
Big data aplicado à cidade digital estratégica: estudo sobre o volume de dados das aplicações smart city

Edson Pedro Ferlin

Pos-Doutorado em Cidade Digital Estratégica pela Pontifícia Universidade Católica do Paraná. Professor das Faculdades Integradas Camões do Paraná
eferlin@live.com

Denis Alcides Rezende

Professor Titular e bolsista de Pesquisa do CNPq Brasil da Pontifícia Universidade Católica do Paraná
dar@denisalcidesrezende.com.br

Editor Científico: José Edson Lara
Organização Comitê Científico
Double Blind Review pelo SER/OJS
Recebido em 30.10.2018
Aprovado em 12.03.2019



Este trabalho foi licenciado com uma Licença Creative Commons - Atribuição – Não Comercial 3.0 Brasil

Resumo

No contexto da Cidade Digital Estratégica, destacam-se as *Smart Cities* que são cidades que utilizam Tecnologias de Informação para facilitar o desempenho dos serviços públicos, reduzir custos e potencializar o contato entre cidadãos e governo. O *Big Data* pode ser a resposta para que governos consigam manipular os grandes conjuntos de informações geradas a partir da digitalização da vida social. O objetivo é avaliar aplicações de *Smart Cities* sob a ótica da incorporação do *Big Data* na gestão de serviços e na tomada de decisões. A metodologia da pesquisa utilizada é o estudo de caso das aplicações *Smart City* e foi utilizada a abordagem quantitativa, envolvendo a análise sobre os dados de aplicações *Smart City*. Os resultados auferidos destacam a importância dos sistemas *Big Data* para aplicações *Smart Cities*. A conclusão reitera que os casos de *Smart Cities* são pontuais e a análise de *Big Data* ainda não faz parte do desenho de políticas das cidades aparecendo também como forma de melhorar planos já estabelecidos em cidades.

Palavras-chave: *Smart City*, *Big Data*, Cidade Digital Estratégica, Gestão de Cidades, Tecnologia da Informação em Cidades.

Big data applied to strategic digital city: study on the volume of data in the smart city applications

Abstract

In the context of Strategic Digital City, the Smart Cities are cities that use information technologies to facilitate the performance of public services, reduce costs and enhance contact between citizens and Government. The Big Data can be the answer to which Governments can manipulate large sets of information generated from the digitalization of social life. The goal is to evaluate applications of Smart Cities from the perspective of the Big Data on service management and decision-making. The research methodology used is the study of Smart City applications and quantitative approach was used, involving the analysis of the data of applications. The results earned the importance of Big Data systems for Smart Cities. The conclusion reiterates that cases of Smart Cities are punctual and Big Data analysis is still not part of the design of policies of cities appearing as well as how to improve plans already established in cities.

Key-words: Smart City, Big Data, Strategic Digital City, Management of Cities, Information technology in Cities.

Big data aplicados la estratégica ciudad digital: estudio sobre el volumen de datos en las aplicaciones de la smart city

Resumen

En el contexto de la Estratégica Ciudad Digital, las ciudades inteligentes, que son ciudades que utilizan tecnologías de la información para facilitar el funcionamiento de los servicios públicos, reducir costos y mejorar el contacto entre ciudadanos y gobierno. El Big Data puede ser la respuesta para que los gobiernos pueden manipular grandes conjuntos de datos generados a partir de la digitalización de la vida social. El objetivo es evaluar aplicaciones de Smart Cities desde la perspectiva de los Big Data en la toma de decisiones y gestión de servicios. La metodología de investigación utilizada es el estudio de caso de aplicaciones de Smart City y se utilizó el enfoque cuantitativo, que implica el análisis de los datos de las aplicaciones de la Smart City. Los resultados obtuvieron la importancia de los sistemas de Big Data de ciudades inteligentes. La conclusión reitera que los casos de ciudades inteligentes son puntuales y análisis de Big Data todavía no es parte del diseño de las políticas de las ciudades que aparecen así como para mejorar los planes ya establecidos en las ciudades.

Palabras clave: Smart City, Big Data, Estratégica Ciudad Digital, Gestión de las ciudades, Tecnología de la información en las ciudades.

1 Introdução

As *Smart Cities*, em português cidades inteligentes, são cidades que utilizam TICs (Tecnologias de Informação e Comunicação) para melhorar o desempenho dos serviços públicos, reduzir custos e potencializar o contato entre cidadãos e governo. Elas estão se tornando uma realidade no atual contexto tecnológico e são uma necessidade para o gerenciamento das cidades e, também, para a definição das políticas públicas. Em Batty (2017) tem-se uma visão geral sobre as *Smart Cities* e sua evolução ao longo dos anos.

A *Smart City* ganhou mais atenção nos últimos anos por conta do aumento da urbanização – em 2014, o equivalente a 54% da população mundial vivia em cidades, com tendência a crescer 1,84% ao ano até 2020 – o que cria demanda por serviços eficientes (Cunha, Przybilovics, Macaya & Burgos, 2016; Hekima, 2016). Como destacado por Rennó (2014), “*um dos grandes desafios das Smart Cities é conectar propostas macro com o contexto urbano local, que está em constante movimento e não opera sob parâmetros generalizantes*”.

O *Big Data* é a resposta para que governos consigam entender, classificar e utilizar positivamente os grandes conjuntos de informações geradas a partir da digitalização da vida

social. *Big Data* é o termo que descreve o grande volume de dados – estruturados e não estruturados – que impactam nas tomadas de decisões (Alecrim, 2015; Dunne, 2012; IBM, 2017; Khan *et al.* 2014).

Em Kon & Santana (2016) e Kudva & Ye (2017) tem-se uma discussão de como a utilização do *Big Data* pode impactar positivamente nas *Smart Cities* tornando-as ainda mais inteligentes e sustentáveis. Algumas soluções de *Smart Cities* utilizando *Big Data* já estão ocorrendo, como descrito em Deloitte (2015) e Nuaimi, Neyadi, Mohamed & Al-Jaroodi (2015).

