

Fornecimento de dados pela IoT, e seu processamento por IA
Universidade Federal de Santa Catarina Ciências da Computação
Agosto de 2018

Matheus Henrique Schaly – 18200436

Título:

Fornecimento de dados pela IoT e seu processamento por IA

Resumo:

A rede da Internet das Coisas (IoT) está em expansão. Com a sua ampliação adquirimos a possibilidade de coleta de grandes volumes de dados (Big Data). Porém, a análise de grandes volumes de dados requer um processamento agilizado, portanto para auxiliar na eficiência do gerenciamento de dados entre a fonte de dados e a nuvem, pode-se usar a tecnologia de Fog Computing. Além disso, os dados coletados pela IoT podem ser usados dentro do campo da Inteligência Artificial e Aprendizagem de Máquina (ML) para gerar informações valiosas. Essas informações podem ser geradas de inúmeras maneiras e possuem várias aplicações distintas. Por exemplo, câmeras ligadas na rede de IoT geram imagens, que por sua vez possuem pixels, e esses pixels podem servir como dados para prever um acidente, analisar o tráfego, detectar movimentos suspeitos, detectar assaltos, reconhecer placas, reconhecer pessoas... Além disso, dispositivos de GPS podem ser transportados por veículos, pessoas, animais... gerando dados espaço-temporais capazes de melhorar o turismo, segurança, marketing, tráfego entre outros aspectos. Portanto, há inúmeras maneiras de coleta de dados a partir de dispositivos conectados à rede, e várias possíveis aplicações para tais dados. Proporcionando um campo de estudo promissor com diversas aplicações práticas.

1. Introdução:

1.1. Motivação:

Internet das Coisas já está embutida em nossas vidas através de smart TVs, smartphones, assistentes de atividade física, sistema de monitoramento de saúde, relógios inteligentes, auxiliares de tráfego em cidades entre outros. Cerca de 6.5 bilhões de dispositivos estavam ligados a IoT em 2016. Estima-se que em 2020 aproximadamente 25 bilhões de dispositivos estarão ligados à internet. Vários setores poderão se beneficiar com a IoT, incluindo principalmente o setor de indústrias, energia, saúde e seguros. Já há empresas que estão alocando seus recursos para o setor de IoT. Entretanto, a enorme quantidade de dados coletados pela IoT são impraticáveis de serem usados manualmente. Portanto, encontra-se a necessidade de unir a quantidade massiva de dados gerados, há uma ferramenta que faça sentido destes dados, neste caso, tal ferramenta seria a IA. O restante deste artigo é organizado da seguinte maneira: Seção II, conceitos básicos: Breve introdução para melhor entendimento dos conceitos de IoT, Fog Computing, big data, IA e ML. Na seção III, trabalhos correlatos: Será apresentado uma breve análise dos artigos que foram utilizados como base para

elaboração do presente artigo. Seção IV, aspectos relevantes: Será destacado os fatores importantes em relação ao artigo, que são abordados nos artigos referenciados, tais como IoT, Fog e IA. Seção V, problemas existentes: Destacará os problemas enfrentados pela coleta e análise de dados nas pesquisas atuais. Seção VI, soluções possíveis: Investigará possíveis soluções aos problemas encontrados nas seção V. Seção VII, conclusões e trabalhos futuros: Resumirá a ideia principal do

1.2. Justificativas:

Há vários cenários em que as enormes quantidades de dados gerados pela IoT e manipulados por IA podem gerar informações para melhorar aspectos relacionados, por exemplo, à segurança, fabricação de produtos, tráfego, agricultura, saúde, reconhecimento de imagem, reconhecimento de voz, trajetórias entre inúmeros outros aspectos. Ao usar dados de GPS podemos detectar o estacionamento ilegal de veículos ^[1], consequentemente colaborando com a efetividade das políticas relacionadas ao gerenciamento de cidades e planejamento urbano por parte do governo. Podemos também usar pixels e trajetórias de pessoas a partir de câmeras ligadas à rede para reconhecermos eventos anormais ^[2], auxiliando na detecção de assaltos, brigas, ou ações suspeitas em geral. Portanto, o uso aliado das tecnologias de IoT para coleta de dados, Fog e big data para melhor gerenciamento da corrente de dados assim como IA e ML para processamento de tais dados, poderia ser útil para diversos aspectos da sociedade.

1.3. Objetivos:

1.3.1 Objetivos Gerais:

Compreender a importância da utilização dos dados gerados pela IoT para geração de informação. Além disso, o artigo analisará métodos para a utilização dos dados adquiridos pela IoT para problemas reais com a utilização de IA e ML. Ferramentas relevantes para o funcionamento do sistema completo também serão abordados, tais como: big data e fog.

1.3.2 Objetivos Específicos:

- Demonstrar exemplos específicos da utilização interdependente de IoT, Fog Computing, big data, IA e ML.
- Explorar particularmente dados de trajetórias unidos às tecnologias citadas acima.

2. Conceitos básicos:

2.1. Internet das Coisas:

Em inglês Internet of Things (IoT) e em português Internet das Coisas pode ser visto como um avançado sistema de automação e análise que explora redes, sensores, big data e inteligência artificial para entregar sistemas completos para um produto ou serviço. Tais sistemas permitem maior transparência, controle e performance quando aplicado em algumas indústrias ou sistemas. Internet das Coisas permite a aprimoração de coleta de dados,

automação, operações, e muito mais através de dispositivos inteligentes ligados à rede.

2.2. Fog Computing:

Em inglês Fog Computing e em português “computação em névoa” é uma infraestrutura de computação descentralizada na qual os dados, armazenamento e aplicações são distribuídas de forma mais lógica e eficiente entre a fonte de geração de dados e a nuvem. Fog Computing essencialmente estende a computação em nuvem e serviços no limite da rede, trazendo as vantagens e o poder da nuvem para mais próximo de onde os dados são gerados. O objetivo do Fog Computing é para melhorar a eficiência e reduzir a quantidade de dados para processamento, análise e armazenamento na nuvem. Isso é usualmente feito para melhorar a eficiência, mas também pode ser usado por razões de segurança e conformidade. As aplicações de Fog Computing incluem smart cities, smart grid, smart building, rede de veículos e software-defined networking.

2.3. Big Data:

Em inglês big data e em português “grande volume de dados” é qualquer quantidade significativa de dados estruturados, semiestruturados ou não estruturados que apresentam o potencial de serem minerados para gerar informação. Big data é normalmente caracterizado por volumes de dados extremos, grande variedade de tipos de dados e a velocidade que tais dados precisam ser processados. Apesar de big data não poder ser definido por um certo tamanho de bytes, ele está no campo de terabytes, petabytes ou até mesmo exabytes de dados capturados no decorrer do tempo.

2.4. Inteligência Artificial:

Inteligência Artificial (IA) é a simulação de processos de inteligência humana a partir de máquinas, especialmente sistemas de computadores. Tais processos incluem o aprendizado (aquisição de informação e regras para usar a informação), raciocínio (usar as regras para chegar a conclusões aproximadas ou definitivas) e autocorreção. Algumas das aplicações de IA incluem reconhecimento de voz e reconhecimento de imagem. IA ganhou mais destaque recentemente, em parte, por conta de big data e melhores hardwares capazes de processar grandes quantidades de dados. Alguns setores em que IA pode ser utilizado inclui assistência médica, negócios, educação, finanças, leis e manufaturas.

