

10-21-2023

Big Data and Traffic Data Analysis: Urban Traffic Forecasting and Management

Luís M. Barata

Instituto Politécnico de Castelo Branco, luis.barata@ipcb.pt

João Gonçalves

Instituto Politécnico de Castelo Branco, joaogoncalves2@ipcbcampus.pt

João Palhares

Instituto Politécnico de Castelo Branco, jpalhares@ipcbcampus.pt

Eurico Lopes

Instituto Politécnico de Castelo Branco, eurico@ipcb.pt

Arminda G. Lopes

Instituto Politécnico de Castelo Branco, aglopes@ipcb.pt

Follow this and additional works at: <https://aisel.aisnet.org/capsi2023>

Recommended Citation

Barata, Luís M.; Gonçalves, João; Palhares, João; Lopes, Eurico; and Lopes, Arminda G., "Big Data and Traffic Data Analysis: Urban Traffic Forecasting and Management" (2023). *CAPSI 2023 Proceedings*. 26. <https://aisel.aisnet.org/capsi2023/26>

This material is brought to you by the Portugal (CAPSI) at AIS Electronic Library (AISeL). It has been accepted for inclusion in CAPSI 2023 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Big Data e Análise de Dados de Tráfego: Previsão e Gestão do Tráfego Urbano

Big Data and Traffic Data Analysis: Urban Traffic Forecasting and Management

Luis M. Barata, Instituto Politécnico de Castelo Branco, Portugal, luis.barata@ipcb.pt

João Gonçalves, Instituto Politécnico de Castelo Branco, Portugal,
joaogoncalves2@ipcbcampus.pt

João Palhares, Instituto Politécnico de Castelo Branco, Portugal, jpalhares@ipcbcampus.pt

Eurico Lopes, Instituto Politécnico de Castelo Branco, Portugal, eurico@ipcb.pt

Arminda G. Lopes, Instituto Politécnico de Castelo Branco, Portugal, aglopes@ipcb.pt

Resumo

O artigo analisa diferentes metodologias e tecnologias utilizadas em sistemas de gestão e análise de tráfego urbano baseados em *Big Data*. São discutidas diversas técnicas, como a instalação de sensores, dispositivos IoT e câmaras de vigilância para recolha de dados, o uso de algoritmos e tecnologias de *machine learning* para identificação e classificação de veículos, o uso de técnicas como *Map Reduce* para o tratamento de dados e a comunicação entre diversas aplicações de redes sociais para obtenção de dados em tempo real sobre a localização e deslocação de utilizadores. Além disso, são propostas modelos baseados em redes neuronais e algoritmos como ABC para a previsão de mudança de sinais e gestão de fluxo de tráfego. São ainda destacadas as vantagens e desvantagens de cada abordagem e como elas podem ser aplicadas para melhorar a gestão de tráfego urbano e reduzir a poluição causada por veículos.

Palavras-chave: *Big Data*; Tráfego Urbano; Análise; Processamento.

Abstract

The paper reviews different methodologies and technologies used in urban traffic management and analysis systems based on Big Data. Several techniques are discussed, such as the installation of sensors, IoT devices and surveillance cameras for data collection, the use of machine learning algorithms and technologies for vehicle identification and classification, the use of Map Reduce techniques for data processing, and the communication between several social network applications to obtain real-time data about the location and movement of users. In addition, models based on neural networks and algorithms such as ABC are proposed for signal change prediction and traffic flow management. The advantages and disadvantages of each approach and how they can be applied to improve urban traffic management and reduce vehicle pollution are highlighted.

Keywords: *Big Data*; Urban Traffic; Analysis; Processing.

1. INTRODUÇÃO

O congestionamento no trânsito, apesar de ser bastante comum na vida da maioria das pessoas, continua a ser atualmente um grande desafio, uma vez que perturba a sua vida, aumenta a poluição, tanto sonora como atmosférica, e aumenta o consumo de combustíveis (Robinson, 1984). Existem várias origens para este fenómeno da atualidade, tais como o aumento da população e do número de veículos de transporte, a má construção de cidades e ruas, a ocorrência de acidentes rodoviários, as mudanças de clima, a existência de eventos comemorativos, entre outros (Chen et al., 2019).