Nesse sentido uma das vias para se ter uma *Smart City* é se estabelecer uma Cidade Digital Estratégica, como descrito em Rezende (2012), pois as informações são essenciais para se gerir os recursos de uma cidade, e as TICs formam os fatores tecnológicos necessários para se alcançar esse objetivo. Alguns exemplos de projetos de Cidade Digital Estratégica podem ser vistos em Rezende *et al.* (2014), Rezende (2016a) e Rezende (2016b).

Reiterando os problemas de pesquisa destaca-se que a promessa de uma *Smart City* leva a um aumento exponencial dos dados por várias ordens de grandeza. Conseqüentemente, tais enormes conjunto de dados são o cerne dos serviços prestados por todos. O fenômeno de grande conjunto de dados há muito tempo tem sido caracterizado por volume, velocidade, e uma variedade de tipos de dados que foram criados em crescentes taxas (Khan *et al.* 2014). Ainda, em Cheng, Longo, Bauer & Kovacs (2015), Hashen (2016) e Silva, Khan & Han. (2017) tem-se algumas soluções para *Smart Cities* utilizando como base os sistemas *Big Data*.

O objetivo é analisar as aplicações do *Big Data* na concepção e operacionalização das *Smart Cities* visando realizar recomendações sobre os recursos computacionais necessários para a sua utilização.

A justificativa da pesquisa está no estudo de aplicações de *Smart Cities* sob a ótica da incorporação do *Big Data* na gestão dos serviços e na tomada das decisões necessárias para o estabelecimento de políticas da Cidade Digital Estratégica (Rezende, 2012). Por isso, para que isso aconteça é necessário que se tenha uma quantidade de informações confiáveis e pertinentes sobre a ação que se deseja tomar.

As *Smart Cities* estão se tornando uma realidade e para a sociedade atual cada vez mais integrada e conectada, que exige uma tomada de decisão imediata. Por isso, é preciso que se consiga manipular esse volume e variedade de dados no menor tempo possível, e isso é um dos objetivos do *Big Data* (Chen, Mao & Liu, 2014; Nuaimi *et al.* 2015; Kon & Santana, 2016).

Nesse contexto, essa pesquisa tem grande relevância para o estabelecimento da Cidade Digital Estratégica, pois permitirá a caracterização dos recursos computacionais necessários na especificação de sistemas *Big Data* para aplicações *Smart Cities*, em específico o banco de dados do sistema.

2 Fundamentação teórica

Nessa seção descrevem-se de maneira geral os conceitos e definições que norteiam as *Smart Cities* e, também, as tecnologias utilizadas nos sistemas *Big Data*.

2.1 *Big Data*

O termo *Big Data* começou a ser usado no final da década de 90, indicando a enorme quantidade de dados que estão sendo gerados todos os dias pelos diversos sistemas e equipamentos. De modo geral, pode-se dizer que *Big Data* é, essencialmente, tudo que é capturado ou gravado digitalmente pelas modernas TICs, tais como IoT (*Internet of Things*) (Combaneyre, 2015), redes de sensores, objetos e dispositivos "inteligentes", a *internet* e mídias sociais.

O fenômeno do *Big Data* tem sido caracterizado pelo Volume, Velocidade e Variedade nos tipos de dados que tem sido gerado em taxas cada vez maiores (Deloitte, 2015). Essa definição, utilizada atualmente, foi cunhada por Doug Laney (Laney, 2001), que diz que o crescimento destes dados é "alto" e está associado a três (3) variáveis conhecidas como três (3) V's: a) alto Volume (quantidade crescente de dados), b) de alta Velocidade (fluxo de dados chegando a alta velocidade, por exemplo, em tempo real) e, c) alta Variedade (muitos diferentes tipos de dados como texto, áudio, vídeo, *etc*).

Além disso, alguns pesquisadores adicionaram três (3) novos V's que também podem ser encontrados na literatura: a) Veracidade (como as organizações confiar nos dados no sentido de integridade e confidencialidade), b) Variabilidade (como estrutura de dados pode mudar) e, c) Valor (valor do negócio em virtude dos dados das organizações) (Chen *et al.*, 2014). Desse modo, pode-se sintetizar as características e funcionalidade do *Big Data* em seis (6) V's:

- Volume: se refere ao tamanho dos dados que foram criados a partir de todas as fontes.

- Velocidade: refere-se à velocidade em que dados são gerados, armazenados, analisados e processados. Uma ênfase será colocada recentemente no apoio a análise de grande volume de dados em tempo real.
- Variada: refere-se aos diferentes tipos de dados a serem gerados. É comum que a maioria dos dados sejam não-estruturados e não podem ser facilmente categorizados ou tabulados.
- Veracidade: se refere a exatidão dos dados capturados, e qual o significado dos resultados gerados a partir dos dados coletados para determinados problemas.
- Variabilidade: refere-se como a estrutura e o significado de dados muda constantemente especialmente quando lidando com dados gerados a partir de análise de linguagem natural por exemplo.
- Valor: refere-se a possível vantagem de grande volume de dados pode oferecer um negócio baseado na análise, gestão e coleta de dados.

Na Figura 1 tem-se um quadro que relaciona os seis (6) V's do *Big Data*, extraídos de Kon & Santana (2016) e Chen *et al* (2014), e as suas respectivas grandezas/características.



Figura 1: Os 6 V's de Big Data

As características do *Big Data* demonstram o grande potencial de ganhos e avanços no uso desse conceito. As possibilidades são infinitas; no entanto, deve-se interligar as tecnologias e ferramentas disponíveis. O *Big Data* para alcançar seus objetivos, e propiciar serviços em *Smart City*, precisa das ferramentas certas e os métodos para que as informações possam ser analisadas e classificadas de forma eficaz e eficiente. Levando em conta os recursos disponíveis

e limitações, pode-se obter muitas oportunidades para a melhoria dos serviços e aplicações para *Smart City* usando o *Big Data* (Chen *et al.*, 2014).