2.5. Aprendizagem de Máquina:

Em inglês Machine Learning (ML) e em português Aprendizagem de Máquina é o que proporciona o computador a habilidade de aprender sem ter sido explicitamente programado, gerando a capacidade de criar algoritmos que pode aprender a partir de dados e gerar previsões. Isso possibilita a rede de IoT a analisar dados, procurar por correlações e determinar a melhor ação a ser tomada. O sistema checka continuamente a qualidade de suas previsões e continua a refinar seu próprio algoritmo. Pode ser dividido em dois principais tipos de aprendizagem: Aprendizagem Supervisionada no qual os dados possuem rótulos dos resultados que deveriam ser gerados, e a Aprendizagem Não-Supervisionada em que os dados

não possuem tais rótulos.

3. Trabalhos Correlatos: Revisão bibliográfica sistemática:

Palavra-chave	Total
“IoT”	733,000
“IoT” “Fog”	12,400
“IoT” “Fog” “Big Data”	4,400
“IoT” “Fog” “Big Data” “Artificial Intelligence”	1,080
“IoT” “Fog” “Big Data” “Artificial Intelligence” “Machine Learning”	713
“IoT” “Fog” “Big Data” “Artificial Intelligence” “Machine Learning” “Trajectory Data”	8

Ao observar a tabela acima, nota-se que há apenas 8 referências no Google Scholar que abrangem IoT, Fog, big data, IA, ML e dados de trajetórias sendo que 5 desses 8 são artigos. Tais artigos são majoritariamente recentes, sendo 3 deles de 2018 e 2 de 2015. Observa-se também que há um grande interesse acadêmico na área de IoT, com um total de 733,000 artigos referenciando o tema. Além disso, apenas 713 trabalhos sobre os temas IoT, Fog, big data, IA e ML, que são áreas que têm seu próprio escopo, mas que podem ser usadas em conjunto.

Os trabalhos escolhidos para a realização do presente artigo estarão devidamente referenciados em “referências bibliográficas”. Alguns desses artigos serão detalhados a seguir.

3.1. “Introducing the new paradigm of Social Dispersed Computing: Applications, Technologies and Challenges” ^[3]

Neste artigo os autores analisam alguns paradigmas da computação como: nuvem, fog, edge e mobile edge, solucionando a ambiguidade de tais definições. Os autores também analisam e discutem as tecnologias de alguns softwares fundamentais, os desafios restantes e oportunidades de pesquisa. O artigo define uma aplicação chamada social computing applications que são aplicações inteligentes aonde os resultados recebidos pelos usuários finais ou a performance que eles experimentam é afetado pelos outros usuários da mesma aplicação. Exemplos desses tipos de roteamento de tráfego incluem o Waze e o Google. Porém, encontra-se atraso no social computing applications causado pela propagação de dados através do backhaul que não são adequados para aplicações que requerem resposta em tempo real ou alta qualidade de serviço. Esses problemas também são analisados ao decorrer do artigo, sendo que parte da solução se encontra na infraestrutura fog, fazendo com que usuários possam deixar tarefas pesadas (como rodar um algoritmo de ML) para locais próximos aos usuários que possuam maior poder de processamento. Trabalhos futuros incluem superar os desafios encontrados no artigo, que incluem: mobilidade, restrições de recursos dos sistemas, sincronização e dessincronização do serviço e, por último, a qualidade das demandas do serviço.

3.2. “A foundational framework for smart sustainable city development: Theoretical, disciplinary, and discursive dimensions and their synergies” ^[4]

Neste artigo os autores exploram cidades inteligentes sustentáveis, que é uma área científica densa e complexa. O Foco do artigo foca em numerosas teorias fundamentais junto a disciplinas acadêmicas, com o objetivo de formar uma estrutura que analiticamente relaciona o desenvolvimento de cidades, sustentabilidade e Tecnologias da Informação e Comunicação (TIC), ao mesmo tempo em que enfatizam como e em que medida sustentabilidade e TIC tornaram-se particularmente influentes no desenvolvimento de cidades na sociedade moderna. O artigo também foca na ideia de interdisciplinaridade e transdisciplinaridade na questão de desenvolvimento de cidades inteligentes sustentáveis, as quais incluem campos de design urbano, planejamento urbano, sustentabilidade, desenvolvimento sustentável, ciência de dados, ciências da computação, teoria de sistemas TIC, entre outros. Portanto o conceito básico seria de alavancar o potencial e onipresença das TIC na transição necessária para o desenvolvimento de cidades inteligentes sustentáveis em um mundo computadorizado e urbanizado.

3.3. “Detecting Vehicle Illegal Parking Events using Sharing Bikes’ Trajectories” ^[1]

Neste artigo os autores analisam dados coletados pelo GPS que está contido nas bicicletas do sistema Mobike. Mobike é uma estação de bicicletas que visa o compartilhamento das mesmas. O artigo usa os dados coletados para detectar estacionamento ilegais. Dois componentes principais são empregados no sistema de detecção: 1) o pré-processamento de trajetórias que filtra os pontos excepcionais de GPS, a execução da correspondência de ruas no mapa e a construção de índices de trajetórias; e 2) detecção de estacionamento ilegal, que modela as trajetórias normais, extrai características das trajetórias de avaliação e utiliza o método de distribuição baseado em teste para descobrir eventos de estacionamento ilegal. A ideia principal é de que estacionamento ilegal normalmente acontecem ao lado da calçada, que bloqueia o caminho dos usuários de bicicletas e afeta significativamente as suas trajetórias. Trabalhos futuros incluem diferenciar o estacionamento ilegal de outros eventos que podem mudar a trajetória das bicicletas. Assim como detectar estacionamentos ilegais que não sejam próximos a calçada.

3.4. “Deep Learning for IoT Big Data and Streaming Analytics: A Survey” ^[5]

Neste artigo os autores exploram uma técnica de ML chamada de Deep Learning (DL), que pode ser usada para facilitar a análise e o aprendizado no campo de IoT. O artigo começa articulando as características dos dados da IoT por uma perspectiva de ML, isto é, análise de big data de IoT e a análise da transmissão de dados de IoT. Também é analisado a tentativa de utilizar DP e IoT, assim como a questão de equipamentos inteligentes de IoT que incorporaram DL. Além disso, as abordagens de implementação de DL em centrais de fog e nuvem que servem como suporte para aplicações de IoT também são avaliados, visto que, por exemplo, carros autônomos precisam tomar decisões rápidas, e para isso é necessário realizar análises rapidamente (com baixíssimo atraso). Trabalhos futuros incluem: 1) Encontrar maneiras eficientes para utilizar grandes volumes de dados advindos de celulares; 2) Melhorar a segurança dos sistemas de IoT e diminuir a possibilidade de falhas; 3) Encontrar maneiras de facilitar a manutenção das comunicações e redes.

