

Explorando Dados Abertos Governamentais Sobre a Mobilidade Urbana na Cidade do Rio de Janeiro

Sérgio Manuel Serra da Cruz^{1,3}, Luan Soares Andrade^{1,2},
Jonice Oliveira Sampaio²,

¹Universidade Federal Rural do Rio de Janeiro - LabData/UFRRJ

²Universidade Federal do Rio de Janeiro - LabCores/UFRRJ

³Programa Educação Tutorial - PET-SI/UFRRJ

serra@pet-si.ufrrj.br, oluan@ufrj.br, jonice@dcc.ufrj.br,

Resumo *Um dos grandes desafios das cidades brasileiras é a busca de alternativas para melhorar a mobilidade urbana. O conceito de cidades inteligentes tem despontado como a combinação do ambiente digital com as comunidades reais com vistas a tornar mais eficiente a gestão pública dos espaços urbanos. Soluções baseadas em dispositivos móveis que utilizem dados abertos governamentais são consideradas importantes diferenciais. Este trabalho propõe uma arquitetura distribuída intitulada BusInRio e experimentos que utilizam dados abertos governamentais que permitem aos usuários de transporte público da cidade do Rio de Janeiro acompanharem rotas e a localização dinâmica dos ônibus de uma linha e o estado do trânsito.*

Abstract *One of the great challenges of the Brazilian cities is the search of best opportunities to improve the urban mobility. The concept of smart cities has emerged as a combination of the digital environment with real communities aiming to making the management public of urban space more efficient. This paper presents a mobile distributed architecture named BusInRio, its prototype and experiments. BusInRio uses open government data to aid user of public transport of the city of Rio de Janeiro monitor and evaluate routes and dynamically recognize the location of the bus lines and the traffic status.*

1. Introdução

A gestão pública vem se modernizando e recebendo diversas demandas. Um dos principais desafios para o avanço da democracia é o desenvolvimento contínuo das ferramentas que aprofundem a interação entre o governo e a sociedade. Neste sentido, ao lado das iniciativas de *e-Governo*, surgiu o conceito de Dados Abertos (Portal Dados Abertos, 2013) numa injunção para assegurar a disponibilidade de dados governamentais para a sociedade.

O conceito se refere aos dados que podem ser usados, modificados e compartilhados livremente por qualquer pessoa e por qualquer propósito, estando sujeito apenas, e no máximo, aos requisitos que preservem sua proveniência e compartilhamento pela mesma licença (Portal Dados Abertos, 2013, Open Data Definition, 2016). Diversos eventos recentes contribuíram para a disseminação do tema na sociedade brasileira.

Citamos a Lei de Acesso à Informação¹, a criação da Infraestrutura Nacional de Dados Abertos², além de diversas iniciativas internacionais tais como o Global Open Data Index (index.okfn.org) e Open Data Institute (theodi.org) que incentivam o uso das TIC (Tecnologia da Informação e Comunicação) e dados abertos como um esforço para tornar os governos mais transparentes, efetivos e confiáveis, com instituições que empoderam os cidadãos e escutam suas necessidades. Atualmente diversos países mantêm portais com diversos tipos de dados abertos governamentais, dentre eles se destacam: data.gov.uk, open.canada.ca, data.gov e data.gov.br.

O conceito de Cidades Inteligentes (Dohler et al. 2011, Schaffers et al, 2011) tem despontado como a combinação do ambiente digital com comunidades reais, tornando mais eficiente a gestão pública dos espaços urbanos, o conceito é bastante amplo e possui diferentes conotações que podem variar segundo as perspectivas do usuário, da tecnologia ou da economia. Apesar de ser um conceito relativamente recente, ele já se consolidou como tema fundamental na discussão global sobre o desenvolvimento sustentável, estima-se que movimentará um mercado global de soluções tecnológicas estimado em US\$ 400 bilhões até 2020.