O *Big Data* pode ser também processado em Computação em Nuvem (*Cloud Computing*) (Tutorials, 2014; White, 2015) usando o modelo computacional denominado de Plataforma como Serviço (PaaS – *Platform-as-a-Service*) ou Infraestrutura como Serviço (IaaS – *Infrastructure-as-a-Service*) (Nuaimi *et al.*, 2015).

O *Big Data* para aplicações *Smart Cities* exibe uma nova característica: a geodistribuição (Bonomi, Milito & Zhu, 2014). Essa nova dimensão do *Big Data* requer que os dados sejam processados na borda próximo aos sensores, em vez dos centros de dados no tradicional modelo de *Cloud Computing*. É necessário oferecer respostas rápidas para proteger a segurança dos componentes críticos da infraestrutura. Nesse sentido, o *Fog Computing* é um modelo adequado, estendendo a computação em nuvem para a borda da rede de computação (Tang, Chen, Hefferman, Wei, He & Yang, 2015). Algumas das principais plataformas computacionais para *Cloud Computing* são: AWS (*Amazon Web Services*) <aws.amazon.com>, Microsoft Azure <azure.microsoft.com>, Google Cloud <cloud.google.com> e IBM Cloud <www.ibm.com/cloud>, dentre outras.

A Computação Paralela é um dos recursos tecnológicos que se pode utilizar no processamento de dados para as *Smart Cities*, em especial no contexto do *Big Data*, e se refere à simultaneidade das operações, utilizando vários recursos de computação para concluir uma tarefa no menor tempo de processamento. Atualmente, alguns modelos clássicos de computação paralela incluem PVM (*Parallel Virtual Machine*) (PVM, 2011), MPI (*Message Passing Interface*) (MPI, 2017), MapReduce (White, 2015) e Dryad (Isard, Budiu, Yu, Birrell & Fetterly, 2007).

2.2 *Smart City*

Uma *Smart City* é uma área urbanizada, onde diversos setores cooperarem para alcançar resultados sustentáveis por meio da análise de informações contextuais, em tempo real. As *Smart Cities* reduzem o congestionamento do tráfego e desperdício de energia, enquanto alocando recursos escassos com mais eficiência e melhoria da qualidade de vida.

O termo guarda-chuva *Smart City* se refere a aplicação desta abordagem em qualquer uma das seis (6) dimensões: a) Economia (*Smart economy*), b) Mobilidade (*Smart mobility*), c)

Ambiente (*Smart environment*), d) Pessoas (*Smart people*), e) Modo de Vida (*Smart living*) e, f) Governo (*Smart governance*) (Kon & Santana, 2016; Silva *et al.*, 2017)

Uma *Smart City* envolve a integração de três (3) fatores básicos: a) Tecnológicos, b) Institucionais e c) Humanos, como mostrado na Figura 2 (Silveira, 2015). Essa relação entre os fatores é que proporciona que realmente ocorra uma *Smart City*, pois não basta ter tecnologia, ou mesmo decisões governamentais se as pessoas não estiverem envolvidas nesse processo. Nesse sentido, é necessário que haja uma cooperação entre os fatores para ter condições propícias para a implantação de uma *Smart City*.

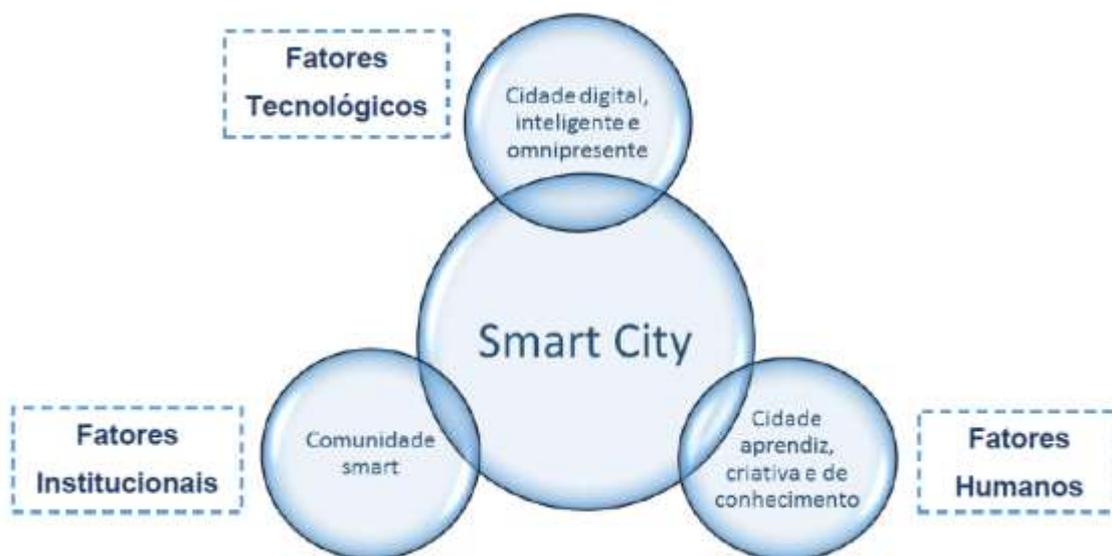


Figura 2: Fatores envolvidos em uma *Smart City*

Fonte: Silveira (2015, p.14)

Na Figura 3 tem-se uma arquitetura típica *Smart City* em que os diversos subsistemas (eletricidade, gerenciamento de água, transporte, estacionamento, resíduos, gás e comunidade) interagem com o sistema de controle e de tomada de decisões.

