

BIG DATA, TÉCNICAS DE CLASSIFICAÇÃO DADOS E SUAS APLICAÇÕES NA ENGENHARIA CIVIL

Bruno Gallo Belluzzo (IC) e Massaki de Oliveira Igarashi (Orientador)

Apoio: PIBIC Mackenzie

RESUMO

A revolução tecnológica, a era da conectividade e a comunicação ubíqua trouxeram novos desafios de atuações para os profissionais de engenharia e tecnologia de informação. O advento da Internet das coisas e as novas gerações de tecnologias e arquiteturas viabilizaram a utilização de diferentes tipos de sensores (sobretudo os sem fio), a nanotecnologia e a necessidade de utilização de técnicas e ferramentas computacionais para a extração de informações a partir de volume de dados muito grande e diverso (Big Data). Isto pode trazer grandes oportunidades; principalmente em relação a: transportes, planejamento urbano e construção civil. Através da análise desses dados e da detecção de padrões, ações estratégicas podem ser tomadas a fim de aprimorar vendas e a qualidade dos produtos de empresas, problemas de gerenciamento urbano e de tráfego são contornados de forma imediata e eficiente, construções e edificações têm um monitoramento em tempo real de suas condições estruturais, com o intuito de prevenir possíveis acidentes. Neste contexto este artigo, que compreende uma revisão bibliográfica em relação a grandes conjuntos de dados (Big Data) e suas aplicações na engenharia civil, pode não só evidenciar tendências e potencialidades ainda pouco exploradas, mas também possibilidades de melhorias para o planejamento urbano, fiscalização ambiental e monitoramento de desastres.

Palavras-chave: Big data. Data mining. Engenharia civil.

ABSTRACT

The technological revolution, the era of connectivity and ubiquitous communication have brought new challenges for engineering and information technology professionals. The advent of the Internet of Things and the new generations of technologies and architectures enabled the use of different types of sensors (especially wireless), nanotechnology and the need to use computational techniques and tools to extract information from data volume very large and diverse (Big Data). This can bring great opportunities; mainly in relation to: transportation, urban planning and construction. By analyzing these data and detecting patterns, strategic actions can be taken to improve sales and product quality of companies, urban management problems and traffic are immediately and efficiently circumvented, buildings and edifications have a realtime monitoring of their structural conditions, in order to prevent accidents.

In this context this article, which comprises a bibliographical review in large data sets (Big Data) and its applications in civil engineering, can not only highlight trends and potentialities that has not yet been explored, but also possible improvements for urban planning, environmental surveillance and disaster monitoring.

Keywords: Big data. Data mining. Civil engineering.

1. INTRODUÇÃO

Desde 1965, quando Gordon E. Moore “profetizou” que o número de transistores praticamente dobraria a cada dois anos (MOORE, 1965), a capacidade de armazenamento também cresceu; e espantosamente, a quantidade de dados gerados cresceu ainda mais. O desafio passou a ser a seleção das informações úteis, interessantes e necessárias em meio a milhares de dados complexos a serem descartados. Logo ocorreu o advento das técnicas, métodos e ferramentas computacionais para auxiliar a extração dessas informações, dentre as quais merecem destaque as técnicas de classificação de dados (Data Mining) para a busca por padrões consistentes, como regras de associação ou sequências temporais em grandes bancos de dados (LAROSE, 2015).

A classificação (mineração) de dados é uma das etapas responsáveis pela descoberta de conhecimento. Os desafios neste campo incluem: captura, arranjo dos dados, pesquisa, compartilhamento, armazenamento, transferência, visualização e análise das informações. Para tal, são utilizados algoritmos, ferramentas e cálculos matemáticos, recuperação de informação e inteligência artificial para entender padrões que tornam possíveis as formulações de hipóteses e as regras. Ela propicia vantagens competitivas às empresas interessadas; através de informações mais relevantes e precisas, que ajudarão nas decisões sobre estratégia de preços, promoções, riscos de investimento, e concorrência; o que afeta também os modelos de negócio, as receitas, as operações e as mídias sociais utilizadas no relacionamento com os clientes (O'BRIEN; 2004; O'BRIEN; MARAKAS, 2007; GALVÃO; MARIN, 2009; WU et al., 2014; JALALI-HERAVI; ARRASTIA; GOMEZ, 2015; LAROSE, 2015).

O termo Big Data advém do crescimento do volume de dados gerados, da dificuldade do seu armazenamento, processamento e análise através das tecnologias tradicionais. Para isso, surgiu a necessidade de novas maneiras de identificação, separação e tradução dessa grande quantidade de informações dispostas de forma aparentemente caótica.

Surge, portanto, o termo Big Data, caracterizado por seus 5 V's: **v**olume, relacionado a quantidade de todos os tipos de dados gerados vindos de diferentes fontes; **v**ariedade, refere-se aos diferentes tipos de dados coletados (por meio de sensores, smartphones, ou redes sociais; eles podem estar na forma de áudio, vídeo, imagem ou texto); **v**elocidade, está ligada a velocidade de transferência dos dados; **v**alor, considerado o mais importante aspecto do Big Data por ser o processo de descoberta de importantes valores ocultos em meio a grandes conjuntos de dados e **v**eracidade, ou seja, a garantia da confidencialidade, integridade e disponibilidade dos dados assegurada pelas novas técnicas de criptografias de dados (HASHIM, 2015).