Neste contexto, é nítida a necessidade de realizar uma boa gestão do tráfego, de forma a evitar alguns dos problemas citados anteriormente. Assim, a gestão de tráfego é uma das principais preocupações dentro de uma cidade da atualidade (Kostakos et al., 2013) aspeto esse cada vez mais visível principalmente devido ao surgimento das *smart cities*. Estas, não obstante, procuram fazer uso desta área de forma altamente eficiente e automatizada. A existência de diversos problemas associados a este tema, nomeadamente a forma de aquisição dos dados, o enorme número de variáveis que podem surgir no dia a dia e a dificuldade de manutenção e tratamento de tantos dados. Eliminar o congestionamento e o fluxo desequilibrado de tráfego em intersecções são um dos maiores desafios que os desenvolvedores de software e os planeadores urbanos podem enfrentar. Para uma melhor funcionalidade dos sistemas de controlo de tráfego urbano e eliminar congestionamentos é necessária uma vigilância robusta e fiável em tempo real. Para isso ser possível é necessário fazer uso de diversos sistemas, altamente complexos. No entanto, surgem outros problemas associados a eles, tais como a grande quantidade de dados que são gerados e analisados (Güven et al., 2021). É exatamente nestes aspetos que é necessária a *Big Data* como uma solução para estes problemas.

As definições e ideias que estão por detrás da *Big Data* começaram a surgir na década de 90 devido ao grande aumento de produção de dados, coisa que se verificou com o desenvolvimento exponencial da internet (Batistič & van der Laken, 2019). Com estas alterações repentinas, começou-se a pensar que seria necessário usar diferentes métodos de tratamento dos dados, tendo então sido criado o conceito de *Big Data*, citado pela primeira vez por Roger Mougalas, em 2005 (Batistič & van der Laken, 2019). Rapidamente o termo foi-se popularizando, ganhando cada vez mais importância e notoriedade dentro da área da tecnologia. *Big Data* é um conceito que está relacionado com uma grande quantidade de dados geralmente com organização e estrutura mais complexa, que permitem a análise de padrões e tendências, sendo este processo chamado de *Big Data Analytics* (BDA).

Assim, é possível verificar uma relação entre *Big Data* e gestão de tráfego, principalmente devido ao facto de ser gerada uma enorme quantidade de dados, mesmo em cidades pequenas com uma menor gestão. Existem diversos programas, usados pelos governos ou instituições próprias, que procuram fazer uma gestão em tempo real do tráfego urbano, de forma a evitar problemas como o

congestionamento e permitir uma resposta rápida perante a ocorrência de imprevistos, tais como acidentes, avarias de sistemas de sinalização ou problemas com o local de circulação (Guven et al., 2021). Mesmo sendo sistemas “eficientes”, são responsáveis pela geração de uma quantidade massiva de dados e informações, que podem provenir de diversas fontes e apresentarem diferentes tipos, o que torna os modelos de tratamento e análise de dados altamente ineficientes perante estas condições.

Neste artigo procura-se apresentar uma relação mais íntima entre os sistemas de gestão de tráfego urbano e os conceitos de *Big Data* bem como tentar verificar o que já está a ser feito na área. Procura-se ainda tentar perceber de que maneira este tema tende a evoluir num possível trabalho futuro.

Este artigo está estruturado da seguinte forma: ponto **2. Metodologia de Pesquisa**, onde é abordado todo o processo de filtragem das pesquisas, bem como escolha de *keywords* e fontes da informação referenciada. No ponto **3. Revisão das Tecnologias de Big Data** onde serão faladas todas as tecnologias encontradas e relevantes para o assunto, sendo explicado em pormenor cada uma delas e de que modo podem ser aplicadas. No ponto **4. Discussões** é discutido alguns aspetos de aplicação de conceitos de *Big Data*, dentro do tema de previsão e gestão de tráfego urbano e o ponto **5. Conclusões** que aborda todas as conclusões tiradas após a realização do trabalho.