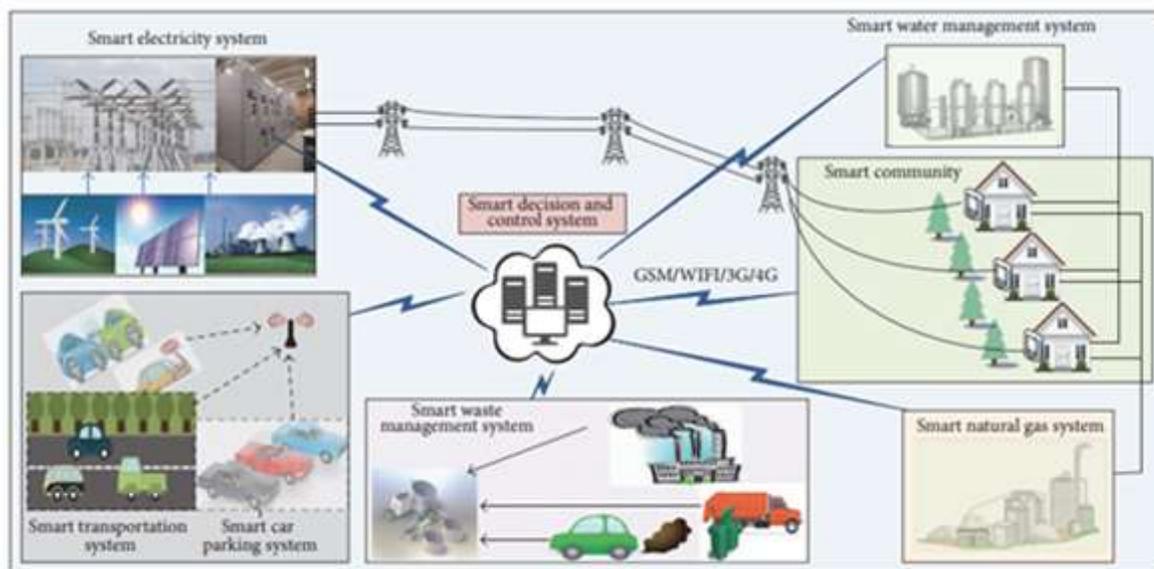


Figura 3: Arquitetura típica de *Smart City*

Fonte: Silva *et al.*, (2017, p.3)

De maneira geral, pode-se definir as Aplicações *Smart City* como sendo a utilização dos três (3) fatores (tecnológicos, institucionais e humanos) em determinado serviço na cidade, dentro do contexto de uma das seis (6) dimensões, compondo os subsistemas da *Smart City*. Diferentemente da cidade digital convencional e da *Smart City*, a Cidade Digital Estratégica pode ser entendida como a aplicação dos recursos da tecnologia da informação na gestão das cidades e, também, na disponibilização de informações e de serviços aos cidadãos. É um projeto mais abrangente do que apenas oferecer *internet* para os cidadãos por meio de recursos convencionais de telecomunicações. Vai além de incluir digitalmente os cidadãos na rede mundial de computadores (Rezende, 2012), pois tem como fundamento as estratégias da cidade para atender os objetivos das diferentes temáticas municipais.

3 Metodologia da pesquisa

Nesse trabalho adotou-se o Estudo de Caso como método de pesquisa para poder analisar os sistemas *Big Data* aplicados às *Smart Cities* e na pesquisa foi adotada a técnica de pesquisa Quantitativa para mensurar os dados coletados (Fonseca, 2002; Gerhardt & Silveira, 2009; Gil, 2007; Padua, 2006; Yin, 2015).

A pesquisa foi composta por duas (2) fases: a Coleta de Dados, que envolve o levantamento dos dados sobre o tema, utilizando para isso o método de pesquisa Estudo de Caso, mais especificamente Multicasos, que no caso é sobre nove (9) Aplicações *Smart Cities*; e a Análise de Dados, que corresponde aos estudos realizados sobre os dados coletados, e que envolveu os dados ligados à utilização de sistemas *Big Data* aplicados às *Smart Cities*.

A amostra da pesquisa contemplou quatro (4) cidades que estão adotando aplicações *Smart Cities* (Surrey (Canadá), Aarhus (Dinamarca), Santander (Espanha) e Chicago – Illinois (Estados Unidos)), e que foram escolhidas devido a abrangência da utilização do conceito *Smart City* e, também, da variedade de aplicações que se deseja analisar para fazer a estimativa da base de dados para a aplicação *Smart City*.

Nesse estudo utilizou-se como unidade de observação as aplicações *Smart Cities*, e envolveu análise sobre os dados dessas aplicações. Inicialmente, buscou-se por bases de dados de diversas aplicações *Smart City* possibilitando uma heterogeneidade, e depois fez-se algumas análises sobre esses volumes de dados. Isto porque o volume de dados (tamanho do conjunto e do registro de dados) depende além da aplicação *Smart City* e, também, da frequência com que ocorre a coleta dos dados pela aplicação, e é medido na unidade computacional *Byte*.

Os dados de aplicações *Smart City* utilizados no estudo estão disponibilizados na internet, nas bases de dados das *Smart City* (Chicago Data Portal <data.cityofchicago.org> e CityPulse Dataset Collection <iot.ee.surrey.ac.uk:8080/datasets.html> e Surrey Datasets <data.surrey.ca/dataset>) coletados em diversos momentos entre 2010 e 2017, e foram sintetizados em 2018.

O conjunto de dados das aplicações *Smart City* analisados nesse estudo inclui:

- (1) dados de 66.523 medidores de consumo de água de casas inteligentes na cidade de Surrey (Canadá) (Surrey, 2014);
- (2) dados relativos à Precipitação Pluviométrica (Chuva) na cidade de Surrey (Canadá) (Surrey, 2017) em seis (6) estações Pluviométricas,
- (3) dados do tráfego foram obtidos do número de veículos em várias estradas na cidade de Aarhus (Dinamarca) (Citypulse, 2014d), verificados em 449 pontos, e este conjunto de dados contém informações muito importantes, por exemplo, o número de veículos e sua velocidade média entre dois pontos em uma estrada;
- (4) conjunto de dados de poluição, coletados em 449 pontos na cidade de Aarhus (Dinamarca), consiste das informações sobre diversos gases tóxicos, tais como