3.5. Comparação do que é apresentado nesse trabalho em relação aos trabalhos correlatos apresentados

Este presente artigo, em comparação com os trabalhos correlatos apresentados, consiste em um survey que contextualiza a área de Redes de Computadores com foco em IoT e IA. Os principais temas abordados incluem IoT, Fog, Big Data, IA e ML. Tal survey também contém a descrição de uma motivação sobre a importância do tema para o meio científico atual. Além disso, trabalhos correlatos foram descritos e analisados culminando na identificação dos seus problemas, suas soluções e seus trabalhos futuros, mostrando também a sua relação com o survey atual. Tendo em mente o foco do artigo e os trabalhos correlatos do mesmo, foi possível explorar aspectos relevantes, assim como os principais problemas existentes relacionados ao tema, suas possíveis soluções e o futuro da pesquisa. Foi destacado o artigo [10] que propõe uma arquitetura para monitorar a saúde de pacientes, utilizando-se das ferramentas descritas acima, armazenando os dados na cloud para então desenvolver o modelo de predição para doenças cardíacas. Contudo, em relação ao presente artigo, nada foi adicionado ao conhecimento científico, visto que se trata apenas de um survey, que sumariza as questões pertinentes sobre o tema, utilizando-se de artigos já realizados sobre o assunto.

4. Aspectos relevantes:

^[7] Dentre as propostas para coleta e gerenciamento de dados, podemos citar a tecnologia Fog. Fog é uma arquitetura emergente para computação, armazenamento, controle e conexões que distribui tais serviços para mais próximo ao usuário final. Fog cobre tanto tecnologia mobile e telefônica, através de hardware e software, residindo em borda das redes, mas também através de acesso a rede entre usuárias finais. Por funcionar como uma arquitetura, suporta a variedade crescente de aplicações, incluindo a IoT, 5G (Fifth-Generation) sistemas sem fio, e inteligência artificial embarcada.

Em décadas passadas, computação em movimento, controle, e armazenamento de dados na nuvem tem sido a tendência. Em particular, computação, armazenamento, e funções de gerência de redes são direcionadas para centros centralizados de dados, redes de IP centrais, e redes celulares centrais. Atualmente, entretanto, computação Cloud está encontrando desafios crescente em referente aos requerimentos emergente em IoT.

^[8] Com o desenvolvimento rápido de tecnologias de aquisição de localizações, massivas trajetórias são continuamente geradas. Muitas aplicações urbanas recaem fortemente na análise e minerações dos resultados de grandes quantidades de dados de trajetórias. Por conta disso, aplicações que envolvem cloud e fog surgem para gerenciar o acesso e a captura dessas informações. Artigos como o de [8] produzem sistemas que são capazes de eficientemente auxiliar em queries de trajetórias, que são baseados em plataformas cloud, como a Microsoft Azure.

^[9] Além disso, desenvolvimentos atuais em IoT e Cyber-Physical Systems (CPS) permitem a criação de soluções relacionadas também a área da saúde, capazes de realizar predições mais inteligentes tanto para a vida cotidiana, assim como em hospitais. Na grande maioria de sistemas de IoT baseados em sistemas de saúde, especialmente em casas inteligentes ou hospitais, uma ligação é necessária entre a infraestrutura de sensores e a

Internet. Para tais aplicações, considera-se a aplicação no limite da rede, isto é, na Edge, pois essas portas de entrada possuem conhecimento benéfico e controle construtivo tanto sobre a rede de sensores como os dados que são transmitidos através da Internet. Artigos como [9] exploram a Edge para oferecer serviços como armazenamento local, processamento de dado local, coleta de dados embarcados, entre outros. Além disso, Fog também pode ser explorado ao formar uma camada inteligente intermediária entre os nodos sensores e a Cloud. De como que o Edge delegaria ao Fog parte de suas tarefas, tornando o sistema capaz de gerenciar problemas de mobilidade, eficiência energética, escalabilidade e confiabilidade. Demonstrando que IoT unido com inteligência pode melhorar tais aspectos.

^[10] Ainda na área de saúde, relacionado a IoT, dispositivos médicos wearables continuamente geram enormes quantidades de dados, que é normalmente chamado de big data, de formas estruturadas e não-estruturadas. Por conta da complexidade do dado, o seu processamento e análise torna-se difícil. Para contornar essa dificuldade, artigos como ^[10] propõem arquiteturas para a implementação de IoT para armazenagem e processamento escalável para as grandes quantidades advindas de sensores relacionados a aplicações na saúde.

Um exemplo de utilização de sensores wearables seria para sugerir exercícios fisiológicos e hábitos alimentares a partir de um monitoramento de dois a três dias de monitoramento fisiológico contínuo dos pacientes. Nesse período sensores wearables iriam continuamente observar e armazenar dados da saúde do paciente em um local de armazenamento de dados. Tais dados ajudariam as diagnoses, em relação à saúde dos pacientes, geradas a partir de médicos. Portanto, não apenas testes laboratoriais seriam utilizados, mas também dados coletados a partir de sensores corporais.

^[6] O espaço de coleta de dados de trajetórias também é afetado. Os avanços nas técnicas de aquisição de locais e computação mobile tem gerado massivas quantidades de dados espaciais de trajetórias, que representa a mobilidade de uma diversidade de objetos móveis, tais como pessoas, veículos e animais. Muitas técnicas vêm sendo propostas para processamento, manipulação e mineração de dados nas décadas passadas. Aspectos precisam ser explorados, tais como a obtenção de dados de trajetórias, o seu pré-processamento e o gerenciamento de dados de trajetórias, assim como uma variedade de tarefas relacionadas a mineração.

Tais trajetórias nos oferecem informações sem precedentes para entender objetos móveis e localizações, favorecendo uma grande variedade de aplicações em redes sociais baseadas em localização, sistemas de transportes inteligentes, e computação urbana. Por sua vez, a prevalência de tais aplicações clama por pesquisas sistemáticas em tecnologias novas de computação para retirar conhecimento de dados de trajetórias.

5. Problemas existentes:

^[10] Em geral, muitos clusters de bancos de dados e recursos adicionais são necessários para armazenagem de big data. Porém, armazenamento e recuperação não são os únicos problemas. A aquisição de padrões significantes a partir de big data, como, por exemplo, informação de diagnóstico de pacientes, também é um problema significativo, que recai na área da IA. Sensores são normalmente utilizados em aplicações críticas para o futuro próximo.

Alguns sensores corporais foram desenvolvidos para o monitoramento contínuo da saúde, saúde física pessoal e consciência de atividade física. Técnicas tradicionais de armazenagem de dados e plataformas não são tão úteis onde o volume, velocidade e variedade de dados está crescendo. Portanto, um problema existente é o de fornecer uma arquitetura e implementação segura de uma arquitetura de IoT escalável para o processamento e proteção de sensores de tempo real usando tecnologias de big data.

[7] A emergência das tecnologias de IoT introduzem vários outros desafios que não podem ser adequadamente endereçados pelas tecnologias atuais de cloud. Tais desafios podem ser:

Requerimento rigoroso em relação a latência: Diversos sistemas de controles industriais, como os sistemas de manufatura, smart grids, sistemas de óleo e gás, normalmente demandam que a latência de ponta-a-ponta entre os sensores e o controle do modo fiquem em um intervalo de poucos milissegundos. Várias outras aplicações de IoT como comunicação entre veículos, comunicação entre veículos e as infraestruturas de trânsito, aplicações utilizando drones, aplicações de realidade virtual, aplicações de jogos, aplicações de transação financeira em tempo real, podem requerer latências abaixo de poucos décimos de milissegundos. Esses requerimentos estão muito distantes da capacidade dos serviços de cloud atuais.