2. MÉTODOS DE PESQUISA

O foco principal deste trabalho é fazer uma análise do papel da *Big Data* na análise de dados de tráfego e verificar algumas tecnologias que estão atualmente a ser aplicadas nesta área, de forma a realizar previsões e análises com base nisso.

Para tal, é preciso primeiro definir alguns aspetos base e fundamentais de forma a ser possível realizar a procura da informação necessária. Todo este processo de filtragem e procura será abordado e explicado neste ponto, onde começamos com a secção **2.1 Fonte de Pesquisa** que pretende demonstrar uma lista das *keywords* utilizadas bem como explicar o método de procura adotado.

2.1. Fonte de Pesquisa

Para iniciar a procura e filtragem de dados relacionados com o assunto alvo, primeiro foi feita uma pesquisa sobre o termo “*Big Data*”, de forma a descobrir a sua origem e artigos que pudessem fornecer informações uteis para realizar uma breve contextualização ao assunto e a algumas definições pertinentes para o mesmo. Essa procura foi realizada inicialmente no *IEEE Explore* e *ScienceDirect*, onde foi possível extrair dados tais como definições básicas e história da *Big Data*. Ao fazer esta pesquisa foram encontrados 946 365 artigos que faziam referência ao tema *Big Data*. Para além disso, ainda numa fase inicial, foi feita uma procura por *urban traffic*, com o intuito de

encontrar dados e informações sobre a sua origem e os problemas que surgem com esta adversidade da atualidade.

Com isso feito, foi necessário acrescentar mais informação/*keywords* no processo de pesquisa, uma vez que nem todos os artigos encontrados remetiam para o tema deste trabalho, tendo sido realizado um foco novamente nas palavras *Big Data* e ainda em *urban traffic*, adicionando posteriormente um filtro de artigos mais recentes, mais especificamente de 2017 até a atualidade.

Ainda sobre o processo de pesquisa, foram realizadas algumas procuras que focavam exclusivamente no assunto de *urban traffic*, com o intuito de conseguir extrair alguns dados que poderiam ajudar a entender melhor o tema. Com esta pesquisa, pretendeu-se ainda conhecer quais as principais dificuldades deste tipo de tecnologias e em que situações estas podem ser aplicadas nos dias atuais.

Com esta pesquisa inicial feita, foi necessário começar a fazer uma filtragem dos resultados obtidos, de forma a estudar e analisar apenas aqueles que eram mais promissores ou relevantes dentro da proposta apresentada. Esta filtragem foi feita primeiro com base no título de cada artigo, visto que havia alguns que mesmo tendo as *keywords* escolhidas, não abordavam exatamente o assunto que se pretendia, fazendo geralmente sempre algum “desvio”. Após isso, foi analisado o *abstract* de cada um dos trabalhos com o objetivo de saber mais detalhes sobre os mesmos e saber até que ponto são realmente os ideais para serem usados como fonte de informação.

Acabando esta etapa, foi atingido um resultado de 9 artigos, com diferentes anos de publicação tal como demonstrado na Figura 1.



Figura 1 – Número de artigos referenciados por ano de lançamento

3. RESULTADOS

Tal como referido anteriormente, um dos principais motivos para a importância da *Big Data* na atualidade é o desenvolvimento rápido e espontâneo da Internet que se tem verificado nos últimos anos. Este desenvolvimento cria diferentes fontes de dados bem como uma enorme variedade dos mesmos, ou seja, há vários tipos de informação que não se consegue analisar e tratar usando bases de dados convencionais. Um outro aspeto fundamental para o fornecimento massivo de informação é o surgimento de dispositivos e tecnologias IoT (*Internet of Things*). Estas tecnologias estão na base de diversos estudos e pesquisas do mundo atual e têm tido cada vez uma maior importância dentro dos sistemas informáticos (Lee, 2017).