- ozônio, carbono, enxofre e dióxido de nitrogênio e assim por diante (Citypulse, 2014c), ainda contempla também alguns outros materiais perigosos;
- (5) conjunto de dados de estacionamento contêm as informações sobre as 55.264 vagas em vários estacionamentos na cidade de Aarhus (Dinamarca) (Citypulse, 2014b);
 - (6) dados referentes às condições climáticas, compreendem informações sobre temperatura, umidade, pressão atmosférica, direção e velocidade do vento, e foram coletados em 449 pontos na cidade de Aarhus (Dinamarca) (Citypulse, 2014a);
 - (7) conjunto de dados sobre luminosidade, temperatura e vagas de estacionamento na cidade de Santander (Espanha) foi obtido em 1.112 pontos de coleta (Cheng *et al.*, 2015);
 - (8) conjunto de dados corresponde à média anual das permissões de construção que foi obtido na cidade de Chicago-Illinois (Estados Unidos) (Chicago, 2018) desde 2006 totalizando 514.216 registros;
 - (9) conjunto de dados corresponde ao consumo de energia elétrica anual e foi obtido na cidade de Chicago-Illinois (Estados Unidos) (Chicago, 2016) em 2010 totalizando 67.051 registros.

4 Análises e resultados

Nessa seção são apresentados estudos realizados sobre sistemas *Big Data* para aplicações *Smart City*, com base em dados disponíveis na *internet* de nove (9) aplicações *Smart City*, provenientes de quatro (4) cidades: a) Surrey (Canadá), b) Aarhus (Dinamarca), c) Santander (Espanha) e, d) Chicago-Illinois (Estados Unidos)). Os estudos analisam as aplicações sob três (3) aspectos: 1) tamanho do banco de dados, 2) tamanho do registro de dados, 3) crescimento do conjunto de dados dos sistemas e, 4) Projeção do conjunto de dados de aplicações *Smart City* para uma cidade com 500.000 habitantes.

4.1 Análise do estudo #1 - banco de dados de aplicações *Smart City*

Utilizando-se dos dados das nove (9) aplicações *Smart* disponíveis na *internet City*, descritas na seção anterior, elaborou-se estimativas de volume de dados (tamanho) que essas aplicações exigiriam ao longo de um ano (12 meses), e permite estimar a base de dados para um sistema *Big Data*.

Em virtude dos conjuntos de dados terem períodos de coleta diferentes, entre 3 a 12 meses, foi primeiramente realizado a normalização e em alguns casos foi necessário realizar a projeção utilizando-se uma tendência linear do conjunto de dados para um período de tempo de 12 meses.

A informação correspondente para cada conjunto de dados das nove (9) aplicações *Smart City* é mostrada na Tabela 1, e os valores apresentados como tamanho do banco de dados são projeções, utilizando uma tendência linear, considerando a coleta de dados por um período de um ano (12 meses).

Tabela 1

Exemplos do tamanho do banco de dados de aplicações *Smart City*

Aplicação <i>Smart City</i>	Tamanho do Banco de Dados (Bytes)*
(1) Consumo de Água - Surrey (Canadá)	10,0 M
(2) Precipitações Pluviométricas (Chuva) - Surrey (Canadá)	3,9 M
(3) Tráfego de Veículos - Aarhus (Dinamarca)	4,0 M
(4) Poluição - Aarhus (Dinamarca)	3,9 M
(5) Vagas de Estacionamento - Aarhus (Dinamarca)	7,1 M
(6) Condição Climática - Aarhus (Dinamarca)	2,8 M
(7) Luminosidade, Temperatura, Vagas de Estacionamento - Santander (Espanha)	81,8 G
(8) Permissões de Construções - Chicago-Illinois (Estados Unidos)	303,1 M
(9) Consumo de Energia Elétrica - Chicago-Illinois (Estados Unidos)	23,8 M

*Estimativas considerando a coleta de dados por um ano

Obs: M (Mega) e G (Giga)

No estudo percebe-se que o tamanho do banco de dados de uma aplicação que utiliza de poucos dados, como no caso da aplicação Condição Climática na cidade de Aarhus (Dinamarca), terá um tamanho do banco de dados pequeno (2,8 MBytes); ao passo que uma aplicação que precisa armazenar uma maior quantidade de dados, como no caso da aplicação Luminosidade, Temperatura e Vagas de Estacionamento na cidade de Santander (Espanha), demandará um banco de dados maior (81,8 GBytes).

Dessa forma, pode-se destacar que o tamanho do banco de dados é decorrente das características da aplicação *Smart City*. Essa dependência é em função da quantidade de dados

que precisam ser coletados, do tempo de armazenamento e, também, da frequência com que essa coleta ocorrerá pela aplicação.

4.2 Análise do estudo #2 – registro de dados de aplicações *Smart City*

Nesse outro estudo foi considerado o tamanho do registro de dados das nove (9) aplicações *Smart City*, e as informações são mostrados na Tabela 2. Nesse estudo foram comparados os tamanhos do registro de dados das aplicações *Smart City*, apresentadas no estudo anterior, provenientes de quatro (4) cidades: a) Surrey (Canadá), b) Aarhus (Dinamarca), c) Santander (Espanha) e, d) Chicago-Illinois (Estados Unidos)).

Tabela 2

Tamanho do registro de dados de aplicações *Smart City*

Aplicação <i>Smart City</i>	Tamanho do Registro de Dados (Bytes)
(1) Consumo de Água - Surrey (Canadá)	157
(2) Precipitações Pluviométricas (Chuva) - Surrey (Canadá)	39
(3) Trafego de Veículos - Aarhus (Dinamarca)	54
(4) Poluição - Aarhus (Dinamarca)	58
(5) Vagas de Estacionamento - Aarhus (Dinamarca)	68
(6) Condição Climática - Aarhus (Dinamarca)	196
(7) Luminosidade, Temperatura, Vagas de Estacionamento - Santander (Espanha)	536
(8) Permissões de Construções - Chicago-Illinois (Estados Unidos)	618
(9) Consumo de Energia Elétrica - Chicago-Illinois (Estados Unidos)	372

Os dados indicam que o tamanho do registro de dados de uma aplicação que necessita armazenar uma menor quantidade de informações, como no caso da aplicação Precipitações Pluviométricas (Chuva) na cidade de Surrey (Canadá), implicará também em uma menor quantidade (39 *bytes*), pois armazena apenas dados numéricos; ao passo que uma aplicação que precisa armazenar uma maior quantidade de informações, como no caso da aplicação Permissões de Construções na cidade de Chicago-Illinois (Estados Unidos), necessitará uma maior quantidade (618 *bytes*) para comportar as informações, pois armazena diversos dados tanto numéricos quanto alfanuméricos (texto).