Limitações de banda de rede: Com o avanço rápido e contínuo de um número crescente de coisas conectadas, há a criação de dados em uma taxa exponencial. Por exemplo, um carro conectado pode gerar dezenas de megabytes por segundo. Enquanto tal dado inclui 1) a mobilidade do carro, como sua rota e velocidade, 2) as condições de operação do carro, como o desgaste de seus componentes, 3) o ambiente em torno do carro, tais como a rua e as condições ambientais, e 4) os vídeos gerados pelas câmeras de segurança do carro. Um carro autônomo gerará ainda mais dados, que foi estimado em ser aproximadamente um gigabyte por segundo. É esperado que o sistema de smart grid dos EUA gerará 1000 petabytes de dados por ano. Por comparação, a US Library of Congress gerou cerca de 2.4 petabytes de dados por mês, Google trafegou aproximadamente um petabyte por mês.

Dispositivos com recursos limitados: Diversos dispositivos de IoT terão recursos limitados. Exemplos incluem sensores, coletores de dados, atuadores, controladores, câmeras de segurança, carros, trens, drones, e dispositivos médicos embutidos em pacientes. Vários desses dispositivos limitados não poderão contar com apenas seus recursos para cumprir com suas necessidades computacionais. Requerendo que todos eles interajam diretamente com a cloud seria irrealístico assim como proibitivo de acordo com o custo, visto que tais interações normalmente requerem intensivo custo de processamento e protocolos complexos. Por exemplo, a grande quantidade de microcomputadores em um veículo moderno precisa de updates de firmware, mas será impraticável requerer que cada um desses dispositivos com recursos limitados performem operações criptográficas e processos sofisticados obtenham updates de firmware e pela cloud.

Além dos problemas citados anteriormente, ainda teríamos a questão de manter as credenciais de segurança e software atualizadas em uma quantidade imensa de dispositivos; proteção de dispositivos com recursos limitados; acessar os status de segurança de um grande

número de sistemas distribuídos de maneira confiável; responder a problemas de segurança sem causar disfunções intoleráveis e, por último, serviços ininterruptos com conectividade intermitente na cloud.

Diferentemente aos campos de tecnologias mais maduros, tais como computação serial, comunicação digital, e a internet, onde sólido e forte fundação arquitetural foi fundamentada, nós ainda estamos perguntando por princípios arquiteturais para muitas aplicações de sistemas emergentes e aplicações tais como IoT, sistemas físicos-cibernéticos, e sistemas embutidos de IA. É necessário realizar decisões fundamentais desde onde há a computação onde armazenar dados ao longo do Cloud-to-Things contínuo para como realizar tarefas computacionais em um substrato capaz, heterogêneo e com nodos disponíveis. Fog proporciona uma direção para explorar tal arquitetura. A IoT emergente introduz vários novos desafios que não podem ser adequadamente resolvidos por modelos de Cloud e host atuais. Portanto é necessário realizar a discussão sobre tais aspectos.

6. Soluções possíveis:

^[10] A utilização apropriada da arquitetura fog poderia auxiliar nos problemas acima descritos. Fog é capaz de habilitar novos, e potencialmente disruptivos, modelos de negócio para computação e redes. Exemplos desses cenários incluem a convergência de roteadores e interruptores, servidores de aplicações e servidores de armazenamento para nodes de uma arquitetura fog. Tal transformação permitiria transformar significativamente a rede, o servidor, e a paradigma da indústria de software.

Além disso, fog-as-a-service (Faas) permitiria que novos modelos de negócios para os clientes. Diferentemente de cloud que são majoritariamente operados por grandes companhias que possuem condições para construir e operar grande centros de dados, Faas poderá permitir companhias, tanto grandes como pequenas, à entrega de computação privada e pública, armazenamento e serviços de controle em escalas diferente para ir ao encontro das necessidades de uma grande variedade de consumidores. Fog também permitiria uma nova maneira para os provedores de serviço de rede acrescentar valor aos consumidores em um novo mundo de neutralidade de rede.

Questões problemáticas em IoT poderiam ser minimizadas com a utilização de tecnologias fog, exemplos incluem: O problema de grande latência de IoT, poderia ser minimizado ao utilizar a análise de dados, o controle, e outros aspectos sensíveis ao tempo, mais próximo aos usuários finais. Isso seria ideal e normalmente a única opção que vai ao encontro ao problema de latência de sistemas de IoT. Problemas de limitações de banda de rede também poderiam ser minimizados com fog, visto que fog permitiria o processamento hierárquico através das camadas de cloud e “coisas”, permitindo que o processamento seja realizado onde ele pode ser balanceado entre requerimentos de aplicações e disponibilidade de redes e recursos computacionais. Isso também reduziria a quantidade de dados necessários para serem enviados para cloud. Fog também poderia realizar o esforço de algumas tarefas que precisam de grandes recursos para serem processadas, ao invés de dispositivos limitados fazerem isso, ou quando também não há a possibilidade de enviar tais tarefas para a cloud, portanto, reduzindo a complexidade desses dispositivos, seus custos de ciclos de vida, assim

como consumo de energia.

^[5] Já em relação ao problema de obtenção de informações relevantes a partir de dados advindos das tecnologias de IoT. Precisamos de melhores práticas relacionadas à IA e ML. Por conta disso, artigos propõem soluções para pesquisadores na área de IoT e desenvolvedores que querem construir análises, sistemas de IA utilizando métodos de deep learning (DL) de ML. O grande sucesso de cloud ao auxiliar DL está sendo sustentado pelo avanço e utilização de processos otimizados de redes neurais, assim como algoritmos distribuídos escaláveis de DL. Ao implementar tal subtipo de ML em plataformas fog para aplicações de IoT, tais como casas inteligentes e smart grids, chamariam a atenção dos usuários finais pelo motivo da facilidade de acessibilidade e tempo de resposta rápido. Mesmo assim, cloud baseado em análises de DL seriam de grande importância para análise de dados que excederiam as capacidades do fog a longo prazo. Algumas aplicações em cidades inteligentes, setores do governo, e dentro da própria nação de implantações de IoT, precisariam utilizar a infraestrutura de cloud unido a infraestrutura de DL.

^[7] Fog pode fornecer uma solução. Um único nodo Fog proporcionou uma plataforma em comum em cada gabinete para todos os serviços, e permitiu aplicações de diferentes fornecedores para coexistirem sem interferir um com o outro. Também proporcionou uma plataforma unificada para suportar comunicação, segurança, e gerenciamento duradouro para todas as aplicações, reduzindo os custos dos sistemas e permitiu fornecedores de aplicações focarem no desenvolvimento de aplicações ao invés de proporcionar hardware especializado e software para hospedar e gerenciar suas aplicações.