Um outro ponto importante a ser abordado são os diferentes aspetos/aplicações que existem na *Big Data*, que são fundamentais para ajudar a definir que tipo de trabalho se pretende fazer e em que etapa vai assentar cada um desses trabalhos. Sendo assim, dentro dos que são relevantes para o trabalho em questão, existe a fase de recolha de dados, responsável pela colheita de informação e meios para atingir esse fim, a fase de análise de dados, responsável por tratar todo os tipos de dados de forma a ser possível adaptados a cada situação e por sua vez permitir aplicá-los na tomada de decisões (Lee, 2017).

São exatamente estas fontes de dados e estes aspetos que são o principal ponto abordado na grande maioria dos projetos e artigos que serão referenciados e abordados nesta secção do trabalho.

3.1. Tecnologias de Big Data

Em (Kulkarni et al., 2023), abordam um método de aplicação de sistemas de gestão e análise de tráfego urbano, utilizando sistemas de transmissão e recolha de dados baseados em dispositivos Bluetooth. Para tal, é proposta a instalação de sensores, em pontos críticos, tais como semáforos ou sinais de trânsito, que teriam a função de recolher diferentes tipos de dados, desde velocidade dos veículos, densidade do trânsito ou possíveis congestionamentos. De maneira a captar com precisão toda a informação relevante, os sensores são programados para detetarem dispositivos Bluetooth, tais como telemóveis ou veículos que tenham essa tecnologia embutida. Segundo os autores, os dados seriam recolhidos 2 vezes ao dia, sendo guardados sobre a forma de um ficheiro CSV e analisados com *R Studio*. O tratamento dos dados é feito tendo em consideração a quantidade de dispositivos detetados, sendo a maioria deles telemóveis e acessórios sem fios e ainda outros dados, tais como a direção em que o dispositivo se está a mover.

Em Hlaing et al. (2020), é preposto um sistema de análise de tráfego, usando tecnologias de *Big Data* e *deep learning*. O método proposto consiste num conjunto de processos que permitem ajudar na gestão do trânsito e redução da poluição causada pelo mesmo, onde é feito uso de sensores, dispositivos IoT e câmaras de vigilância para a recolha de informação relacionada com os veículos,

fazendo uso de mecanismos de *Map Reduce* para o tratamento dos dados. Para finalizar a sua proposta, são usados algoritmos e tecnologias de *machine learning* para ajudar na identificação e classificação de cada veículo.

Zhang e Jiang (2020) deram a ideia de design de um sistema de controlo de tráfego inteligente com base em tecnologias e métodos de *Big Data*. O sistema funciona com base numa arquitetura específica, definida pelos autores, que fornece um conjunto de 3 camadas, cada uma responsável por uma parte fundamental para o seu normal funcionamento. Estas camadas são: a camada de suporte de dados, camada de processamento de dados e camada de aplicação. A função da primeira camada do sistema é a recolha de dados, onde são usadas câmaras para captura de vídeo e imagens em tempo real. Na camada seguinte, processamento dos dados, é realizado todo o processo de análise e cálculo dos dados, de forma a ser possível tirar conclusões relevantes dos mesmos e é ainda tratada toda a etapa de armazenamento de forma a criar um histórico e estado atual da situação que está a ser analisada. Dentro das análises realizadas, é mais relevante salientar a velocidade dos automóveis e a sua variação entre os diferentes pontos de monitorização, a nível de ocupação dos veículos e a previsão de congestionamento causada por cada automóvel.

No projeto Vidya e Ashwini (2018), são discutidas diversas técnicas de *Big Data* que podem ser aplicadas em sistemas de controlo de tráfego, ajudando a prevenir e reduzir alguns dos problemas associados a este tema. Uma das técnicas propostas é o *Map Reduce*, onde é tido como input uma grande quantidade e variedade de dados, que são armazenados num único local HDFS (*Hadoop Distributed File System*). De seguida os dados são separados com base nos eventos de tráfego que estão na base do mesmo, onde vão posteriormente sofrer uma redução, resultando numa quantidade menor e dados mais refinados. Uma outra estratégia abordada pelos autores é fazer uso da comunicação entre diversas aplicações de redes sociais, fazendo assim uso de informações transmitidas pelos telemóveis de diversos utilizadores. Este método permite obter dados em tempo real sobre a localização e deslocação de cada utilizador, o que permite prever quando poderá ocorrer algum evento de congestionamento. Para finalizar, é proposto um terceiro método, baseado em sistemas de redes neuronais que fazem uso do algoritmo ABC. Este modelo, de acordo com os autores, permite realizar previsões relacionadas com a mudança de sinais, de forma a otimizar ao máximo os sistemas de gestão de tráfego. Para além disso, esta abordagem procura ainda criar uma gestão de fluxo, permitindo alterar esse mesmo fluxo de acordo com as necessidades do momento.