Atualmente, cidades de países emergentes estão investindo bilhões de dólares em produtos e serviços baseados na economia do conhecimento para sustentar o crescimento econômico e as demandas materiais da nova classe média. No entanto, um dos principais problemas enfrentados pelos gestores e que afeta grande parte população é a precariedade da mobilidade urbana. Problemas tais como engarrafamentos, falta de investimentos em infraestrutura e qualidade precária dos serviços de transporte afetam diretamente a economia e tornam a vida do cidadão mais estressante. Com vistas a explorar as oportunidades de negócio advindas deste conceito, diversas empresas ligadas à economia do conhecimento (Uber, Cabify, Waze, entre outros) oferecem serviços centrados em plataformas digitais distribuídas baseados em dados proprietários que visam oferecer alternativas aos desafios da mobilidade urbana.

A cidade Rio de Janeiro, assim como outras grandes cidades do mundo, já apresenta iniciativas interessantes sobre os conceitos de Cidades Inteligentes e Dados Abertos, por exemplo, ele anualmente promove o concurso de criatividade RioApps³ que estimula o desenvolvimento de *apps* que contribuirão para tornar a cidade inteligente. Além disso, ele mantém o Centro de Operações (centrodeoperacoes.rio) que monitora a cidade e realiza o gerenciamento de crises (chuvas, deslizamentos, acidentes). Ela também mantém repositórios de dados abertos governamentais através do portal (data.rio) sobre os mais diversos assuntos tais como: educação, saúde, impostos administração pública, meteorologia e transportes em formatos abertos para que empreendedores e desenvolvedores criem suas aplicações de interesse público. Cada dataset possui sua estrutura sintática própria e são ofertados em vários formatos abertos, a saber: XML, JSON, CSV, entre outros.

Este trabalho tem como objetivo contribuir com a oferta de soluções tecnológicas relacionadas aos desafios da mobilidade urbana nas grandes cidades brasileiras.

¹http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm

²<http://www.governoeletronico.gov.br/acoes-e-projetos/Dados-Abertos/inda-infraestrutura-nacional-de-dados-abertos>

³<http://portalrioapps.com.br/>

Apresentamos uma arquitetura distribuída denominada *BusInRio*, seu protótipo e experimentos baseados em nuvem de computadores, que utiliza dados abertos governamentais sobre as viagens dos ônibus que são transformados através de regras Extração, Transformação e Carga (ETL) em bases NoSQL interligadas que oferecem informações acerca das linhas de ônibus para que o usuário do serviço possa escolher o modal que irá utilizar em função não só da localização na cidade do Rio de Janeiro como também sobre o estado do trânsito por onde a linha circulará.

O trabalho está organizado da seguinte forma: a seção 2 apresenta os trabalhos relacionados. A seção 3 discute os fundamentos teóricos, já a seção 4 apresenta os materiais e métodos adotados. A seção 5 apresenta a arquitetura *BusInRio* e a seção 6 discute os principais resultados. Por fim, a última seção apresenta a conclusão, limitações e discute os trabalhos futuros.

2. Trabalhos Relacionados

Diversas cidades do mundo oferecem *apps* que utilizam dados abertos sobre mobilidade urbana. As cidades Nova Iorque, Chicago e Londres publicam uma variedade de dados através das suas autoridades de trânsito sobre as rotas dos transportes públicos para que desenvolvedores, pesquisadores, planejadores urbanos, jornalistas de dados e *startups* os utilizem.

Por exemplo, na cidade de Londres o *app CityMapper* também está integrado com o metrô, em Nova Iorque existe o *app OneBusAway* que apresenta dados de transporte público integrando dados abertos sobre os modais trem, metrô e ônibus. Algumas cidades do Brasil, dentre elas o Rio de Janeiro, começam a oferecer alguns *apps* comerciais que utilizam dados abertos governamentais (RioBus, Cadê o Ônibus?, Buus⁴). O *app* RioBus é um sistema colaborativo de monitoramento de ônibus em tempo real que provê algumas funcionalidades semelhantes às primeiras versões do Buus. Os *apps*, utilizam as APIs de dados abertos de mobilidade urbana disponibilizada pela Prefeitura da cidade do Rio de Janeiro. Devemos ressaltar que esses *apps* não provêm informações técnicas, publicações ou mesmo maiores detalhes sobre sua arquitetura, algoritmos, estrutura de dados nem mesmo sobre o funcionamento pois se tratam de programas protegidas por interesses comerciais.