Enquanto Fog não precisa contar com a virtualização ou ser centrado em informação ou definido por software, podemos realizar um Fog centrado em informações e um Fog virtualizado já que esses ramos são complementares e podem ser disponibilizados para o Fog. O Fog também inclui redes móveis e com fio, além de borda transversal, acesso e wearables. O suporte à computação de borda móvel dentro de uma RAN exigirá muitas das mesmas funções de uma arquitetura de Néon de ponta a ponta para, por exemplo, distribuir, orquestrar, gerenciar e proteger os aplicativos e proteger os aplicativos e as plataformas de ativação de aplicativos. O Fog, no entanto, é mais amplo do que apenas suportar a computação de borda móvel. Fog é uma arquitetura para distribuir serviços de computação, armazenamento, controle e rede em qualquer lugar ao longo do contínuo do Cloud-to-Thing, dentro e fora de redes com e sem fio, e suportando aplicativos de redes móveis e com fio.

7. Projeto e Desenvolvimento de uma Proposta:

^[10] Este artigo em questão propõe uma arquitetura para monitorar a saúde de pacientes. A arquitetura proposta consiste nas fases de aquisição dos dados, fase de transferência de dados e fase de armazenamento de big data. Uma vez que os dados estão armazenados na cloud, o sistema de monitoramento utiliza o algoritmo de gradiente estocástico com regressão logística para desenvolver o modelo de predição para doenças cardíacas.

7.1. Arquitetura proposta:

A fim de promover comunicação eficiente entre dispositivos na internet, uma arquitetura é identificada por diferentes camadas, tais como camada de aplicação, camada de comunicação, camada de segurança, camada implantada, camada de hardware, camada de integração e camada de BD (Fig. 1). As tecnologias usadas na IoT incluem rótulos de RFID, sensores, atuadores e celulares. Tais objetos interagem entre si usando esquemas de endereço único para alcançar um objetivo em comum. Vários protocolos vêm sendo desenvolvidos em todas as camadas da pilha ISO para tornar possível as operações dos dispositivos da IoT. Constrained Application Protocol (CoAP) Routing Protocol for Low-Power and Lossy Networks (RPL) protocolos de mensagens são mais familiares. Tais protocolos são projetados a fim de preservar energia, assim como diminuir os requerimentos de computação e memória.

Como mostrado na Fig. 1, o foco da camada de segurança é prover segurança as coisas, comunicações e aplicações. Camada de segurança embarcada foca no desenvolvimento de funções seguras para coisas embarcadas tais como sensores, controladores, coisas wearables e não wearables. Camada de segurança de comunicação é usada para assegurar a transferência de dados pelas coisas. Camada de segurança de BD é focada na proteção dos dados e na segurança das conexões de BD para servidores remotos/cloud. Camada de segurança de aplicação é a de maior importância para as camadas de IoT. Camadas de aplicação são implantadas nela e protegidas com o auxílio de diferentes ferramentas de rede. O foco da camada de BD é nas conexões da base de dados de aplicações de IoT para a centros remotos de Fog/Cloud. Serviços de Java web, tais como JDBC, AWS EBS, e AWS EMR são usadas para prover conexões da base de dados entre IoT e aplicações.

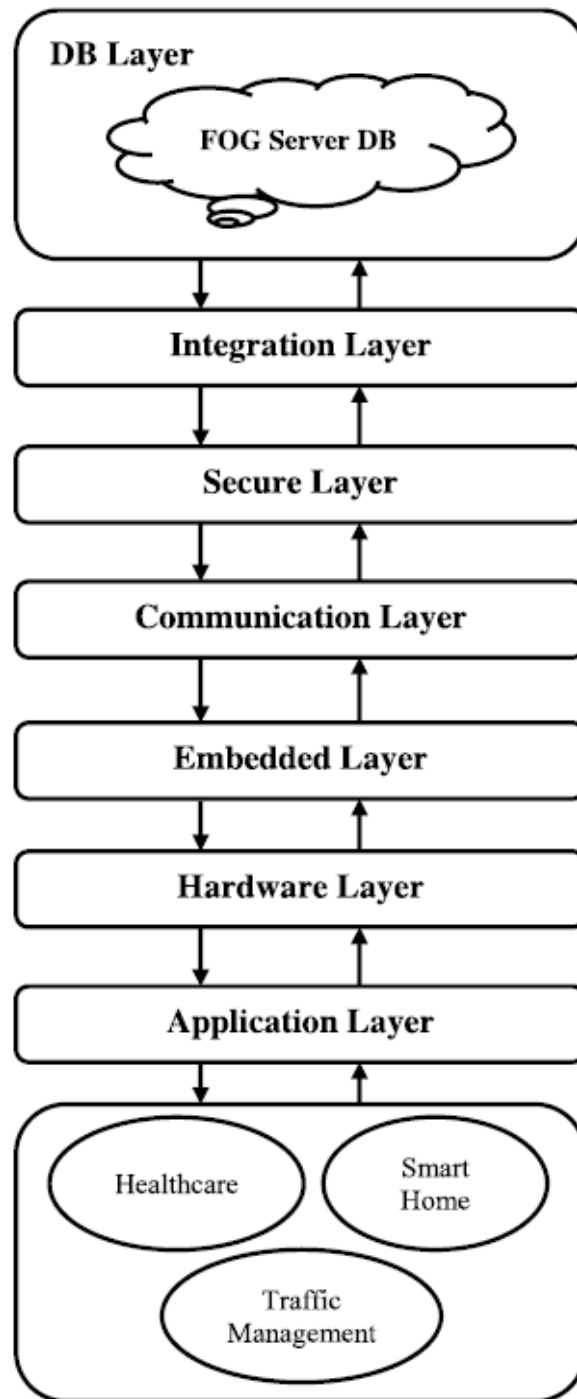


Fig. 1. The layered architecture of Internet of Things.

Nos tempos atuais, dispositivos de IoT estão continuamente gerando grandes quantidades de dados, que é ocasionalmente chamado de big data (tanto estruturado quanto desestruturado). Em geral, é difícil processar e analisar big data a fim de encontrar informações significativas. Além disso, a segurança sobre os dados é de extrema importância para sistemas de dados de saúde. Para superar tais problemas, uma nova arquitetura para a implementação de

IoT para armazenar e processar dados escaláveis de sensores (big data) em aplicações de saúde. A comunicação típica entre as coisas pode ser vista no diagrama de fluxo da Fig. 2.

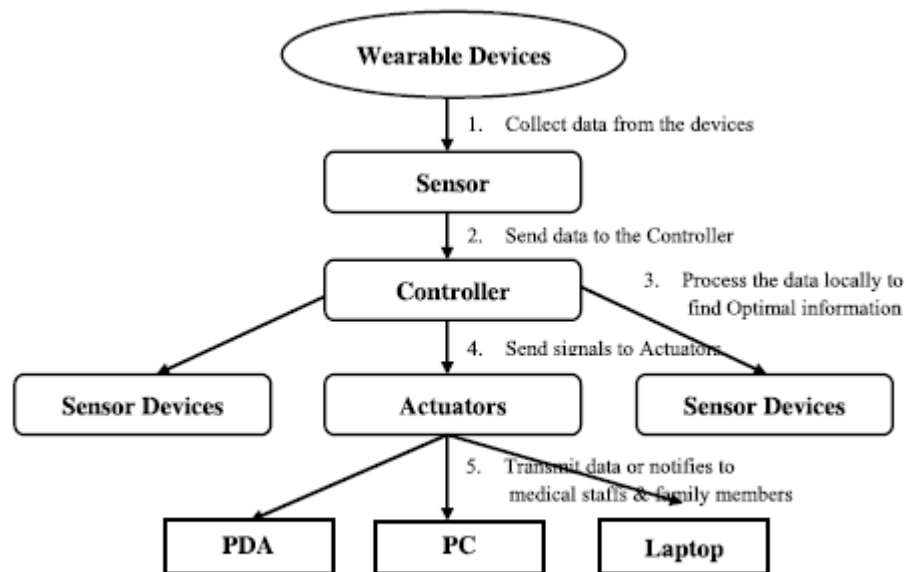


Fig. 2. Communication among things in healthcare Internet of Things.