Uma outra abordagem é proposta por Chen et al. (2019) onde é demonstrada a possibilidade de criar sistemas que tenham como base uso de tecnologias de *Big Data* e dispositivos IoT. Estes dispositivos são os responsáveis por fornecerem e recolherem os dados que serão tratados nas etapas seguintes do sistema, permitindo ajudar a fazer uma melhor gestão de semáforos em situações críticas de congestionamento. Para além disso, é ainda tido em conta a existência de alguns veículos e variáveis que são mais difíceis de controlar, tais como veículos não motorizados ou peões que não cumprem

as regras de circulação na via pública. Para o tratamento de dados é usado um modelo baseado no algoritmo *K-Means*, onde todo o input recebido é visto como uma série de eventos temporais, que podem sofrer um processo de segmentação, de forma a facilitar a previsão de tráfego por parte de modelos de redes neuronais. Desta forma, é possível fazer a análise do tráfego de acordo com o histórico de dados, que estão armazenados no sistema e com os dados que são recebidos em tempo real. Com estas etapas realizadas, o sistema faz uso das capacidades dos sistemas IoT para funcionar como é pretendido, através de um controlo organizado por níveis e setores: o nível central que permite controlar toda uma vasta área como uma cidade e todos os seus serviços associados, o nível regional que utiliza as informações do nível anterior e complementa-as de acordo com a situação do momento, fornecendo uma solução para o próximo nível. Por fim existe o nível de interseção que é responsável por fazer o controlo final do tráfego, ajustando as suas medidas de acordo com a demanda e necessidade do momento, como por exemplo o tempo de uma luz verde ou vermelha para os automóveis em determinado semáforo.

Gong et al. (2017) propõem uma metodologia de controlo de tráfego com um sistema que se baseia numa arquitetura SMASH, que é um software muito usado dentro de trabalhos e assuntos relacionados com *Big Data*, uma vez que fornece a capacidade de analisar enormes conjuntos de dados, altamente variáveis de forma escalável, podendo com isso se adaptar a situações de recolha em tempo real. Para além destas características, esta arquitetura ainda demonstra ser ideal para a realização das suas tarefas de forma remota, uma vez que os seus componentes e serviços se encontram alojados na Cloud.

Em Sreekumar et al. (2017) os autores propuseram um modelo que utiliza Contextual Vídeo Data, ou seja, informações relacionadas ao vídeo, neste caso a semelhança do cenário entre *frames*, a direção dos objetos/veículos no vídeo, informação de oclusão, entre outros (Sreekumar et al., 2017), a fim de aplicar diferentes algoritmos para alcançar uma vigilância de tráfego superior. Foi utilizada a estrutura de rede neuronal *DarkNet* para executar o modelo YOLOv2 e CNN para que o sistema compreenda quando um veículo entra ou sai das faixas de rodagem e em geral o fluxo de tráfego.

Quanto ao modulo de deteção e classificação, foram selecionados os modelos YOLOv2 e o Faster-RCNN, que foram testados/avaliados com uma ferramenta própria dos autores, onde o YOLOv2 teve as melhores avaliações.

Os autores utilizaram o código *open-source DarkNet* e adicionaram camadas de software para suportar o seu modulo de rastreio de objetos.

No modulo de deteção de alteração de cena, foi utilizado um detetor de margem, mais especificamente *Canny* para visualizar o movimento feito um por objeto entre os *frames* adjacentes, utilizam uma *border following technique* para obter os contornos das bordas detetadas. Depois são

representados os objetos em movimento recolhendo os contornos usando um algoritmo de aproximação Douglas-Peucker.