O sistema UbiBus (Tito et al, 2012) tem como foco a cidade do Recife (PE), ele está baseado em um *middleware* que faz uso de informações contextuais de *tweets* e de computação ubíqua para se adaptar às situações do trânsito e recomendar rotas alternativas dos ônibus. O sistema Busão (Leite et al, 2013) tem como foco a cidade de Campina Grande (PB) e oferece aos usuários os itinerários dos ônibus e as melhores rotas dos ônibus. Ambos não utilizam dados abertos, mas dados do GPS coletados pelos passageiros.

Barbosa et al., (2014) e Bessa et al., (2015) desenvolveram soluções baseadas em dados abertos sobre os transportes do Rio de Janeiro, os autores oferecem recursos de visualização das trajetórias de ônibus do Rio de Janeiro. Este último trabalho tem como principal característica a detecção de *outliers* através de uso de redes neurais. Diferentemente, do *BusInRio*, essas soluções não discutem os aspectos arquiteturais das soluções apresentadas nem relatam possíveis integrações com aplicativos móveis nem

⁴ Desenvolvedor Tormenta Labs; Luis Picanço e Buus Serviços de Tecnologia (respectivamente).

representam os dados dinamicamente. Por fim, esses trabalhos não consideram a proveniência de dados abertos.

3. Fundamentação Teórica

A qualidade e a proveniência dos dados abertos (Mendonça et al., 2014) foram os principais fatores considerados para a estratégia de desenvolvimento do *BusInRio*. Nesta seção descreveremos os principais fundamentos teóricos utilizados neste trabalho.

3.1 Dados Abertos Governamentais

Os dados abertos governamentais dissem respeito a publicação e a disseminação das informações do setor público na Internet, sendo compartilhadas em formato bruto e aberto, compreensíveis logicamente, possibilitando o reuso por aplicações desenvolvidas pela sociedade.

A atividade de abrir dados ainda é realidade para uma pequena parcela dos governos nacionais. Se consideramos os governos municipais essa parcela é ainda menor. Para grande parte dos órgãos públicos, a falta de pessoal capacitado ainda é o principal motivo que contribui com essa realidade. Abrir dados não é uma tarefa trivial. Eles devem ser completos, primários, atuais, não proprietários, compreensíveis por máquina e não proprietários (OpenGovData, 2016). Os principais efeitos dos dados abertos nas políticas públicas são a inclusão social, a transparência governamental e a efetividade das políticas públicas.

3.2 Proveniência em Dados Abertos

A proveniência de dados representa a ancestralidade de um objeto e pode ser descrita em diferentes termos, dependendo do domínio abordado (Gil et al. 2010). Em geral, ela fornece a documentação essencial para registrar a origem, a qualidade e a autoridade dos dados, assim como a capacidade de reproduzir e validar resultados de processos comerciais ou científicos. Os diferentes tipos de proveniência e seu nível de detalhamento (Cruz et al. 2009) são facetas que podem ser exploradas no contexto da qualidade dos dados abertos. Quanto menor a granulosidade, mais refinadas serão as possibilidades de investigação entre os dados publicados e seus descritores de proveniência. Nesse trabalho, adotamos a abordagem proposta por Mendonça et al. (2013) consideramos apenas a proveniência retrospectiva de baixa granulosidade relativa ao processo de ETL dos dados abertos.

3.3 Cidades Inteligentes

Apesar do conceito de Cidade Inteligente ser recente (Hernández-Munõz et al., 2011), ele já se consolidou como temática fundamental nas discussões sobre economia do conhecimento e mobilidade urbana. Ele diz respeito aos sistemas onde pessoas interagem e usam energia, materiais, serviços e financiamento para catalisar o desenvolvimento econômico e a melhoria da qualidade de vida. Esses fluxos de interação são considerados inteligentes por fazer uso estratégico das TICs juntamente com o planejamento e gestão urbana para buscar resposta às necessidades da sociedade (Dohler et al. 2011).