A arquitetura proposta consiste em três fases diferentes, nomeadamente, fase de coleta de dados, fase de transferência de dados e fase de armazenamento de big data.

- A fase de coleta de dados é utilizada para o monitoramento constante da saúde do paciente. Portanto, no momento em que a taxa de respiração, taxa de batimento cardíaco, pressão sanguínea, temperatura corporal e açúcar no sangue excedam os valores normais, os dispositivos mandam uma mensagem de alerta com valor clínico para o médico que está usando a rede sem fio através da computação fog. Na fase de coleta de dados, dados de sensores coletados da computação fog são conectados com a cloud para transferir e armazenar nos bancos de dados.
- Na fase de transferência de dados, a arquitetura proposta usa o método ‘s3cmd utility’ para mover os dados clínicos para o container Amazon S3. Uma organização de saúde utiliza a instância EC2 na nuvem da Amazon para armazenar os dados de sensores (dados fisiológicos) dos pacientes. Entretanto, esse conjunto de dados de sensores não está disponível na Amazon EMR por conta dos dados armazenados apenas no disco local rodando em uma instância EC2. Portanto, os dados precisam ser transferidos para o EMR. Amazon EMR fornece serviço Hadoop na nuvem para processar big data.
- Em geral, armazenamento de grandes quantidades de dados no contêiner da Amazon S3 não é possível. Para contornar esse obstáculo, a arquitetura proposta utiliza Apache Pig na fase de armazenamento de dados para transferir grandes quantidades de dados (dados clínicos) do contêiner da Amazon S3 para o Apache HBase.

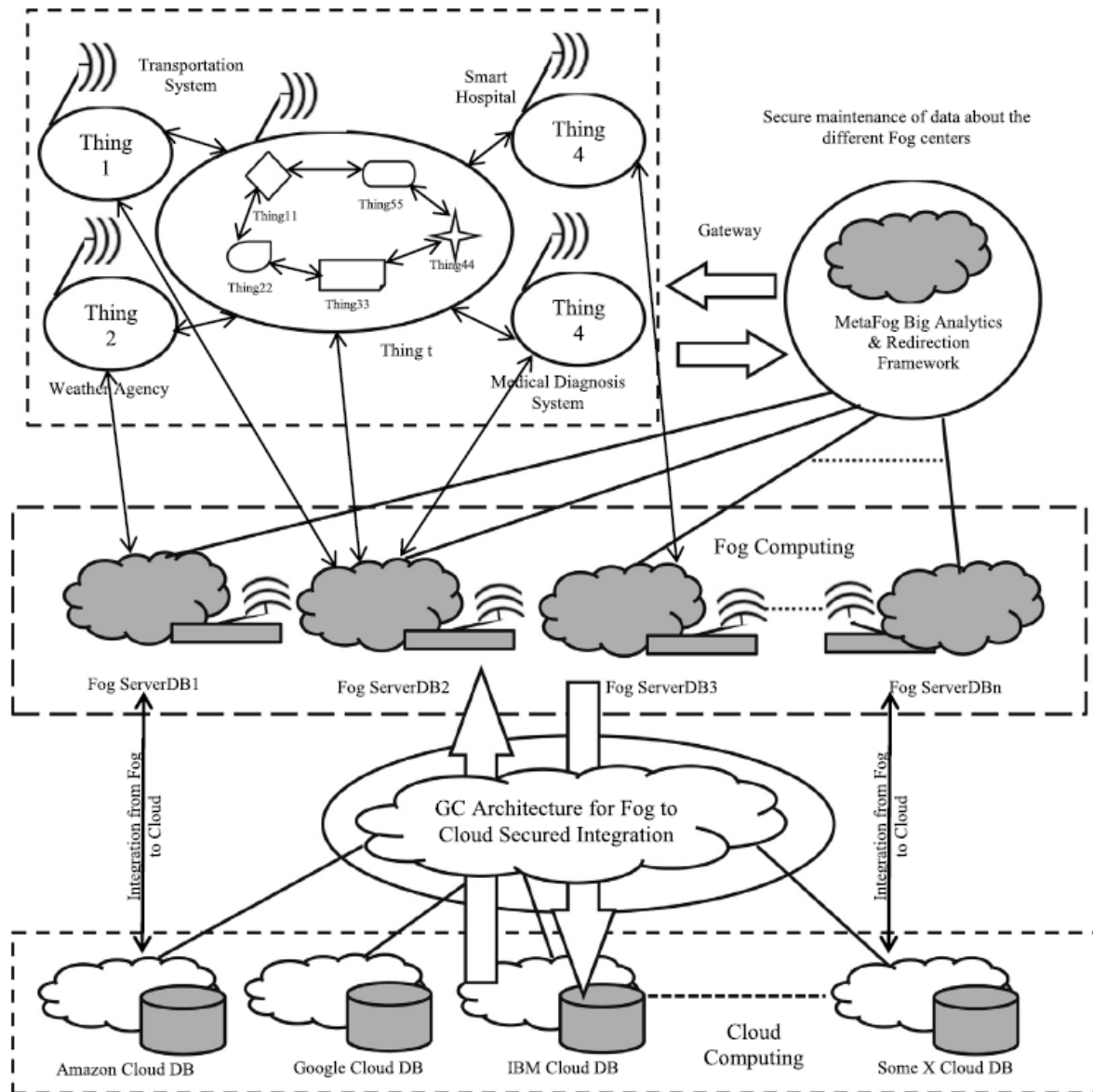


Fig. 3. MF-R with GC architecture for Internet of Things and big data ecosystem.

Como mostrado na Fig. 4, dados de sensores coletados da computação fog são conectados à cloud para transferência e armazenamento no banco de dados. A arquitetura proposta move os dados para a Amazon S3, onde a Amazon EMR pode acessar tais dados. Isso é para habilitar a utilização de conjuntos de dados de sensores médicos na nuvem.