Apesar da importância das tecnologias relacionadas ao Big Data para processar dados em grande escala, os atuais sistemas de informação da engenharia civil ainda são escassos, assim como as aplicações de novos métodos e ferramentas.

O objetivo deste artigo é realizar uma revisão bibliográfica em relação a grandes conjuntos de dados (Big Data), a técnicas de mineração de dados (Data Mining) e suas aplicações na engenharia civil, a fim de apresentar desafios de atuação desses profissionais e evidenciar tendências e potencialidades ainda pouco exploradas.

2. DESENVOLVIMENTO DO ARGUMENTO

Para esse estudo foi elaborado um ensaio teórico a partir do levantamento bibliográfico de produções nacionais e internacionais referentes ao tema de pesquisa. Na primeira etapa de levantamento de dados foram selecionados artigos na base de dados Scopus, Scientific Electronic Library Online (SciELO), e periódicos CAPES. Foram utilizados como critério de busca as palavras big data, data mining, civil engineering. Não foi feita restrição quanto ao índice de impacto das publicações, pois o objetivo foi alcançar uma visão mais ampla da produção científica relacionada ao tema, sem o filtro dos critérios editoriais. Selecionou-se artigos que apresentavam relação direta com os objetivos desse estudo.

O processo de análise de dados, consiste na coleta informações, o seu processamento e na saída filtrada desses dados. Para a realização desse processo os dados precisam passar por uma sequência de passos definidos, entre eles estão a coleta, armazenamento, análise e apresentação. Algumas plataformas e linguagens que se destacam nesse meio são o MapReduce, Hadoop e o NoSQL.

Map - Reduce é um modelo de programação que foi criado pelo Google visando facilitar o processamento dos grandes volumes de dados gerados. A partir dele é permitida uma manipulação desses grandes volumes de dados de forma distribuída, paralela e com grande tolerância a falhas. O modelo de programação baseia-se em duas operações básicas: map e reduce. A operação map é usada para localizar algo, de forma que recebe um par no formato chave/valor e gera um conjunto intermediário de dados. A operação reduce é utilizada para fazer a sumarização do resultado, sendo executada para cada chave intermediária, com todos os conjuntos de valores intermediários àquela chave intermediária.

Hadoop é *framework* (ou seja, uma biblioteca) escrita em linguagem Java, para processamento distribuído, sendo uma de suas implementações as funções Map-Reduce, voltada para clusters acionados por modelos de simplificação de programação e processamento de grandes massas de dados. O Hadoop funciona através do processamento paralelo de conjuntos de dados e foi projetado para escalar a partir de um único servidor a

milhares de máquinas. O resultado disso serão várias pequenas parcelas de um todo distribuídas para diferentes máquinas através da função Map, e então recombinaadas (através da função reduce).

Os NoSQL, abreviação do termo em inglês *Not Only SQL*, ou seja, os bancos de dados não relacionais, surgiram da necessidade de uma performance superior e de uma alta escalabilidade. Eles têm uma grande facilidade na distribuição horizontal, ou seja, mais dados, mais servidores, não necessariamente de alta performance. Um grande utilizador desse conceito é o Google, que usa computadores de pequeno e médio porte para a distribuição dos dados; essa forma de utilização é muito mais eficiente e econômica. Além disso, os bancos de dados NoSQL são muito tolerantes a erros e propiciam, inclusive, uma redundância dos dados.

Os bancos de dados NoSQL executam atividades para os quais os bancos de dados relacionais tradicionais, conhecidos como SQL, seriam muito restritos, pois necessitariam a distribuição vertical de servidores, ou seja, quanto mais dados, mais memória e mais disco de servidor seriam necessários.

O modelo tradicional (banco de dados relacional) é popularmente conhecido pela sigla SQL, abreviação para *Structured Query Language* (Linguagem de Consulta Estruturada), uma linguagem padrão de gerenciamento de dados que interage com os principais bancos de dados baseados no modelo relacional. Ela é uma linguagem essencialmente declarativa; isso significa que o programador necessita apenas indicar qual o objetivo pretendido, através de comandos, para buscar informações no banco de dados, adicionar registros, criar e alterar tabelas. Alguns dos principais comandos SQL para manipulação de dados são: INSERT (inserção), SELECT (consulta), UPDATE (atualização), DELETE (exclusão). SQL possibilita ainda a criação de relações entre tabelas e o controle do acesso aos dados.

2.1 Aplicações na Engenharia Civil

As técnicas analíticas avançadas suportadas pelas tecnologias Big Data têm sido vantajosas para minimizar o impacto ambiental dos resíduos gerados pela indústria de construção civil. No passado, os resíduos eram gerenciados somente após sua produção; mas com a utilização de softwares para gerenciamento de resíduos, as decisões relativas ao seu manuseio ocorrem desde a fase de concepção do projeto, o que possibilita uma minimização considerável (BILAL et al., 2016).

Os engenheiros e as instituições responsáveis pela segurança, manutenção e operação das estruturas de uma construção civil enfrentam o desafio de desenvolver uma metodologia para identificar automaticamente os danos. Uma proposta apresentada para a

solução desta dificuldade, consiste em analisar mudanças bruscas e permanentes detectadas através da monitorização, a longo prazo, de sensores instalados nas edificações, desenvolvendo métodos para processar e analisar uma grande quantidade de dados (POSENATO et al., 2010).