4. Materiais e Métodos

Esta seção apresenta não só a metodologia proposta para o desenvolvimento do *BusInRio*, como também descreve os datasets utilizados nesse trabalho.

4.1. Datasets e suas estruturas

Nesse trabalho utilizamos vários datasets abertos disponibilizados pela Prefeitura (percursos de cada linha de ônibus e dados em tempo real dos GPS de cada ônibus de cada linha). Resumidamente, os volumes de dados diários produzidos são medidos em dezenas de megabytes/dia. Os dados são de natureza textual e são caracterizados por milhares de registros produzidos pelos GPS de mais de 8.000 ônibus e arquivos das 490 linhas de ônibus que atendem o município.

Os dados contendo a localização das paradas (ponto de ônibus) de cada linha de ônibus são oferecidos pela FETRANPOR-RJ em formato CSV (<http://data.rio/dataset/pontos-dos-percursos-de-onibus>) e possuem o seguinte estrutura: Descrição de início e fim daquela linha; agência que disponibiliza os dados; número de sequência em que o ônibus passa em um determinado ponto; coordenadas (latitude e longitude) do ponto de ônibus.

Os dados dos GPS instalados nos ônibus são de alta frequência, os dados de GPS com a localização atual de cada ônibus são produzidos a cada minuto e disponibilizados em tempo real em formato JSON (<http://data.rio/dataset/gps-de-onibus>). Este conjunto de dados cobre parte da frota das concessionárias da Cidade do Rio de Janeiro e algumas linhas intermunicipais. O *dataset* possui a seguinte estrutura: Data e hora da geração do dado; identificação alfanumérica encontrada na lateral do ônibus; identificador da linha do ônibus; coordenadas (latitude e longitude) do ônibus no trajeto; velocidade do ônibus no momento da coleta e medida em graus em relação ao norte, que representa a direção do veículo.

4.2. Metodologia proposta

Os dados de GPS das trajetórias dos ônibus apresentam alguns problemas tais como: ausência de informação da direção, erros de representação das coordenadas, erros de coleta (apesar da frequência ser a cada minuto, algumas medidas são superiores a 10~12 minutos). Neste trabalho, adotamos a abordagem ETL composta por três fases para garantir a qualidade e a integridade dos dados abertos, a saber:

1. *Coleta de dados brutos*: nesta fase os dados brutos são coletados a partir de diversas fontes do portal rio.data, eles são temporariamente armazenados em disco (*data staging area*) antes que qualquer tipo de manipulação;

2. *Limpeza e consolidação de dados*: nesta fase os dados brutos são corrigidos, transformados e consolidados em dados curados de acordo com os requisitos da aplicação e dos princípios do registro de proveniência de dados abertos ligados conforme apresentamos em trabalho anterior (Reiche, Cruz e Campos, 2014). Dentre as tarefas de limpeza realizamos exclusões de diversos tipos de registros que apresentam falhas: duplicados, falhos ou sem coordenadas; coordenadas erradas, trajetórias errôneas com distâncias superiores a 100 metros a partir da rota esperada; trajetórias com intervalos de GPS inferiores a 1 minuto ou superiores a 5 minutos. Como consolidação de dados calculamos a direção rota do ônibus e ligamos os dois datasets.

3. *Cargade dados curados*: nesta fase os dados consolidados são (re)agrupados em documentos ou família de colunas e carregados para os repositórios NoSQL apropriados.

4. *Anotação de proveniência retrospectiva*: nesta fase os metadados de proveniência do tipo retrospectiva de baixa granulosidade (Cruz et al, 2009) são produzidos pelos processos computacionais de limpeza e consolidação de dados abertos, os metadados também são armazenados nos repositórios de dados curados. As anotações de proveniência descrevem como o dado foi coletado e processado antes da publicação e em última análise garantem a transparência dos dados e das suas transformações.