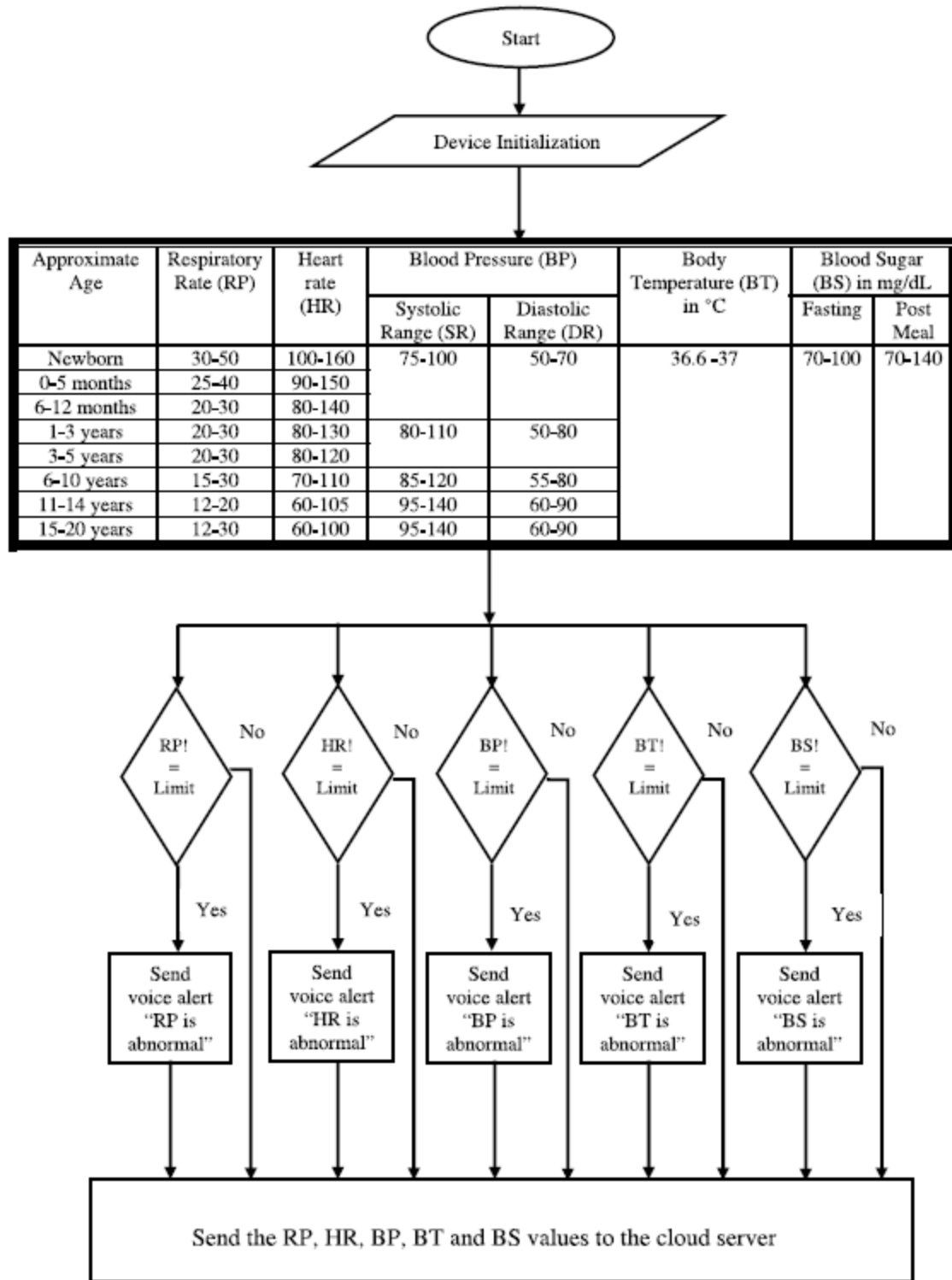


Fig. 4. Data flow diagram for body sensor devices in continuous IoT health monitoring system.

7.2. Validação e Resultados:

O sistema de monitoramento baseado em IoT proposto utiliza o algoritmo gradiente estocástico com regressão logística para desenvolver modelos de predições de doenças cardíacas. Registros clínicos anteriores e dados de sensores dos pacientes são coletados através da Cleveland Heart Disease Database (CHDD). Neste processo, o sistema de monitoramento de saúde utiliza dados de sensores atuais obtidos a partir de sensores corporais através da utilização da cloud e tecnológicas de big data. Em geral, a base de dados CHDD contém 76 atributos; mas todos os experimentos publicados referenciam o uso de um subconjunto de 14 desses 76 atributos. A tabela de classificação de previsão pode ser vista na Table 6.

Table 6
Training and testing data.

Training sample				Validation sample			
	Yes (1)	No (0)	Total		Yes (1)	No (0)	Total
Yes (1)	89	33	122	Yes (1)	36	11	47
No (0)	25	64	89	No (0)	14	31	45

A análise de performance do modelo proposto, baseado em MapReduce, é mostrado na Table 7. O modelo proposto de predição de doenças cardíacas classifica eficientemente doenças cardíacas com a acurácia de treinamento e amostra de validação sendo de 72.51% e 72.82% respectivamente. A performance do modelo de predição é avaliada com ajuda da sensibilidade, especificidade, precisão, recordação e F-Measure que podem ser definidos como mostrado abaixo.

$$\text{Sensitivity} = \frac{\text{True Positive}(TP)}{\text{True Positive}(TP) + \text{False Negative}(FN)} \quad (1)$$

$$\text{Specificity} = \frac{\text{True Negative}(TN)}{\text{False Positive}(FP) + \text{True Negative}(TN)} \quad (2)$$

$$\text{Accuracy} = \frac{(TP) + (TN)}{(TP) + (FN) + (FP) + (TN)} \quad (3)$$

$$\text{Precision} = \frac{\text{True Positive}(TP)}{\text{True Positive}(TP) + \text{False Positive}(FP)} \quad (4)$$

$$\text{Recall} = \frac{\text{True Positive}(TP)}{\text{True Positive}(TP) + \text{False Negative}(FN)} \quad (5)$$

Table 7
Performance evaluation.

Training sample					
Accuracy	Sensitivity	Specificity	Precision	Recall	F-measure
72.51%	78.07%	65.98%	72.95%	78.07%	75.42%
Validation sample					
Accuracy	Sensitivity	Specificity	Precision	Recall	F-measure
72.82%	72%	73.8%	76.59%	72%	74.22%

8. Conclusões e Trabalhos Futuros:

A temática da união entre os dados coletados pela rede de IoT, e a posterior análise por métodos de IA é bastante atual e de suma importância para o futuro. Nota-se a necessidade da formulação de arquiteturas viáveis para a conexão correta entre os dispositivos dos usuários finais com a plataforma cloud, que pode ser auxiliada por estruturas de fog e edge, fazendo com que o processamento que será realizado por algoritmos de ML e técnicas de IA seja feito eficientemente e tenha resultados úteis para a aplicação em questão.

^[5] Deep learning e IoT continuamente chamam a atenção de pesquisadores e empresas nos anos recentes, visto que essas duas tendências tecnológicas já foram provadas em terem efeitos positivos nas nossas vidas, cidade e o mundo como um todo. IoT e DL constituem uma corrente de produção de dados e aplicação ao consumidor, no qual IoT gera o dado bruto que é então analisado por modelos de DL, e modelos de DL produzem abstração de alto-nível e introspecção que é fornecido aos sistemas de IoT para aprimorar seus serviços. DL baseado em infraestruturas fog e cloud para auxiliar aplicações de IoT também possuem seu espaço de pesquisa.