O planejamento urbano e o futuro das cidades inteligentes foi viabilizado, principalmente, devido à revolução causada pela conectividade entre os aparelhos eletrônicos, os dispositivos móveis, os smartphones, as centrais multimídias e as máquinas inteligentes; onde há um grande número de dispositivos se comunicando entre si e através de redes corporativas ou internet e gerando um massivo volume de dados. Um grande número de objetos está sendo conectado à internet, o que compreende a internet das coisas (tradução do termo em inglês *Internet of Things – IoT*). A internet das coisas é uma revolução tecnológica a fim de conectar aparelhos eletrônicos do dia-a-dia, máquinas industriais e meios de transporte à internet (RATHORE et al., 2016).

Devido ao rápido crescimento da densidade populacional nas cidades urbanas, infraestrutura e serviços são requeridos para fornecer as necessidades dos moradores da cidade. Houve um aumento significativo dos dispositivos digitais, por exemplo, sensores, atuadores e smartphones que conduzem a grandes potenciais de negócios para a IoT, uma vez que todos os dispositivos podem interconectar e comunicar uns com os outros na Internet. São objetos inteligentes e autoconfiguráveis que estão conectados uns aos outros através de uma infraestrutura de rede global. Estes objetos são reais, amplamente dispersos, com baixa capacidade de armazenamento e processamento; mas com o objetivo de melhorar a confiabilidade, o desempenho e a segurança das cidades inteligentes e suas infraestruturas. Baseado neste conhecimento, cabe também uma visão geral do termo cidades inteligentes

As cidades tornaram-se dotadas de inteligência graças aos desenvolvimentos recentes das tecnologias digitais. Uma cidade inteligente é equipada com diferentes elementos eletrônicos empregados por várias aplicações, como câmeras de rua para sistemas de observação, sensores para sistemas de transporte, etc. Além disso, há a possibilidade do uso de dispositivos móveis individuais. Portanto, considerando o ambiente heterogêneo, diferentes termos, tais como características de objetos, contribuintes, motivações e regras de segurança devem ser investigados.

No contexto IoT, os dispositivos podem ser integrados com base na localização geográfica e avaliados através de um sistema de análise. Os serviços de sensores para a

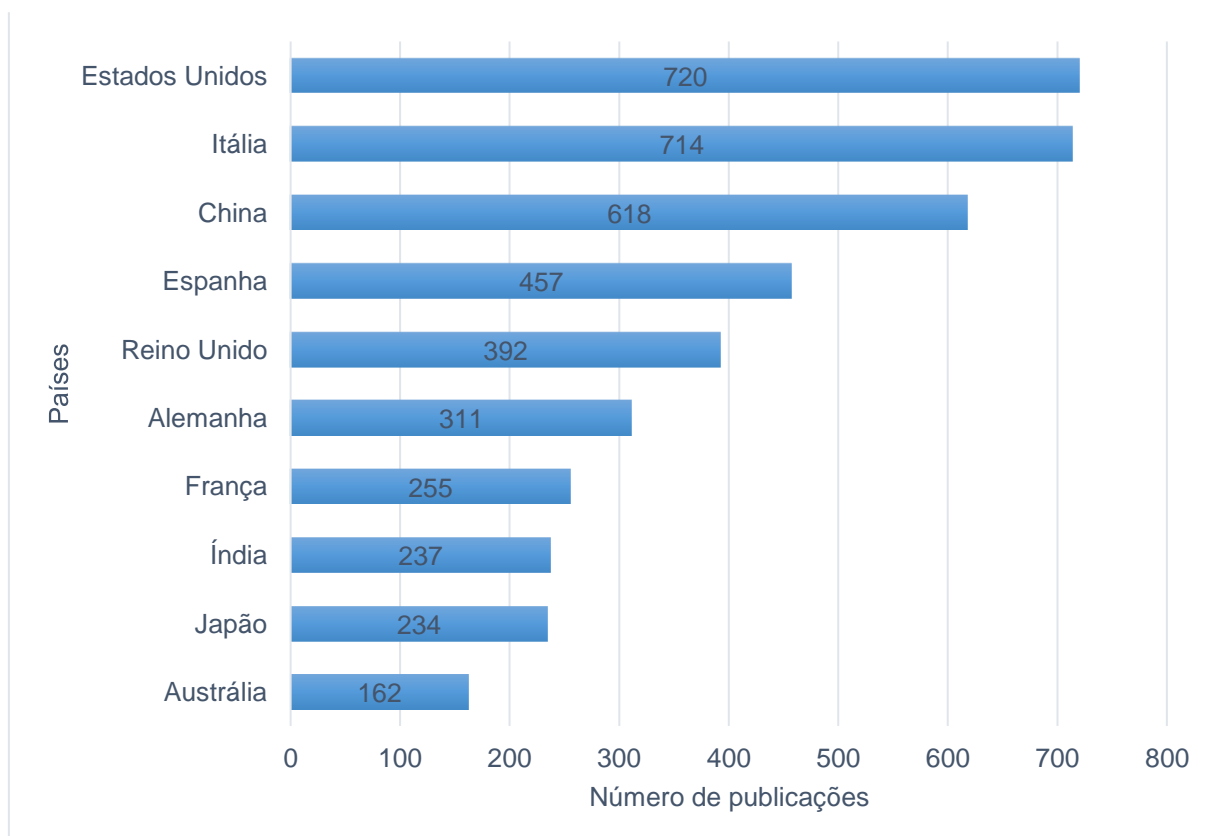
coleta de dados específicos podem ser utilizados com vários projetos em curso relacionados com a monitorização de ciclistas, veículos, estacionamentos públicos, etc. Existem muitas aplicações de domínio de serviço que utilizam uma infraestrutura IoT para

