-Security	-
GAL; TERDIK, 2017	CEP	-	-	Security	-
KOTENKO; SAENKO; KUSHNEREVICH, 2017	CEP	-	Apache Spark, Hadoop, Java, MySQL	IOT-Big Data-Security	Parcial
KOUSIOURIS et al., 2018	CEP	Sim	Apache Spark, Node-RED, Openstack Swift	Smart Cities	Parcial
SOTO et al., 2016	CEP	Sim	Java, Mosquitto, Weka	IOT-ML	Parcial

Figura 7 – Comparativo entre os artigos selecionados.

3.4 Considerações do Capítulo

Neste capítulo foi executado um mapeamento sistemático com o objetivo de identificar o estado da arte em processamento de eventos para redes IoT. Como resultado deste mapeamento identificou-se alguns trabalhos de interesse onde foi elaborado um detalhamento de suas motivações e justificativas para a elaboração do mesmo, também foi detalhado a solução que cada um destes trabalhos propuseram e por ultimo foi apresentado os resultados que os autores obtiveram com o desenvolvimento destes trabalhos. Por fim foi demonstrada uma comparação entre as tecnologias empregadas em cada um destes trabalhos e apresentando criticas e desafios identificados durante esta revisão.

4 CONSIDERAÇÕES FINAIS

O presente trabalho teve como objetivo fazer o levantamento dos trabalhos científicos elaborados sobre processamento de eventos aplicados a internet das coisas, onde para se atingir este objetivo foi executado um mapeamento sistemático o qual seguiu o algoritmo proposto por Petersen et al. (2008). Após a execução do mapeamento sistemático identificou-se dez trabalhos de interesse onde foi elaborada uma discussão comparando estes trabalhos entre si, onde identificou-se que a maioria dos autores destes trabalhos não se preocuparam com a reusabilidade de suas propostas o que dificulta a elaboração de novos trabalhos que façam uso ou mesmo que se baseiem nestas soluções.

REFERÊNCIAS

AGRAWAL, S.; VIEIRA, D. A survey on Internet of Things. **Abakos**, [S.l.], v.1, n.2, p.78–95, 2013.

AKBAR, A.; KHAN, A.; CARREZ, F.; MOESSNER, K. Predictive analytics for complex IoT data streams. **IEEE Internet of Things Journal**, [S.l.], v.4, n.5, p.1571–1582, 2017.

APPEL, S.; FRISCHBIER, S.; FREUDENREICH, T.; BUCHMANN, A. Event stream processing units in business processes. In: **Business Process Management**. [S.l.]: Springer, 2013. p.187–202.

ASLAM, A.; CURRY, E. Towards a Generalized Approach for Deep Neural Network Based Event Processing for the Internet of Multimedia Things. **IEEE Access**, [S.l.], v.6, p.25573–25587, 2018.

CHEN, Y.; BORDBAR, B. Dress: A rule engine on spark for event stream processing. In: IEEE/ACM INTERNATIONAL CONFERENCE ON BIG DATA COMPUTING, APPLICATIONS AND TECHNOLOGIES, 3., 2016. **Proceedings...** [S.l.: s.n.], 2016. p.46–51.

DAYARATHNA, M.; PERERA, S. Recent advancements in event processing. **ACM Computing Surveys (CSUR)**, [S.l.], v.51, n.2, p.33, 2018.

DWARAKANATH, R. et al. Trustcep: Adopting a trust-based approach for distributed complex event processing. In: IEEE INTERNATIONAL CONFERENCE ON MOBILE DATA MANAGEMENT (MDM), 2017., 2017. **Anais...** [S.l.: s.n.], 2017. p.30–39.

ETZION, O.; NIBLETT, P.; LUCKHAM, D. C. **Event processing in action**. [S.l.]: Manning Greenwich, 2011.

FITZGERALD, E. et al. Common Event Expression (CEE) Overview. **Report of the CEE Editorial Board**, [S.l.], 2010.