No artigo de Essien et al. (2019) tiveram como objetivo melhorar a previsão da velocidade do tráfego nas redes rodoviárias urbanas, fazendo uso de *deep learning* e *data fusion*. Para isso os autores recolheram e utilizaram um conjunto de dados de tráfego e meteorológicos de uma rede rodoviária urbana em Grande Manchester para treinar e testar o algoritmo de *deep learning*, neste caso o LSTM-NN (Long Short-Term Memory Neural Network), que foi utilizado para a previsão da velocidade do tráfego visto que pode lidar com dados sequenciais. Depois foi feita a comparação do desempenho do modelo LSTMN-NN com a utilização de fontes de dados de apenas tráfego, ou seja, sem as condições meteorológicas, para a previsão de velocidade de tráfego. Os resultados experimentais mostraram que a abordagem proposta superou o desempenho da utilização de fontes de dados, indicando o valor da função de dados a nível de input e *Deep Learning* para a previsão de tráfego.

Em Duan et al. (2020) foi realizado um estudo sobre os fatores que contribuem para o congestionamento do tráfego urbano nas cinco áreas urbanas do centro de *Chengdu*. Para tal, os autores recolheram dados de várias fontes, incluindo 1,4 mil milhões de registos GPS de táxis em *Chengdu*, dados de fluxo de tráfego rodoviário e dados POI (*Point of Interest*). Os dados POI referem-se a dados baseados na localização que identificam pontos de interesse específicos, tais como restaurantes, lojas e atrações turísticas nas áreas urbanas de *Chengdu* da plataforma AMAP (AutoNavi Maps). Em seguida, utilizaram um *Python crawler* para extrair estes dados e estabeleceram uma análise de componentes principais, análise de correlação cinzenta, e modelo de previsão da rede neuronal para analisar os dados e identificar os principais fatores que contribuem para o congestionamento do tráfego. Os autores ainda utilizaram uma Rede Neuronal de *BackPropagation* para prever o congestionamento de tráfego. Por fim descobriram que sete indicadores, tais como o período de tempo, o número de semáforos, a largura da estrada e o volume de tráfego, tem um impacto significativo no congestionamento do tráfego rodoviário.

4. DISCUSSÃO

Este trabalho apresenta diversas maneiras e metodologias que podem ser usadas para a criação de sistemas de gestão de tráfego eficientes e precisos, que geram uma quantidade tremenda de dados com uma grande variedade e tipo que, portanto, necessitam de aplicar algumas técnicas de *Big Data* para tratarem de todo esse processo e reduzir os problemas que estão associados a isso. De todos os trabalhos revistos neste artigo, apenas algumas técnicas são aplicadas, abrindo caminho para num futuro serem exploradas outras abordagens que podem se mostrar melhores do que as exploradas até então. Com tudo isto revisado, surgem algumas dúvidas que ficaram por responder mesmo após toda a análise feita ao longo do trabalho, questões essas que vão ser agora discutidas, de forma a tentar esclarecer o máximo possível e tentar encontrar uma possível resposta para as mesmas.

Começando pela primeira e talvez mais óbvia de todas, que desafios existem em sistemas de controle e previsão de tráfego urbano processo. Como já visto anteriormente, há diversos desafios associados a este tipo de tecnologias, sendo um dos principais a recolha dos dados. Este problema acontece visto que atualmente existem poucos métodos que sejam muito eficazes no processo de recolha de dados necessários nestes tipos de sistemas, de forma precisa. Isto deve-se principalmente devido à enorme quantidade de variáveis que podem aparecer durante esse processo, uma vez que existem, não apenas automóveis, mas também peões e outros fatores que podem influenciar os dados. Tendo como exemplo os peões e veículos não motorizados, como bicicletas/*skates*, estes são propícios a apresentar um comportamento mais irregular, que dificulta a fazer a previsão do seu movimento ou ações, especialmente porque estes tipos de variáveis nem sempre seguem o padrão de respeitarem as regras devidas e circularem apenas onde devem, característica essa que dificulta bastante durante a fase de realizar previsões. Estes aspetos, apesar de cada vez menos notórios, ainda são um dos principais desafios encontrados por este tipo de tecnologia, coisa que mesmo com a aplicação de *Big Data*, ainda se demonstra altamente pouco desenvolvido.