facilitar as operações de poluição atmosférica e sonora, a mobilidade dos veículos e dos sistemas de vigilância.

A revolução da Internet fornece uma infraestrutura em que muitas pessoas são capazes de interconectar uns aos outros. A próxima revolução da Internet tornará possível fornecer interconexões adequadas entre os objetos. Em 2011, o número de objetos que estão interconectados juntos era muito mais do que o número de pessoas. Por conseguinte, a IoT irá afetar os vários aspectos da vida dos cidadãos das cidades inteligentes como: saúde, segurança e transporte; além de desempenhar um papel importante em relação às decisões políticas (como economia de energia, redução da poluição, etc.), monitorização remota e infraestruturas necessárias. Neste contexto, a IoT contribuirá para proporcionar um funcionamento mais eficiente, econômico e seguro do sistema baseado em diferentes aspectos, tais como políticas de economia de energia, considerações econômicas, níveis de fiabilidade.

A Itália tem uma grande expressividade nas pesquisas sobre a internet das coisas; superada apenas, em número de publicações, pelas grandes potências no Big Data: Estados Unidos e China. Porém, quando o assunto se refere às cidades inteligentes, a Itália, nos últimos 5 anos, teve um crescimento ainda mais expressivo, atingindo um total de 714. Esse número é muito próximo ao dos Estados Unidos (720). (Figura 1)

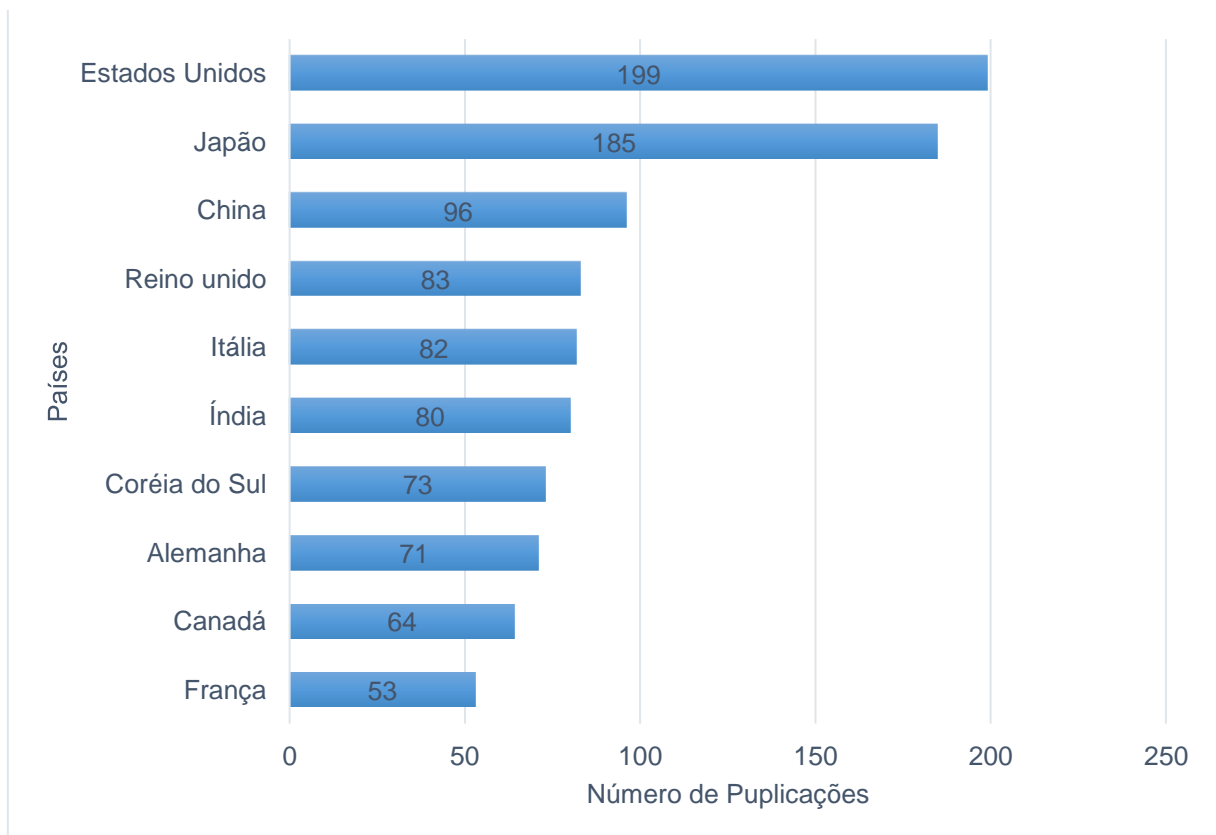
Figura 1: Número de publicações a respeito de cidades inteligentes por país.