No estudo constatou-se que o tamanho do registro de dados depende diretamente da quantidade de dados (campos) e do tipo dos dados armazenados (numérico, texto, lógico, *etc*),

pois quanto maior for a quantidade de dados a serem armazenados, maior será o tamanho do registro das informações coletadas pelas aplicação *Smart City*.

Nesse sentido, destaca-se que é importante especificar claramente o registro de dados da aplicação *Smart City*, tanto a quantidade de campos quanto o tipo de dados, pois o tamanho do registro de dados impactará no tamanho do banco de dados do sistema *Big Data*.

4.3 Análise do estudo #3 - crescimento do conjunto de dados de uma aplicação *Smart City*

Nesse estudo considerou-se a aplicação Consumo de Energia Elétrica em uma *Smart City*, obtido por 1 milhão de medidores inteligentes em intervalo de tempo de um ano, e é mostrado na Tabela 3. *Obs.*: A quantidade de dados coletados foi calculada considerando o tamanho do registro de dados de 1 *KBytes*.

Foram comparados os volumes de dados gerados pela aplicação *Smart City* considerando quatro (4) frequências de coleta: 1) uma por dia, 2) uma por hora, 3) uma a cada 30 minutos e, 4) uma a cada 15 minutos.

Com base nesse estudo percebe-se claramente o crescimento exponencial da geração de dados, pois se for utilizado uma frequência de uma medição por dia tem-se ao final de um ano um conjunto de dados de ~0,37*TBytes*, ao passo que se for utilizado uma medição a cada 15 minutos serão armazenados um conjunto de dados de 35,04*TBytes*.

Tabela 3

Total de dados coletados em um ano (12 meses) em diferentes frequências de coleta

Frequência de Coleta	1/dia	1/hora	1/30 min	1/15 min
Registros Coletados	365M	8,75B	17,52B	35,04B
Tamanho da Coleta	~0,37 <i>TBytes</i>	~8,76 <i>TBytes</i>	~17,52 <i>TBytes</i>	~35,04 <i>TBytes</i>

M (milhões), B (bilhões)

Isto demonstra a importância de se definir apropriadamente além dos dados que serão coletados (registro de dados), também a frequência de coleta, pois isto tem impacto direto na especificação do tamanho do banco de dados do sistema *Big Data* para a aplicação *Smart City*. De modo que, quanto maior for a frequência de coleta, maior será o tamanho do banco de dados.

4.4 Análise do estudo #4 – projeção do conjunto de dados de aplicações *Smart City* para uma cidade com 500.000 habitantes

Com base nos estudos anteriores, em específico o tamanho do registro de dados das aplicações, estabeleceu-se uma projeção do tamanho do conjunto de dados de algumas aplicações *Smart City* para uma cidade com 500.000 habitantes. Nesse estudo analisou-se oito (8) aplicações *Smart City* (Consumo de Água Residencial, Precipitação Pluviométrica (chuva), Tráfego de Veículos, Poluição, Vagas de Estacionamento, Condição Climática, Construções e Consumo de Energia Elétrica Residencial). Nessa projeção foi considerado quatro (4) pessoas por residência para a definição dos pontos de coleta do consumo de água e de energia elétrica residencial, e de uma vaga de estacionamento para cada seis (6) habitantes para a definição dos pontos de coleta de vagas de estacionamento. Além disso, estipulou-se uma frequência de coleta em função das características da aplicação como mostrado na Tabela 4.

Tabela 4

Projeção do conjunto de dados de aplicações *Smart City* para uma cidade com 500.000 habitantes

Aplicação <i>Smart City</i> (cidade de 500.000 habitantes)	Frequência de Coleta	Pontos de Coleta	Conjunto de Dados (Bytes)
Consumo de Água - Residencial	1/mês	125.000	224,6 M
Precipitações Pluviométricas (Chuva)	1/dia	10	139,0 K
Trafego de Veículos	1/30 min	500	451,1 M
Poluição	1/dia	500	10,1 M
Vagas de Estacionamento	1/hora	80.000	44,4 G
Condição Climática	1/hora	500	818,7 M
Permissões de Construções	1/dia	125.000	26,3 G
Consumo de Energia Elétrica - Residencial	1/15 min	125.000	1,5 T

Obs: K (Kilo), M (Mega), G (Giga), T (Tera)

A projeção do tamanho do conjunto de dados (em bytes) considerando o registro de dados, a frequência e os pontos de coleta está mostrada Tabela 3, e se percebe o impacto que esses elementos tem sobre o conjunto de dados. Por exemplo, a aplicação Consumo de Energia Elétrica devido a frequência de coleta (uma medição a cada 15 minutos) e, também, pela quantidade de pontos de coleta (125.000 residências) exigirá 1,5TBytes de armazenamento ao longo de 12 meses. No outro extremo tem-se a aplicação Precipitações Pluviométricas (Chuva)

que devido a frequência de coleta (uma por dia) e de ser previstos apenas 10 pontos de coleta necessitará de pouco espaço de armazenamento (139 KBytes).

4.5 Relação dos estudos

Os estudos realizados demonstram o impacto que a aplicação *Smart City* tem sobre o banco de dados de um sistema *Big Data*. Esses estudos estão inter-relacionados, pois analisam as informações sobre o conjunto de dados de aplicações *Smart City* diversas óticas.

O primeiro estudo analisou a questão do tamanho do conjunto de dados de nove (9) aplicações e está relacionado com a quantidade de informações a serem armazenadas e, bem como, como a frequência com que essa coleta é realizada, além do tempo de armazenamento pela aplicação *Smart City*.