5. Solução Proposta

Este trabalho propõe uma arquitetura, intitulada *BusInRio*, para a publicação de dados abertos curados e anotados com proveniência, baseia-se na descoberta de dados abertos governamentais da área de transporte público. O *BusInRio* é um serviço que não depende de plataformas tecnológica ou *demiddlewares* comumente utilizados em aplicações do tipo Cidades Inteligentes (Hernández-Munõz et al., 2011).

A arquitetura *BusInRio* é do tipo distribuída e escalável. A Figura 1 ilustra sua representação conceitual. Sua porção servidora executa em ambiente do tipo nuvem de computadores do tipo pública e modelo orientado a serviços SaaS (Chee, Franklin Jr, 2013) e a porção cliente executa em *smartphones* com diversos tipos de sistema operacional.

A arquitetura adota o padrão arquitetônico cliente-servidor. Os códigos da aplicação servidora foram desenvolvidos em linguagem PHP, já o cliente é totalmente desenvolvido tecnologias HTML5-JavaScript e os repositórios de dados curados e de descritores de proveniência são do tipo não relacional.

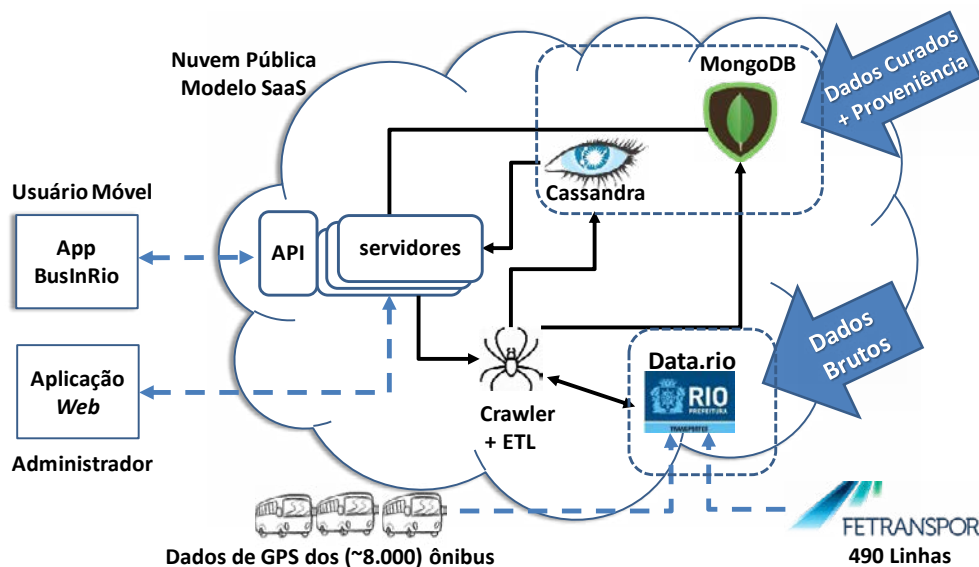


Figura 1. Representação conceitual da arquitetura *BusInRio*.

O servidor de aplicação tem como principais funcionalidades conectar-se ao *data.rio*, utilizar a abordagem ETL para coletar, tratar os dados abertos de transporte público e armazená-los juntamente com sua proveniência retrospectiva em suas

respectivas bases NoSQL (Cassandra e MongoDB) e disponibilizar uma API para que a aplicação móvel consulte e exiba os dados curados e mapas aos usuários.

O cliente é totalmente baseado do framework Ionic⁵, sendo capaz de oferecer uma aplicação híbrida capaz de operar em diversos tipos de *smartphones* ou navegadores web. O cliente possui as seguintes funcionalidades: coletar os dados através da API e exibir para o usuário lista de linhas de ônibus e as representações gráficas através de mapas das APIs do Google Maps que permitem a leitura e visualização de dados sobre o ônibus e sobre trânsito do entorno.

BusInRio é uma arquitetura que explora o princípio da elasticidade das nuvens, à medida que aumentam as requisições dos clientes ou percentagem de uso de CPU dos servidores e de memória variam acima (ou abaixo) de nível pré-definidos, novas instâncias são dinamicamente iniciadas (ou desabilitadas) conforme o caso.