A internet das coisas direciona o leitor ao conceito de casas inteligentes, onde aparelhos eletrônicos diferentes estão interligados uns aos outros, alcançando serviços multimídia interativos de alta qualidade. Nesse sistema, há um grande número de dispositivos comunicando-se entre si, gerando um massivo volume de dados, também denominado Big Data. Casas inteligentes com sensores de fumaça e temperatura, podem, por exemplo, evitar incêndios, reduzir o consumo de água, de eletricidade e de gás e possibilitar o gerenciamento mais eficiente, reduzindo os impactos ao meio ambiente.

O Japão se destacou quanto ao número de publicações acerca desse tema, ficando apenas atrás dos EUA. Essa pesquisa e estudo intensivos direcionados a casas inteligentes podem estar relacionados a algumas necessidades relevantes presentes no cotidiano da população, como seu envelhecimento, grandes cargas horárias de trabalho, pequeno tamanho das moradias devido à falta de espaço urbano, renda per capita e densidade demográfica.

(Figura 2).

Figura 2: Número de publicações a respeito de casas inteligentes por país.

Além das casas inteligentes, a internet das coisas compreende os estacionamentos inteligentes, sistemas de segurança inteligentes, tráfego veicular inteligente, os sistemas de água e de temperatura inteligentes. As informações climáticas, de temperatura e água podem aumentar a eficiência da previsão e fornecimento de informações referentes à umidade do ar, chuva, níveis de rios, lagos e barragens, pressão atmosférica, velocidade do vento, temperaturas máximas e mínimas. Esses fatores permitem, por exemplo, evitar enchentes e gerenciar melhor o consumo de água pelos cidadãos.

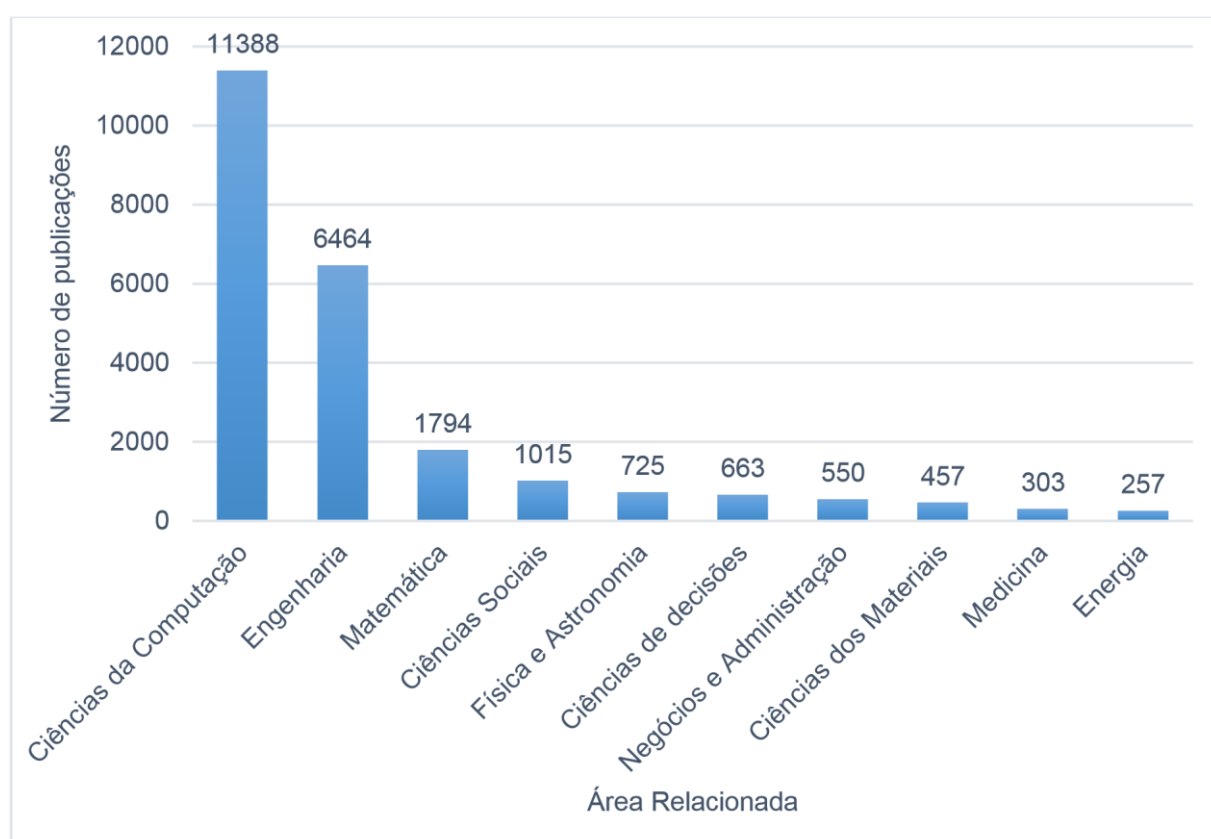
Outro aspecto importante é o controle ambiental, particularmente, a análise de poluição do ar; que viabiliza alertar as pessoas nas proximidades a respeito do risco e dos possíveis danos à saúde, principalmente crianças e idosos, que são mais frágeis. Indústrias que estiverem produzindo muitos poluentes poderão ser notificadas pelo governo para que tomem medidas corretivas, além da limitação de circulação de veículos de baixa tecnologia, que emitem grandes quantidades de poluentes.