Uma das possíveis soluções para isto é a aplicação de algoritmos de inteligência artificial, de forma a tentar alcançar um resultado cada vez mais preciso e que consiga contornar estes desafios, coisa que já está a ser implementado, tal como foi possível verificar em alguns dos estudos revistos ao longo do documento, como em Hlaing et al. (2020), onde é abordado um modelo à base de *machine learning* para identificar e diferenciar os diferentes tipos de veículos. Um outro exemplo disso, é proposto em Essien et al. (2019), onde neste caso o modelo de *deep learning* tem como função melhorar a previsão do tráfego mesmo perante diferentes velocidades. Apesar disso, muitos desses algoritmos ainda são pouco eficazes e necessitam de sofrer algum tipo de alteração para se tornarem mais viáveis de aplicar.

Um outro ponto, ainda dentro do assunto da recolha de dados, é a maneira como os dados são recolhidos, onde foi observado que, atualmente o principal método para recolher os dados é o uso de dispositivos IoT ou através da comunicação gerada por dispositivos móveis, que estão constantemente a fornecer e partilhar dados, funcionando como uma fonte de informação. Isto é proposto em Hlaing et al. (2020) e em Chen et al. (2019), onde são usados sistemas IoT para captação dos dados de forma precisa e eficaz.

Uma outra questão que surge é de que maneira é possível aplicar *Big Data* nos sistemas de previsão e controle de tráfego, visto ser um dos principais focos deste artigo. Com isso em mente, sabendo qual é a principal função da *Big Data* dentro dos sistemas atuais, é possível tirar algumas conclusões referentes a este ponto. Primeiramente, é importante voltar a salientar que qualquer programa ou sistema de controlo e previsão de tráfego em tempo real vai gerar uma enorme quantidade de dados, que dependendo da fonte que está a ser usada, pode resultar em uma enorme variedade dos mesmos. É nestes casos, quando o assunto é o tratamento e análise destes dados, que se torna importante fazer

uso de algumas metodologias de *Big Data*. Um dessas tecnologias é o *Reduce*, que foi uma das propostas feitas em Vidya e Ashwini (2018) com o intuito de diminuir a quantidade massiva de dados que eram gerados pelos seus sistemas, ou em Hlaing et al. (2020), que é usado com os mesmos propósitos.

Com relação ao trabalho futuro, achamos que este é um assunto com bastante importância e potencial para o desenvolvimento desta área, especialmente pelo facto de ser algo altamente recorrente quando se fala em cidades do futuro. Por esse motivo, achamos importante haver um maior foco em solucionar alguns dos problemas dos sistemas atuais, fazendo uso de modelos e inteligência artificial para ser possível aumentar a sua precisão na obtenção de resultados. Para além disso, o uso de *Big Data* para tratamento de dados, se mostrou muito eficiente, sendo um dos principais fatores para ser possível tornar estes sistemas totalmente autónomos e viáveis.

5. CONCLUSÃO

Os estudos e projetos analisados mostram que as tecnologias de Big Data têm um papel importante na gestão do tráfego urbano, permitindo recolher e analisar grandes quantidades de dados provenientes de sensores, dispositivos IoT e câmaras de vigilância. O uso de algoritmos de *machine learning* e *deep learning* permite identificar e classificar cada veículo, analisar a velocidade e ocupação dos mesmos e prever possíveis congestionamentos. Também foi possível constatar que a técnica de *Map Reduce* é uma ferramenta útil para tratar e reduzir a quantidade de dados a serem analisados. A comunicação entre diferentes aplicações de redes sociais e o uso de sistemas de redes neuronais também são métodos interessantes para prever a ocorrência de eventos de congestionamento e otimizar a gestão de tráfego. Em suma, essas tecnologias têm o potencial de ajudar a melhorar a eficiência e a segurança do trânsito nas áreas urbanas, permitindo uma melhor gestão dos recursos, reduzindo a poluição, uma melhoria da mobilidade urbana e qualidade de vida nas cidades.