O protótipo do *BusInRio* utiliza o serviço de nuvem pública baseado no modelo SaaS oferecido pelo Koding⁶ que disponibiliza instâncias em servidores da AmazonEC2 para que os clientes possam testar aplicações em nuvem. As instâncias utilizadas nesse trabalho foram do tipo t2.medium (2 CPU e 4GB de RAM virtuais), essa configuração de instâncias fornece um ambiente de uso geral de custo reduzido onde se pode hospedar servidores web, ambientes de desenvolvimento e os repositórios de dados e de proveniência.

5.1. *BusInRio* Crawler

O crawler é um módulo central da *BusInRio*, ele executa no servidor e consiste em acessar recursivamente os servidores do data.rio, recuperar e tratar os dados através das regras da abordagem ETL e retroalimentar as bases de dados NoSQL em tempo real com os dados curados sobre os GPS dos ônibus e suas linhas. Para isso, ele coleta e trata os dados em formato JSON e armazena-os sob a forma de documentos no MongoDB e os dados do percurso em CSV e armazena-os no Cassandra.

5.2. Bases NoSQL

Nessa arquitetura utilizamos dois bancos do tipo NoSQL. O Cassandra é um banco de dados não relacional distribuído altamente escalável que se apoia na arquitetura do DynamoDB, da Amazon Web Services e modelo de dados baseado no BigTable⁷. Este sistema apresenta grande velocidade pois guarda os dados curados na memória RAM. Além disso, é orientado a colunas e por esses motivos foi escolhido para representar os dados dos percursos dos ônibus. Foi criada uma *keyspace* chamada transporte e as tabelas "percurso" e "linha" para representar os dados. Em "linha" armazena-se os dados sobre a linha do ônibus (identificação e descrição) e no "percurso" os pontos pelo qual o ônibus passa.

O MongoDB também é uma aplicação não relacional de alta performance, código aberto, sem esquemas e orientado a documentos. Além de ser orientado a documentos,

⁵ <http://www.ionicframework.com>

⁶ <https://koding.com/>

⁷ <https://cloud.google.com/bigtable/>

ele é formado por um conjunto de documentos JSON. Muitas aplicações podem, dessa forma, modelar informações de modo muito mais natural, pois os dados podem ser aninhados em hierarquias complexas e continuar a ser indexáveis e mais fáceis de buscar. Graças ao formato JSON dos dados (visto que a maioria dos dados seriam utilizados em JSON pelo *app*) este servidor foi o escolhido para armazenar e manipular os dados de GPS dos ônibus. Foi criado um banco "transporte" no MongoDB e uma coleção "onibus" e, pode-se facilmente inserir um documento que representa as informações coletadas pelo GPS do ônibus.

5.3. API BusInRio

A API BusInRio é composta por funções acessíveis somente por programação, e que permitem utilizar características da arquitetura, ela permite que a aplicação cliente que executa no *smartphone* interprete e converta os dados dos transportes para o formato JSON que é mais facilmente interpretado pelas aplicações móveis e navegadores web.

Através da API BusInRio o aplicativo envia requisições HTML parametrizadas para o servidor de aplicação, por exemplo, número da linha escolhida pelo usuário e data; então o servidor responde com dados sobre os ônibus (em formato JSON) que será interpretado pelo aplicativo e disposto na interface para visualização do usuário.

5.4. Prova de Conceito

Nesta sub-seção, serão apresentados detalhes sobre a avaliação da solução proposta, incluindo tecnologias utilizadas para o desenvolvimento do protótipo do *app*.

O *app BusInRio* está baseado no framework Ionic. Este é uma SDK que permite o desenvolvimento de aplicativos híbridos capazes de executar em *smartphones* com diversos tipos de sistema operacional e navegadores web apenas utilizando os recursos do HTML5, Javascript e CSS. O framework é baseado no AngularJS e seu núcleo compõe funcionalidades do Apache Cordova. O aplicativo móvel *BusInRio* foi desenvolvido com auxílio deste framework utilizando os dados NoSQL disponibilizados pela API.