O estacionamento inteligente, possibilita controlar o fluxo de veículos em diferentes zonas, o que é relevante para planejar onde se deve desenvolver novos estacionamentos. Os motoristas também podem ser privilegiados com a possibilidade de saber onde existem locais

disponíveis para estacionar perto de seu destino, reduzindo o consumo de combustível e o tempo.

O gráfico da Figura 3 destaca os tópicos mais citados junto ao termo Internet das Coisas, entre eles estão: Ciências da Computação. Engenharia em geral, Matemática, Ciências Sociais, Física e Astronomia. A partir dela, é possível perceber que apesar da maioria das citações estarem na área de engenharia e computação, outros meios acadêmicos já demonstram um interesse significativo ao tema.

Figura 3: Número de publicações de Internet das coisas em cada área.



O monitoramento veicular provavelmente é uma das fontes mais importantes das cidades inteligentes, por meio do qual as pessoas recebem, em tempo real, informações a respeito da melhor rota; conseqüentemente, menos congestionamentos e emissão de poluentes.

As entidades de fiscalização governamentais podem receber, instantaneamente, dados a respeito de acidentes ou outros ocorridos em ruas e rodovias, portanto, diminui o tempo das ações corretivas, das interrupções no trânsito; além de propiciar um maior número de vidas salvas em acidentes.

Sensores GPS (sigla do termo em inglês *Global Positioning System*, um sistema de posicionamento por satélite que fornece a um aparelho receptor móvel a sua posição), sensores de ultrassom e acelerômetros dispostos internamente e na parte frontal dos carros podem fornecer informações de localização do veículo, número de automóveis entre dois sensores, detecção de acidentes, através do impacto aos sensores do veículo, enviando imediatamente um alerta à polícia e aos hospitais próximos. Os dados podem ser coletados e transmitidos através da tecnologia GPRS (sigla do termo em inglês *General Packet Radio Service*), um padrão popular de rápida transferência de dados pela rede celular.

Nas cidades existem as preocupações referentes a segurança, porém analisar todas as gravações em busca do ocorrido é uma tarefa muito difícil. Dessa forma, é sugerido não apenas o espalhamento das câmeras de segurança por toda uma cidade, mas também botões de emergência para que os indivíduos possam notificar as agências de segurança sobre uma anomalia. Estas, por sua vez, irão monitorar o acontecimento, acionar a polícia, bombeiros ou ambulâncias. Estas informações detectadas a partir de diferentes sensores poderão também ser utilizadas para prever e evitar novas preocupações, através das ferramentas de inteligência computacional.

Shi e Abdel-Aty (SHI; ABDEL-ATY, 2015) e colaboradores realizaram estudo de viabilidade do monitoramento de tráfego em tempo real, avaliando operação e segurança simultaneamente. O objetivo deste estudo foi aumentar a performance do sistema de vias expressas urbanas, para reduzir congestionamentos e risco de colisão. Neste estudo utilizaram o *Microwave Vehicle Detection System* (MVDS) implantado numa rede de vias expressas em Orlando. Este sistema é uma fonte ideal para o Big Data graças a um sistema de detecção complexo, à riqueza de informações e à coleta contínua de dados. Os resultados da avaliação de congestionamento ajudam na análise de segurança em tempo real, na compreensão dos efeitos da dinâmica do tráfego nas ocorrências de colisão. Com este sistema podem ser instaladas câmeras de monitoramento num conjunto de rodovias para fornecer dados com precisão para toda a rede; além de ser criados métodos matemáticos para se estimar dados ausentes em trechos sem câmeras e, portanto, sem dados (SHAN; Q. ZHU, 2015).

Outra maneira de controle de tráfego de menor custo, é através do uso da infraestrutura de câmeras já instaladas nas cidades, criando uma rede de sensores sem fio e construindo uma plataforma de sistemas de informações geográfica para gerenciar e monitorar a mobilidade urbana, sem custos de hardwares adicionais. Este sistema integrado foi estendido com uma interface para a Simulação de Mobilidade Urbana (SUMO), um pacote de simulação de tráfego rodoviário em roteamento contínuo e microscópico de código aberto. Os fluxos de entrada são gerados pelo sistema a partir de dados reais de mobilidade. Desta forma, o

sistema gere e monitora a mobilidade urbana em tempo real proporcionando um instrumento útil para o planejamento das intervenções de planos de mobilidade / infraestruturas, sendo capaz de simular o futuro estado e verificar ainda a efetividade do que é realizado. O objetivo deste trabalho é projetar um sistema integrado de mobilidade com capacidades de simulação de trânsito. Esta plataforma baseia-se na integração de muitas fontes diferentes de dados de tráfego em tempo real. A principal novidade é a exploração de outros sensores não utilizados principalmente para os dados de tráfego, que podem ser processados para extrair informações sobre o tráfego.