^[9] Outros aspectos que foram citados incluem o conceito de fog sendo utilizados em sistemas inteligentes relacionados à área da saúde unidos a IoT. Entradas inteligentes próximos aos sensores de casas inteligentes ou hospitais podem explorar a estratégia única para resolver diversos desafios relacionados a IoT, como mobilidade, eficiência energética, escalabilidade e problemas de confiabilidade. Foi investigado os serviços de auto nível que pode ser oferecido por entradas inteligentes aos sensores e usuários finais de maneira geograficamente distribuídas, na edge da rede (por exemplo, processamento local, armazenamento, notificação, padronização, firewall, serviços web, compressão, etc.). Foi formado uma camada intermediária de computação fog para formar uma camada de processamento intermediária. O sistema inclui todo o fluxo de dados desde a aquisição de dados nos sensores nodos até o cloud e usuários finais.

^[3] Novos paradigmas de comutação coexistem com outros paradigmas mais clássicos, e cada um deles trouxeram para auxiliar conjuntos de ferramentas e tecnologias. Esse grande número de diferentes designs de computação nos permitem implementar sistemas de complexidade que nunca foi imaginado, e que aumenta dia a dia. Alguns dos paradigmas mais recentes ainda não foram suficientemente aplicados em prática e até criam confusão sobre o

que eles significam para diferentes comunidades científicas. Nesse artigo em questão, foi explorado os paradigmas da computação que aparecem na comunidade científica sobre sistemas distribuídos, focando especialmente em computação cloud, computação edge e computação fog. Também foi descrito um conjunto de tecnologias computacionais ou serviços, que quando incrementada com os paradigmas de computação, podem habilitar aplicações computacionais sociais de interesse. Alguns exemplos foram descritos, partindo do domínio de transpores, e englobando o campo da energia. Essas aplicações podem ser realizadas com sucesso tanto em computação de dispositivos edge assim como em dispositivos de computação fog. Porém, ao imaginarmos aplicações mais complexas e integradas, nós precisamos considerar os desafios mencionados anteriormente. Tecnologias computacionais atuais resolvem apenas parcialmente tais desafios, dando à comunidade uma grande oportunidade de explorar esse grande universo de pesquisa.

[7] Borda ativada por fog e rede de acesso: o Fog pode ser usado para suportar a rede na borda. Por exemplo, Fog pode fornecer serviços para ajudar dispositivos de ponta de rede e dispositivos de usuário final (por exemplo, veículos, drones, robôs industriais e de consumo, smartphones e realidade virtual) a formar redes locais, fornecendo credenciais de segurança temporárias a esses dispositivos locais para ajudar eles a estabelecer comunicações confiáveis e atuarem como servidores de aplicativos locais e servidores de armazenamento de dados para as redes de borda. Algumas funções do Fog para suportar tal rede de ponta podem ser implementadas nos dispositivos do usuário final. Em tais casos, como as funções do Fog interagem com os sistemas operacionais e o hardware dos dispositivos do usuário final torna-se essencial. Mais do que apenas usar o D4D para agrupar recursos de borda inativa, podem ser necessárias novas pilhas de protocolo para que os dispositivos do usuário final suportem a rede de borda ativada pelo Fog.

Comparado com o Cloud, Fog apresenta novos desafios de segurança. Sistemas distribuídos, como o Fog distribuído, são geralmente mais vulneráveis a ataques do que os sistemas centralizados, como Cloud. Enquanto a Cloud opera em instalações altamente protegidas, selecionadas e controladas por operadores de Cloud, Fog frequentemente precisa operar em ambientes mais vulneráveis - onde eles podem atender melhor aos requisitos do cliente e sempre que os usuários desejarem. Muitos sistemas de Fog serão significativamente menores que nuvens (por exemplo, um Fog em um veículo) e, portanto, podem não ter tantos recursos quanto as Clouds para se protegerem. Além disso, cada sistema Fog pode não ter a inteligência global necessária para detectar ameaças.

[10] Neste artigo, a arquitetura MetaFog-Redirection redireciona os dados de saúde para diferentes servidores de Fog com base na categorização de dados (sensível, crítica e normal). As medidas realizadas durante o artigo não foram checadas com outras arquiteturas similares. Tal arquitetura executa o processamento de segurança e de arquivos de log para impedir os dados do usuário mal-intencionado e não autorizado. Também discute uma arquitetura segura e a implementação de uma arquitetura de IoT escalonável para processar e proteger dados de sensores em tempo real usando tecnologias de Big Data que sejam escaláveis. As medidas realizadas durante o artigo não foram checadas com outras arquiteturas similares.

9. Referências bibliográficas:

Quantidade de referências por ano: 6 de 2018, 5 de 2017, 1 de 2016, 1 de 2015, 1 de 2010 e 1 de 2003.

- [1] HE, T.; BAO, J.; RUAN, S.; LI, Y.; TIAN, C.; ZHENG, Y.; Detecting Vehicle Illegal Parking Event using Sharing Bikes' Trajectories. **The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining**, London, ago. 2018.
- [2] COSAR, S.; DANATIELLO, G.; BOGORNY, V.; GÁRATE, C.; ALVARES, L. O.; BRÉMOND, F.; Toward Abnormal Trajectory and Event Detection in Video Surveillance. **IEEE Transactions on Circuits and Systems for Video Technology**, p. 683-695, 2017.
- [3] GARCÍA-VALLS, M.; DUBEY, A.; BOTTI, V.; Introducing the new paradigm of Social Dispersed Computing: Applications, Technologies and Challenges. **Journal of Systems Architecture**, maio. 2018.
- [4] BIBRI, S. E.; A foundational framework for smart sustainable city development: Theoretical, disciplinary, and discursive dimensions and their synergies. **Sustainable Cities and Society**, v. 38, p. 758-794, abr. 2018.
- [5] MOHAMMADI, M.; AL-FUQAHA, A.; SOROUR, S.; GUIZANI, M.; Deep Learning for IoT Big Data and Streaming Analytics: A Survey. **IEEE Communications Surveys & Tutorials**, jun. 2018.
- [6] ZHENG, Y.; Trajectory Data Mining: An Overview. **ACM Trans. Intell. Syst. Technol.** **6**, **3**, Article 29, maio. 2015.
- [7] CHIANG, M.; ZHANG, T.; Fog and IoT: An Overview of Research Opportunities. **IEEE Internet of Things Journal**, v. 3, p. 854-864, abr. 2016.
- [8] LI, R.; RUAN, S.; BAO, J.; ZHENG, Y.; A Cloud-Based Trajectory Data Management System. **The 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems**, p 96, nov. 2017.
- [9] RAHMANI, A. M.; GIA, T. N.; NEGASH, B.; ANZANPOUR, A.; AZIMI, I.; JIANG, M.; LILJEBERG, P.; Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: A fog computing approach. **Future Generation Computer Systems**, **78**, p.641-658, fev. 2017.
- [10] MANOGARAN, G.; VARATHARAJAN, R.; LOPEZ, D.; KUMAR, P. M.; SUNDARASEKAR, R.; THOTA, C.; A new architecture of Internet of Things and big data ecosystem for secured smart healthcare monitoring and alerting system. **Future Generation Computer Systems** **82**, p.375-387, out. 2017.