O segundo estudo analisou o tamanho do registro de dados, que é a unidade básica de informação do banco de dados do sistema *Big Data*, de nove (9) aplicações *Smart City*, e é determinado em função da quantidade de dados que precisam ser armazenados a cada coleta pela aplicação.

No terceiro estudo analisou-se o conjunto de dados de uma aplicação *Smart City* sob a ótica do crescimento do banco de dados em função da frequência com que ocorre essa coleta dos dados pela aplicação *Smart City*.

Por último, o quarto estudo projetou o conjunto de dados de oito (8) aplicações *Smart City* considerando uma cidade com 500.000 habitantes, e que demonstra a influência do tamanho do registro de dados, a frequência e os pontos de dados sobre o espaço de armazenamento.

5 Conclusão

O tema *Smart City* está se tornando muito difundido, não somente no meio acadêmico ou de pesquisa, mas também nos órgãos de governos e empresas como uma solução para os problemas crescentes das cidades. Esses problemas precisam de decisões acertadas e que ocorram no menor período de tempo para a que a definição de políticas públicas seja eficaz. É nesse contexto que se apresenta o *Big Data*, como um recurso tecnológico a ser utilizado na composição das *Smart Cities* para se trabalhar com um volume de dados cada vez maior e, também, considerando a diversidade de dados que são gerados nas cidades.

O objetivo deste trabalho foi alcançado, pois foram realizados estudos da utilização do *Big Data* na concepção e operacionalização da Cidade Digital Estratégica, que possibilitou realizar apontamentos sobre os recursos computacionais, em específico o banco de dados, necessários para a sua utilização nas aplicações *Smart City*.

A pesquisa contemplou o estudo de nove (9) aplicações de *Smart Cities* de quatro (4) cidades (Surrey (Canadá), Aarhus (Dinamarca), Santander (Espanha) e Chicago-Illinois (Estados Unidos)) com base nos dados dessas aplicações e que pudessem servir como elemento norteador para as definições de sistemas *Big Data* para aplicações *Smart City*.

Os resultados auferidos indicam que a aplicação *Smart City* impacta diretamente no banco de dados do sistema *Big Data*, pois ela é decorrente do tamanho do registro de dados e, também, da quantidade de registros armazenados pela aplicação *Smart City*. A primeira definição a ser realizada é a do registro de dados, em que se especificam os dados a serem coletados e armazenados, como os dados e os tipos de dados (numérico, texto, etc). A segunda definição depende da frequência com que a coleta dos dados ocorrerá, pois isso impactará na quantidade de registros de dados que precisam ser armazenados no sistema *Big Data*.

Ainda, destaca-se que o período de armazenamento também afeta o banco de dados, pois quanto maior for o período de coleta e, conseqüente, o armazenamento, maior será o tamanho do banco de dados. Por isso, uma aplicação que precisa trabalhar com um histórico muito grande de dados, naturalmente demandará um maior banco de dados.

Salienta-se também que o tamanho do conjunto de dados tem impacto direto no tempo de processamento da aplicação, pois quanto maior for o conjunto de dados, maior será o tempo gasto para consultar e processar os dados e, conseqüentemente, no tempo de resposta do sistema. Ainda, destaca-se que o tempo de resposta do sistema para *Smart City* depende das características da aplicação e, principalmente, do tempo para a tomada de decisão, por isso quanto menor for esse tempo mais rápido ocorrerão as ações dos órgãos competentes.

Outro ponto a ser realçado é que o banco de dados do *Big Data* afeta a especificação do *Data Center*, como a quantidade de computadores/servidores e dos demais equipamentos (*switchs*, roteadores, unidades de armazenamento, etc) a serem utilizados e, por conseqüência, no consumo de energia elétrica e na refrigeração/dissipação de calor dos equipamentos.

No que se refere à contribuição desse trabalho tem-se a ampliação da pesquisa voltada à Cidade Digital Estratégica, em particular à *Smart City*, além de fornecer subsídios para definição dos serviços e recursos que serão gerenciados na cidade.

A principal limitação desse trabalho está relacionada com a metodologia da pesquisa que reduziu as análises e considerações a nove (9) aplicações *Smart Cities* de quatro (4) cidades, não expressando a realidade das cidades que estão utilizando soluções *Smart City*.

A conclusão reitera a importância dos sistemas *Big Data*, na composição das soluções *Smart City*, como elemento facilitador da gestão das cidades, visando uma gestão pública que propicie um melhor aproveitamento dos recursos públicos, com sustentabilidade, na ampliação da qualidade de vida dos cidadãos.

Referências

- Alecrim, E. (2015). *O que é big data?*. Disponível em <http://www.infowester.com/big-data.php>.
- Batty, M. (2017). *The age of the smart city*. Disponível em <http://www.spatialcomplexity.info/files/2017/06/BATTY-Working-Paper-The-Age-of-the-Smart-City.pdf>.
- Bonomi, F., Milito, R., & Zhu, J. (2014). *Fog computing: a platform for internet of things and analytics*. Big Data and Internet of Things: A Roadmap for Smart Environments, 169-186.
- Chen, M., Mao, S., & Liu, Y. (2014). Big data: a survey. *Mobile Netw Appl*. Springer: Switzerland, 171-209.
- Cheng, B., Longo, S., Bauer, M., & Kovacs, E. (2015). Building a big data platform for smart cities: experience and lessons from Santander. *Anais do 2015 IEEE International Congress on Big Data*. New York-USA, 592-599.
- Chicago Data Portal (2018). *Building permits*. Disponível em <https://data.cityofchicago.org/Buildings/Building-Permits/ydr8-5enu>.
- Chicago Data Portal (2016). *Energy usage 2010*. Disponível em <https://data.cityofchicago.org/Environment-Sustainable-Development/Energy-Usage-2010/8yq3-m6wp>.
- CityPulse Dataset Collection (2014a). *Weather data*. Disponível em <http://iot.ee.surrey.ac.uk:8080/datasets.html#weather>.
- CityPulse Dataset Collection (2014b). *Parking data*. Disponível em <http://iot.ee.surrey.ac.uk:8080/datasets.html#parking>.
- CityPulse Dataset Collection (2014c). *Pollution data*. Disponível em <http://iot.ee.surrey.ac.uk:8080/datasets.html#pollution>.
- CityPulse Dataset Collection (2014d). *Road traffic data*. Disponível em <http://iot.ee.surrey.ac.uk:8080/datasets.html#traffic>.