O rápido crescimento da população e das atividades urbanas aumenta também a demanda dos sistemas de transporte urbano. Este rápido crescimento determinou diversos problemas na gestão dos sistemas de transporte urbano, uma vez que não são adequados a esta demanda crescente, para ser processado em um curto espaço de tempo. Como consequência surgem defeitos e dificuldades para utilização dos sistemas de transporte metropolitano, cujo bom funcionamento é um aspecto chave para a sociedade civil evoluída. O bom funcionamento do transporte urbano proporciona melhor controle do congestionamento de trânsito, conforto, sem superlotação, melhor planejamento de rotas, e a segurança do trânsito. A fim de superar alguns dos problemas de manutenção, uma intervenção válida é a de efetuar uma gestão eficiente do trânsito e também um planejamento.

Outro estudo sobre trânsito, desenvolvido por Zhou e colaboradores (ZHOU et al., 2016), cujo objetivo foi desenvolver métodos para detectar más condutas de motoristas de táxi. Os passageiros podem sofrer desvio de rota ou serem sobretaxados por motoristas não éticos, especialmente quando se encontram em cidades desconhecidas. Eles desenvolveram análise em tempo real da localização dos táxis, a fim de melhorar seus serviços. Este trabalho investigou um método de detecção online de trajetória anômala, denominado OnATrade (que significa atividades no comércio de corrida de taxi), a fim de melhorar o serviço de táxi usando GPS e Big Data. Este método consiste em duas etapas: recomendação de rota e detecção online. Na primeira etapa, offline, as possibilidades de rota são geradas a partir de várias trajetórias históricas, usando um algoritmo de recomendação de rota. Na segunda etapa, é apresentada uma abordagem de detecção online de trajetória para encontrar táxis que apresentam anomalias de condução, que consiste na comparação da trajetória em andamento com as possibilidades de rota. Além disso, passageiros ou o escritório da companhia de táxi podem participar deste processo para melhorar a qualidade do serviço. Os Passageiros podem avaliar com o tablet presente no táxi e o escritório também pode intervir no caso de uma emergência.

A diferença do planejamento urbano para as cidades inteligentes consiste na forma como são utilizados os dados gerados pelos sensores. As cidades inteligentes tomam

decisões em tempo real, a partir de dados gerados naquele instante, em contrapartida, o planejamento urbano irá fazer uso de dados previamente gerados pela mesma cidade e as decisões serão tomadas futuramente. Um exemplo seria analisar o consumo de energia do último ano, de forma a tomar iniciativas para atender essa demanda no ano seguinte, inclusive nos meses de maior consumo e os de menor produção energética.

Os sistemas de cidades inteligentes e planejamento urbano, através da internet das coisas, usando a análise e ferramentas Big Data, podem deixar um grande impacto no desenvolvimento das nações, uma vez que aumentam o poder de decisão da sociedade, através de atitudes inteligentes e efetivas no momento apropriado.

Big Data se tornou viável na área de transportes através da rápida popularização do *Intelligent Transportation System* (ITS) nas últimas décadas. Grande em tamanho e rico em informação, os dados aparentemente desorganizados podem melhorar consideravelmente a compreensão dos especialistas sobre seus sistemas (SHI; ABDEL-ATY, 2015).

3. CONSIDERAÇÕES FINAIS

A principal contribuição deste artigo é evidenciar os novos desafios que o engenheiro civil enfrenta e enfrentará em breve, e para os quais as instituições de formação desse profissional devem estar atentas, a fim de implantar melhorias nos métodos de ensino, visando preparar profissionais qualificados para as novas exigências do mercado.

Diante da vastidão de dados gerados a todo instante (Big Data), o engenheiro precisa ter conhecimento de programas que poderão auxiliá-lo na manipulação dessas informações (Data Mining) e na tomada de decisões técnicas e coerentes para uma rápida resposta e melhor eficiência de todo o processo. O domínio de tecnologias de seleção, manipulação e gerenciamento de informações, suas aplicações e contribuições nas diversas áreas de atuação do engenheiro, é essencial para atender as atuais exigências do mercado.

Nesta revisão bibliográfica destacou-se as contribuições dessas tecnologias em áreas estratégicas de atuação do engenheiro civil, como o gerenciamento de resíduos da construção civil, que possibilitará a otimização dos resíduos gerados. A consequência será a redução nos custos da obra, no impacto e poluição do meio ambiente.

A extração de informações e as técnicas de mineração e classificação de dados podem ser úteis, por exemplo, no monitoramento da estrutura de edifícios, pontes, barragens, ferrovias, hidráulica e ciências de materiais, problemas de geoprocessamento; que geram uma enorme massa de dados, definidos em domínios multidimensionais, cuja análise e interpretação pode selecionar informações importantes para compreensão do comportamento

dessas estruturas, permitindo a previsão e a resolução de eventuais problemas que possam ocorrer.

Outra possibilidade de aplicação se dá na aquisição, transformação e análise de dados geográficos através de técnicas matemáticas e computacionais para tratamento destas informações para cartografia, análise de recursos naturais, transportes, energias e planejamento urbano e regional. Algumas dessas ferramentas computacionais, por exemplo os sistemas de informações geográficas (SIG), permitem realizar análises complexas ao integrar dados de diversas fontes e ao criar bancos de dados georreferenciados (ZHAO; FOERSTER; YUE, 2012).