O *app BusInRio* permite ao usuário escolher, a partir de um *smartphone*, a linha de ônibus que desejar e consultar a localização dos ônibus daquela linha em tempo real. A Figura 2 ilustra algumas telas da aplicação. Por exemplo, quando o *app* passa como parâmetro o "número da linha" de ônibus, o servidor atende a requisição produzindo um documento JSON contendo as últimas informações sobre cada carro daquela linha, tais como "*ordem, latitude, longitude, data hora e velocidade*". Essas informações são então interpretadas pelo *app* que produz uma visualização amigável para o usuário (figura 2B).



Figura 2. O app *BusInRio* smartphone - consulta às linhas disponíveis (A) e visualizaçãoda localização de cada ônibus da linha e total de ônibus da linha em operação(B).

6. Experimentos e Discussão

Através dos dados abertos disponibilizados pelo data.rio sobre os percursos dos ônibus e GPS em tempo real e com o auxílio de bancos de dados NoSQL MongoDB e Cassandra foi possível estabelecer a ligação entre esses diferentes datasets e fornecer informações úteis ao usuário e de maneira simples. Nossos experimentos utilizaram dados relativos ao mês de julho de 2015.

Através do app *BusInRio* pode-se exibir a localização exata de cada ônibus de uma linha no aplicativo. Ao analisar os dados de localização e velocidade de cada ônibus foi possível gerar um relatório do tipo gráfico exibindo um mapa de calor que representa o trânsito nas principais ruas da malha viária da cidade do Rio de Janeiro através do Google Maps.

Neste caso o usuário, através do navegador do seu *smartphone*, pode identificar claramente os pontos onde o tráfego é mais lento como mostra a Figura 3 (esquerda). As regiões onde o trânsito está mais intenso está representado na cor vermelha e onde flui em melhores condições há um gradiente (do vermelho ao verde). Caso a velocidade média de um conjunto de ônibus que percorram um determinado trajeto comum sejam inferiores a 10Km/h o trânsito na região é considerado intenso (vermelho), e superior 20km/h (verde) é considerado aceitável. Utilizou como valor de referência a velocidade média dos ônibus da cidade de Curitiba-PR (~17Km/h) por ser considerada uma cidade com transporte urbano modelo para o Brasil.

O *appBusInRio* utiliza os dados abertos governamentais curados, ligados e armazenados no MongoDB para estimar a velocidade média dos ônibus localizados nas principais regiões da cidade e para identificar as condições do trânsito, quanto menor a velocidade média do ônibus, maiores as chances de estar retido em um congestionamento.



Figura 3. Mapa de calor com as condições do trânsito calculadas pelo *BusInRio* (A) em comparação com as condições do trânsito avaliadas pelo Google Maps (B). As setas indicam pontos geográficos equivalentes em mapas distintos (Centro do RJ e Zona Sul).

Verificou-se empiricamente se os cálculos do *BusInRio* (sobre as condições do trânsito e linhas do município do Rio de Janeiro) se aproximavam com os de outros produtos, efetuamos dezenas de consultas em dias/horários e comparamos de consultas contra o Google Maps sobre as condições do trânsito calculadas. Verificamos que houve grande correlação. Por exemplo, na Figura 3(A e B) é possível verificar trânsito lento no centro da Cidade do Rio no dia 06/04/2016 por volta das 20:35h, neste dia havia uma manifestação convocada pelos Servidores Públicos do Estado no centro do RJ. Através dos mesmos mapas de densidade de trânsito se verificou que ele estava pouco congestionado nas suas principais vias de escoamento à cidade (av. Brasil) e em diversas partes da Zona Sul da cidade.

7. Conclusões, Limitações e Trabalhos Futuros

Cidades inteligentes são resultantes de conjuntos de ações que objetivam tornar a gestão dos espaços públicos e recursos mais eficientes por meio das TICs. Diversas soluções têm sido propostas para melhorar o problema da mobilidade urbana. Contudo, essas soluções carecem de estudos mais aprofundados.