GAL, Z.; TERDIK, G. Anaysis of controller based IEEE 802.11 system with similarity measure clustering: Evaluation of channel allocation efficiency by change detection.

In: DIGITAL FORENSIC AND SECURITY (ISDFS), 2017 5TH INTERNATIONAL SYMPOSIUM ON, 2017. **Anais...** [S.l.: s.n.], 2017. p.1–6.

GONÇALVES, A. R. S. M. **Research of the internet of things business models in Portugal**. 2017. Tese (Doutorado em Ciência da Computação) — .

JUN, C.; CHI, C. Design of complex event-processing ids in internet of things. In: MEASURING TECHNOLOGY AND MECHATRONICS AUTOMATION (ICMTMA), 2014 SIXTH INTERNATIONAL CONFERENCE ON, 2014. **Anais...** [S.l.: s.n.], 2014. p.226–229.

KOTENKO, I. V.; SAENKO, I.; KUSHNEREVICH, A. Parallel big data processing system for security monitoring in Internet of Things networks. **JoWUA**, [S.l.], v.8, n.4, p.60–74, 2017.

KOUSIOURIS, G. et al. An integrated information lifecycle management framework for exploiting social network data to identify dynamic large crowd concentration events in smart cities applications. **Future Generation Computer Systems**, [S.l.], v.78, p.516–530, 2018.

MINBO, L.; ZHU, Z.; GUANGYU, C. Information service system of agriculture IoT. **automatika**, [S.l.], v.54, n.4, p.415–426, 2013.

NOCERA, F.; DI NOIA, T.; MONGIELLO, M.; DI SCIASCIO, E. Semantic IoT Middleware-enabled Mobile Complex Event Processing for Integrated Pest Management. In: CLOSER, 2017. **Anais...** [S.l.: s.n.], 2017. p.610–617.

PETERSEN, K.; FELDT, R.; MUJTABA, S.; MATTSSON, M. Systematic Mapping Studies in Software Engineering. In: INTERNATIONAL CONFERENCE ON EVALUATION AND ASSESSMENT IN SOFTWARE ENGINEERING, 12., 2008, Swindon, UK. **Proceedings...** BCS Learning & Development Ltd., 2008. p.68–77. (EASE'08).

SANCHEZ, G. Wireless sensor network deployment for integrating video-surveillance and data-monitoring in precision agriculture over distributed crops. **Computers and Electronics in Agriculture**, [S.l.], v.75, n.2, p.288–303, 2011.

SOTO, J. A. C.; JENTSCH, M.; PREUVENEERS, D.; ILIE-ZUDOR, E. CEML: Mixing and moving complex event processing and machine learning to the edge of the network for IoT applications. In: INTERNATIONAL CONFERENCE ON THE INTERNET OF THINGS, 6., 2016. **Proceedings...** [S.l.: s.n.], 2016. p.103–110.

VIJAYARAGHAVAN, A.; DORNFELD, D. Automated energy monitoring of machine tools. **CIRP annals**, [S.l.], v.59, n.1, p.21–24, 2010.

WEINER, H. S. et al. **Health care patient status event processing and reporting**. [S.l.]: Google Patents, 2008. US Patent 7,439,856.

WEISER, M. The computer for the 21st century. **Mobile Computing and Communications Review**, [S.l.], v.3, n.3, p.3–11, 1999.

WU, E.; DIAO, Y.; RIZVI, S. High-performance complex event processing over streams. In: ACM SIGMOD INTERNATIONAL CONFERENCE ON MANAGEMENT OF DATA, 2006., 2006. **Proceedings...** [S.l.: s.n.], 2006. p.407–418.

XAVIER, M. S. R. d. B. **Smart Homes no mercado downstream de Oil & Gas**. 2016. Dissertação (Mestrado em Ciência da Computação) — FEUC.

ZIMMERLE, C.; GAMA, K. A web-based approach using reactive programming for complex event processing in internet of things applications. In: ANNUAL ACM SYMPOSIUM ON APPLIED COMPUTING, 33., 2018. **Proceedings...** [S.l.: s.n.], 2018. p.2167–2174.