A velocidade com a qual a tecnologia se difunde na vida das pessoas é um desafio para o profissional que se propõe a atender as exigências da sociedade em que atua. Nesse contexto o conhecimento sobre as inovações que surgem no âmbito da internet das coisas e sua aplicação no sentido de ampliar a segurança, praticidade e eficiência de tarefas comuns do cotidiano ou imprevistos que possam ocorrer, é importante para o engenheiro que vai atuar no planejamento urbano (cidades, edifícios e trânsito inteligentes), meio ambiente (energia, água) e segurança.

A adequação do currículo das escolas de engenharia para a formação de um profissional com domínio dessas tecnologias é essencial para uma atuação mais eficiente e preparado para atender às exigências atuais e futuras.

A partir dessa análise, constatou-se que houve um aumento significativo de pesquisas sobre internet das coisas e cidades inteligentes, a partir de 2011. O tema Internet das Coisas possui a China como líder em pesquisas, seguida pelos Estados Unidos. Itália, Reino Unido e

Índia também têm expressão significativa nesse tema. Uma exemplificação da crescente importância dessa área seria que, em 2006 foram encontradas 131 referências, em 2011 chegaram a 1001, e em 2016 atingiu 4936 publicações, caracterizando uma curva de crescimento exponencial.

A China e Estados Unidos se mostraram as duas maiores potências nesse meio acadêmico que engloba o Big Data, A Internet das Coisas, controle de tráfego e a utilização de GPS, estando sempre entre os primeiros e com um grande volume de publicações. O Brasil teve um breve ressaltado no monitoramento de tráfegos em tempo, estando entre os 10 países que mais possuem artigos relacionados ao tópico.

Ao abordarem conceitos específicos dentro de Internet das Coisas, como por exemplo, cidades inteligentes, alguns países começam a se destacar mais que outros, mostrando que existem diferentes interesses e focos de pesquisa principais para cada nação. Para determinados assuntos, Itália e Japão estiveram à frente dos EUA e da China, da mesma

forma que Reino Unido, Índia, Espanha, Alemanha e Coréia do Sul se destacaram, mesmo que em menores proporções.

4. REFERÊNCIAS

BELLUZZO, B. G. et al. Big data e suas aplicações na engenharia civil. In: Brazilian Technology Symposium - BTSym 2016, 2016, Campinas - SP. Brazilian Technology Symposium. Campinas: LCV, 2016. v. 1. p. 1-1.

BELLUZZO, B. G.; IGARASHI, M. O. BIG DATA, MINERAÇÃO DE DADOS E APLICAÇÕES NA ENGENHARIA CIVIL. In: Anais do XLIV Congresso Brasileiro de Educação em Engenharia (COBENGE 2016). Brasília: ABENGE. 2016.

BILAL, M. et al. Big data architecture for construction waste analytics (CWA): A conceptual framework. *Journal of Building Engineering*, v. 6, p. 144-156, 2016.

GALVÃO, N. D.; MARIN, H. F. Data mining: a literature review. *Acta Paulista de Enfermagem*, v. 22, n. 5, p. 686-690, 2009.

HASHEM, I. A. T. et al. The rise of "big data" on cloud computing: Review and open research issues. *Information Systems*, v. 47, p. 98-115, 2015.

JALALI-HERAVI, M.; ARRASTIA, M.; GOMEZ, F. A. How can chemometrics improve microfluidic research?. *Anal Chem*, v. 87, p. 3544-55, 2015.

KENNETH, C. L.; JANE, P. L. *Management Information System Managing the Digital Firm*. Printice-Hall New Delhi, 2006.

LAROSE, D. T. *Discovering knowledge in data: an introduction to data mining*. John Wiley & Sons, 2014.

MOORE, G. Moore's law. *Electronics Magazine*, v. 38, n. 8, p. 114, 1965.

O'BRIEN, J. A. *Sistemas de informação e as decisões gerenciais na era da internet*. Saraiva, 2004.

O'BRIEN, J. A.; MARAKAS, G. M. *Administração de sistemas de informação: uma introdução*, São Paulo: McGraw-Hill, 2007.

POSENATO, D. et al. Methodologies for model-free data interpretation of civil engineering structures. *Computers & structures*, v. 88, n. 7, p. 467-482, 2010.

RATHORE, M. M. et al. Urban planning and building smart cities based on the internet of things using big data analytics. *Computer Networks*, v. 101, p. 63-80, 2016.

SHAN, Z.; ZHU, Q. Camera location for real-time traffic state estimation in urban road network using big GPS data. *Neurocomputing*, v. 169, p. 134-143, 2015.

SHI, Q.; ABDEL-ATY, M. Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. *Transportation Research Part C: Emerging Technologies*, v. 58, p. 380-394, 2015.

WU, X. et al. Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*, v. 26, n. 1, p. 97-107, 2014.

ZHAO, P.; FOERSTER, T.; YUE, P. The geoprocessing web. *Computers & Geosciences*, v. 47, p. 3-12, 2012.

ZHOU, Z. et al. A method for real-time trajectory monitoring to improve taxi service using GPS big data. *Information & Management*, v. 53, n. 8, p. 964-977, 2016.

Contatos: brugabell@gmail.com e massaki.igarashi@mackenzie.br