REFERÊNCIAS

- Batistič, S., & van der Laken, P. (2019). History, Evolution and Future of Big Data and Analytics: A Bibliometric Analysis of Its Relationship to Performance in Organizations. *British Journal of Management*, 30(2), 229–251. Doi: 10.1111/1467-8551.12340
- Chen, C., Liu, H., & Wang, Z. (2019). Analysis and design of urban traffic congestion in urban intelligent transportation system based on big data and internet of things. *ACM International Conference Proceeding Series*, 659–665. Doi: 10.1145/3349341.3349486
- Duan, X., Xu, J., Chen, Y., & Jiang, R. (2020). Analysis of influencing factors on urban traffic congestion and prediction of congestion time based on spatiotemporal big data. *Proceedings - 2020 International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering, ICBAIE 2020*, 75–78. Doi: 10.1109/ICBAIE49996.2020.00022
- Essien, A., Petrounias, I., Sampaio, P., & Sampaio, S. (2019). Improving Urban Traffic Speed Prediction Using Data Source Fusion and Deep Learning. *2019 IEEE International Conference on Big Data and Smart Computing, BigComp 2019 - Proceedings*. Doi: 10.1109/BIGCOMP.2019.8679231

- Gong, Y., Rimba, P., & Sinnott, R. O. (2017). A big data architecture for near real-time traffic analytics. *UCC 2017 Companion - Companion Proceedings of the 10th International Conference on Utility and Cloud Computing*, 157–162. Doi: 10.1145/3147234.3151010
- Guyen, C., Karaduman, O., & Avci, E. (2021). A Review on Urban Intelligent Traffic Management Problems: Sensors and Methods. *2021 2nd International Informatics and Software Engineering Conference (IISEC)*, 1–6. Doi: 10.1109/IISEC54230.2021.9672412
- Hlaing, S. S., Tin, M. M., Khin, M. M., Wai, P. P., & Sinha, G. R. (2020). Big traffic data analytics for smart urban intelligent traffic system using machine learning techniques. *2020 IEEE 9th Global Conference on Consumer Electronics, GCCE 2020*, 299–300. Doi: 10.1109/GCCE50665.2020.9291790
- Kostakos, V., Ojala, T., & Juntunen, T. (2013). Traffic in the smart city: Exploring city-wide sensing for traffic control center augmentation. *IEEE Internet Computing*, 17(6), 22–29. Doi: 10.1109/MIC.2013.83
- Kulkarni, A. R., Kumar, N., & Rao, K. R. (2023). Efficacy of Bluetooth-Based Data Collection for Road Traffic Analysis and Visualization Using Big Data Analytics. *Big Data Mining and Analytics*, 6(2), 139–153. Doi: 10.26599/BDMA.2022.9020039
- Lee, I. (2017). Big data: Dimensions, evolution, impacts, and challenges. *Business Horizons*, 60(3), 293–303. Doi: 10.1016/J.BUSHOR.2017.01.004
- Robinson, R. (1984). Problems in the urban environment: traffic congestion and its effects. *Wollongong Studies in Geography*. <https://ro.uow.edu.au/wollgeo/14>
- Sreekumar, U. K., Devaraj, R., Li, Q., & Liu, K. (2017). TPCAM: Real-time traffic pattern collection and analysis model based on deep learning. *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, 1–4. Doi: 10.1109/UIC-ATC.2017.8397674
- Vidya, K. T., & Ashwini, B. P. (2018). Survey on Real Time Traffic Analysis on Big Data. *Proceedings of the 2nd International Conference on Trends in Electronics and Informatics, ICOEI 2018*, 852–858. Doi: 10.1109/ICOEI.2018.8553757
- Zhang, D., & Jiang, Y. (2020). Design of urban intelligent traffic congestion situation monitoring system based on big data. *Proceedings - 2020 International Conference on Intelligent Transportation, Big Data and Smart City, ICITBS 2020*, 12–15. Doi: 10.1109/ICITBS49701.2020.00011