Este trabalho apresenta a arquitetura *BusInRio*, totalmente baseada em softwares e dados abertos governamentais e em tecnologias NoSQL. As regras da abordagem ETL mostraram-se adequadas, porém limitadas para o tratamento de dados abertos em ambiente de nuvem. Com relação aos repositórios de dados, verificou-se que o MongoDB se mostrou muito superior no que diz respeito a velocidade de resposta das consultas do *app*, especialmente considerando-se que tinha um volume muito maior de registros das viagens dos ônibus. Por questões de limitações de escopo, neste trabalho não apresentamos maiores detalhes sobre a geração de metadados de proveniência dos processos de transformação e ligação dos dados abertos, nem sobre desempenho e volumes de dados manipulados pela arquitetura.

Como trabalhos futuros é sugerido que se desenvolvam novas regras para detecção de *outliers* e dados ausentes. Além disso se buscará que aplicação integre dados de metrô, barcas e trens para oferecer ao usuário uma estimativa aproximada dos melhores percursos e das rotas comparando com outros modais de transportes públicos na cidade do Rio de Janeiro.

Referências Bibliográficas

- Barbosa. L.; Kormaksson, M.; Vieira, M. R.; Tavares, R. L.; Zadrozny. B.; (2014), “Vistradas: Visual Analytics for Urban Trajectory Data”. 15th Brazilian Symposium on Geoinformatics.
- Bessa, A.; Silva F. M.; Nogueira R. F.; Bertini. E, Freire, J. (2015), “RioBusData: Visual Data Analysis of Outlier Buses in Rio de Janeiro”. Symposium on Visualization in Data Science.
- Chee, B. J. S; Franklin Jr. Curtis (2013) “Computação em Nuvem – Tecnologias e estratégias” Ed. M Books.
- Cruz, S.M.S. et al. (2009) “Towards a Taxonomy of Provenance in Scientific Workflow Management Systems”, In: Congress on Services, IEEE, p. 259–266.
- Dohler, M.; Vilajosana, I.; Vilajosana, X.; Llosa, J. (2011) “Smart Cities: An action plan,” in Proc. Barcelona Smart Cities Congress, Barcelona, Spain, pp. 1–6.
- Gil, Y. et al. (2010) "Provenance XG Final Report", W3C, www.w3.org/2005/Incubator/prov/XGR-prov
- Hernández-Muñoz, J. M.; Vercher, J. B.; Muñoz, L.; Galache, J. A.; Presser, M.; Hernández Gómez, L. A.; Pettersson J., “Smart Cities at the forefront of the future Internet,” The Future Internet, LNCS., v. 6656, pp. 447–462
- Leite, D. F. B; Rocha, J. H.; Baptista, C. S.; (2013) “Busão: um Sistema de Informações Móvel para Auxílio à Mobilidade Urbana Através do Uso de Transporte Coletivo”. In: SBSI, 2013, João Pessoa.
- Mendonça, R. R.; Cruz, S. M. S, Campos, M. L. M (2014) “Gerência de Proveniência Multigranular em Linked Data com a Abordagem ETL4LinkedProv”. In SBBD, 2014 Curitiba.
- Open Government Data (2016). <https://opengovdata.org/>. Acesso em março de 2016.
- Open Definition (2016) <http://opendefinition.org/>. Acesso em março de 2016.
- Portal Brasileiro De Dados Abertos. (2013) “Maturidade em dados abertos, entenda as 5 Estrelas”, <http://dados.gov.br/noticia/maturidade-em-dados-abertos-entenda-as-5-estrelas/>. Acesso em: março de 2016.
- Schaffers, H. et al. (2011) “Smart cities and the future internet: Towards cooperation frameworks for open innovation”. The Future Internet, LNCS., v. 6656, pp. 431–446.
- Tito, A. O., et al (2012) “UbiBus: Um Sistema de Informações Inteligentes para Transporte Público”. In: Workshop Tecnologias da Informação e Comunicação nos Grandes Eventos Esportivos (WTICEE), Aracaju-SE.