- Cunha, M.A., Przeybilovics, E., Macaya, J.F.M., & Burgos, F. (2016). *Smart cities: transformação digital de cidades*. (1. ed.). São Paulo-SP: PGPC.
- Combaneyre, F. (2015). *Understanding data streams in IoT*. White Paper. SAS.
- Deloitte (2015). *Smart cities, big data*. UK: Deloitte & Touche.
- Dunne, T. (2012). Big data, analytics, and energy consumption, *Lavastorm Agle Analytics*, Boston, Mass, USA.
- Fonseca, J.J.S. (2002). *Metodologia da pesquisa científica*. Apostila. Fortaleza-CE: UEC.
- Gerhardt, T.E., Silveira, D.T. (Org.). (2009). *Métodos de pesquisa*. Série Educação a Distância. (1. ed.). Porto Alegre-RS: Editora da UFRGS.
- Gil, A.C. (2007). *Como elaborar projetos de pesquisa*. (4. ed.). São Paulo-SP: Atlas.
- Hashen, I.A.T. (2016). The role of big data in smart city. *International Journal of Information Management*, 5(36).
- Hekima. (2016). *Como smart cities usam big data*. Disponível em <http://www.bigdatabusiness.com.br/como-smart-cities-usam-big-data/>.
- IBM (2017). *Big data*. Disponível em <http://www.ibm.com/big-data/us/en/>.
- Isard, M., Budiu, M., Yu, Y., Birrell, A., & Fetterly, D. (2007). *Dryad: distributed data-parallel programs from sequential building blocks*. *Anais do Proceedings of the 2007 Eurosys Conference*.
- Khan, N., Yaqoob, I., Hashem, I.A.T., Inayat, Z., Ali, W.K.M., Alam, M., Shiraz, M., & Gani, A. (2014). Big data: survey, technologies, opportunities, and challenges. *The Scientific World Journal*, 18. doi: 10.1155/2014/712826.
- Kon, F., & Santana, E.F.Z. (2016). *Cidades inteligentes: conceitos, plataformas e desafios*. JAI 2016 – Jornadas de Atualização em Informática 2016.
- Kudva, S., & Ye, X. (2017). Smart cities, big data, and sustainability union. *Big Data and Cognitive Computing*, 1(4), 2017. doi: 10.3390/bdcc1010004.
- Laney, D. (2001). *3D data management: controlling data volume, velocity, and variety*. Application Delivery Strategies. Meta Group. Disponível em <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>.
- MPI (2017). *Message passing interface standard (MPI)*. Disponível em <https://computing.llnl.gov/tutorials/mpi/>.
- Nuaimi, E.Al., Neyadi, H.A., Mohamed, N., & Al-Jaroodi, J. (2015). Applications of big data to smart cities. *Journal of Internet Services and Applications*.

- Pádua, E.M.M. (2006). *Metodologia da pesquisa: abordagem teórico-prática*. (12. ed.). Fortaleza-CE: Papirus Editora.
- PVM (2011). *PVM - parallel virtual machine*. Disponível em <https://www.csm.ornl.gov/pvm/>.
- Rennó, R. (2014). Smart cities e big data: o cidadão produtor de dados. *Revista de Estudios Urbanos y Ciencias Sociales*, 6(2), 13-24.
- Rezende, D.A. (2012). *Planejamento de estratégias e informações municipais para cidade digital: guia para projetos em prefeituras e organizações públicas*. São Paulo-SP: Atlas.
- Rezende, D.A. (2016a). Digital city projects: information and public services offered by Chicago (USA) and Curitiba (Brazil). *International Journal of Knowledge Society Research*, 7, 16-30, 2016a.
- Rezende, D.A. (2016b). Digital city: Chicago and Schaumburg (USA) information, sustainability indicators and public services projects. *Revista de Gestão e Secretariado*, 7, 55-79.
- Rezende, D.A., Madeira, G.S., Mendes, L.S., Breda, G.D., Zarpelão, B.B. & Figueiredo, F.C. (2014). Information and telecommunications project for a digital city: a Brazilian case study. *Telematics and Informatics*, 31, 98-114.
- Silva, B.N., Khan, M., & Han, K. (2017). Big data analytics embedded smart city architecture for performance enhancement through real-time data processing and decision-making. *Hindawi Wireless Communications and Mobile Computing*. 2017, article ID 9429676.
- Silveira, A.T.Z.N. (2015). *Importância dos sistemas inteligentes de água nas cidades para a eficiência do serviço*. Técnico Lisboa. Dissertação de Mestrado. Lisboa-Portugal.
- Surrey Datasets (2017). *Rainfall monitoring 2017*. Disponível em <http://data.surrey.ca/dataset/rainfall-monitoring-2017>.
- Surrey Datasets (2017). *Water meters*. Disponível em <http://data.surrey.ca/dataset/water-meters.2014>.
- Tang, B., Chen, Z., Hefferman, G., Wei, T., He, H. & Yang Q. (2015). A hierarchical distributed fog computing architecture for big data analysis in smart cities. *Anais do 2015 ACM BD&SI*. Kaohsiung-Taiwan.
- Tutorials Point (2014). *Hadoop: big data analysis framework*. Ebook.
- White, T. (2015). *Hadoop: the definitive guide*. (4. ed.). Sebastopol-CA-USA: O'Reilly Media.
- Yin, R.K. (2015). *Estudo de caso: planejamento e métodos*. (5. ed.). Porto Alegre-RS: Bookman.