

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE ELETRONICA
CURSO DE ESPECIALIZAÇÃO EM MBA EM GESTÃO DA TECNOLOGIA DA
INFORMAÇÃO E COMUNICAÇÃO**

LUCIANO AUGUSTO POZZO

**BIG DATA E SUAS VANTAGENS NO RAMO DE MONITORAMENTO DE
VEÍCULOS**

MONOGRAFIA DE ESPECIALIZAÇÃO

CURITIBA - PR

2017

LUCIANO AUGUSTO POZZO

**BIG DATA E SUAS VANTAGENS NO RAMO DE MONITORAMENTO DE
VEÍCULOS**

Monografia apresentada como requisito parcial para obtenção do grau de especialista em Gestão da Tecnologia e Comunicação, do Departamento Acadêmico de Eletrônica da Universidade Tecnológica Federal do Paraná.

Orientador: Prof. MSc. Luiz Pinheiro Junior

CURITIBA - PR

2017



TERMO DE APROVAÇÃO

Big Data e suas Vantagens no Ramo de Monitoramento de Veículos

Por

Luciano Augusto Pozzo

Esta monografia foi apresentada às **20 h** do dia **29/09/2017** como requisito parcial para a obtenção do título de Especialista no CURSO DE ESPECIALIZAÇÃO EM GESTÃO DE TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO, da Universidade Tecnológica Federal do Paraná, **Campus Curitiba**. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho:

1		Aprovado
2		Aprovado condicionado às correções Pós-banca, postagem da tarefa e liberação do Orientador.
3		Reprovado

Prof. MSc. Alexandre Jorge Miziara
UTFPR - Examinador

Prof. MSc. Luiz Pinheiro Junior
UTFPR – Orientador

Prof. MSc. Alexandre Jorge Miziara
UTFPR – Coordenador do Curso

RESUMO

Esta pesquisa tem como tema as vantagens que a tecnologia pode propiciar para um ramo específico de atividade. Sendo assim o objetivo desta pesquisa é explorar como a análise dos dados coletados por dispositivos *Internet of Things (IoT)* podem contribuir para as empresas de transportes rodoviários. Com isso estruturou-se um referencial teórico sobre os temas de *Internet of Things (IoT)*, análise de dados e *Big Data*, e a tecnologia atualmente utilizada nas empresas de monitoramento de veículos, para conhecer em profundidade quais são os autores sobre estes assuntos e o que essa literatura recomenda. Na sequência para compreender o fenômeno investigado realizou-se um estudo de caso qualitativo numa empresa do setor de monitoramento de veículo denominada como *Car Monitor*, nome fictício para garantir a integridade e fidelidade da marca, por se tratar de um trabalho acadêmico. Neste estudo de caso foram coletados dados dos sistemas de Rastreamento e Telemetria veicular, e dados secundários de documentos, manuais e procedimentos técnicos. Na análise dos dados construíram-se tabelas e gráficos visando exemplificar como a coleta de informações sobre os veículos podem ser analisadas e obtidos *insights*, por exemplo, obter os veículos e motoristas que mais consomem combustível na frota; identificar os motoristas que não fazem o período de descanso adequado na sua jornada; criar rotas mais inteligentes agregando informações de postos de combustíveis com preços melhores; otimizar a gestão da frota, evitando veículos ociosos; antecipar problemas ou manutenções de veículos analisando comportamento do motorista. Conclui-se que a análise de dados coletadas por meio de dispositivo e sensores *Internet of Things (IoT)* armazenados em um grande volume de dados conceituado *Big Data*, possam agregar valor competitivo para as organizações de monitoramento de veículos por meio da redução de custos para as frotas de veículos, e até impactando o meio ambiente tendo uma redução na emissão de gases promovendo a sustentabilidade ambiental das empresas e econômica com a redução de custos que a tecnologia. As limitações da pesquisa foram o volume restrito de dados cedidos pela empresa, principalmente dados relacionados aos clientes e a quantidade de fontes de dados, de modo que se sugerem pesquisas futuras como comparar outras fontes de dados, diferenciar os tipos de frotas e agregar novos dados acerca deste tema.

Palavras-chave: análise de dados; big data; monitoramento de veículos.

ABSTRACT

This research has as subject-matter the advantages that technology can provide for a specific line of business. Thus, the purpose of this research is how an analysis of the data collected by Internet of Things (IoT) devices can contribute to road transport companies. For this reason it was structured a theoretical reference to the topics of Internet of Things (IoT), data analysis and Big Data, and the current technology in the vehicle monitoring companies, to know in depth, which are the authors on these subjects and what this literature recommends. Following was made a qualitative case study of a company in the vehicle tracking sector named as Car Monitor, fictitious name to ensure the integrity and fidelity of the brand, because it is an academic work. In this case study, we collected data from Vehicle Tracking and Telemetry systems, and secondary data from documents, manuals and technical procedures. In the analysis of the data were constructed tables and graphs aiming to exemplify how the collection of information about the vehicles can be analyzed and provide insights, for example, to obtain the vehicles and drivers that most consume fuel in the fleet; identify drivers who do not make the adequate rest period on their journey; create smarter routes by aggregating information from fuel stations with better prices; optimize fleet management, avoiding idle vehicles; anticipate problems or maintenance of vehicles by analyzing driver behavior. It is concluded that an analysis of data collected through devices and Internet of Things (IoT) sensors stored in a large volume of data conceptualized as Big Data can add competitive value to vehicle monitoring organizations by reducing vehicle fleet costs, and even impacting the environment with a reduction in the emission of gases promoting environmental sustainability of the companies and economical with the reduction of costs and technology. As limitations of the research, the restricted volume of data provided by the company, mainly customer-related data and quantity of data sources, so that future research is suggested as comparing other data sources, differentiating fleet types and aggregating new data on this subject.

Palavras-chave: data analysis; big data; vehicle tracking.

LISTA DE ILUSTRAÇÕES

Figura 1- Tipos de equipamentos utilizados no monitoramento veicular	19
Figura 2- Arquitetura de funcionamento da solução de monitoramento atual	24
Figura 3- Arquitetura de funcionamento da solução de monitoramento após mensageria distribuída.....	25
Figura 4 - Visão geral consumidores e produtores Kafka	26
Figura 5- Comunicação entre Cluster Kafka e Consumidores	27
Figura 6- Funcionamento do processo de escrita das mensagens na mensageria Kafka	27
Figura 7- Gráfico apresentando piores motoristas da frota	35
Figura 8- Gráfico apresentando comparação entre piores e melhores motoristas da frota	36
Quadro 1- Frota nacional de veículos	18
Quadro 2- Resumo dos dados de pesquisa	28
Quadro 3- Informações do pacote enviado pelos equipamentos e sensores de monitoramento	30
Quadro 4- Resumo após processamento consumidores Kafka dos pacotes de monitoramento	31
Quadro 5- Análise consumo combustível a partir do resumo dos pacotes de monitoramento	32
Quadro 6- Análise média de consumo combustível a partir do resumo dos pacotes de monitoramento	32
Quadro 7- Análise divergência de consumo combustível a partir do resumo dos pacotes de monitoramento	33

LISTA DE SIGLAS

CRM	Customer Relationship Management ou Gestão do Relacionamento com o Cliente
GB	Gigabyte
GPRS	General Packet Radio Service ou Serviço de Rádio de Pacote Geral
GPS	Global Positioning System ou Sistema de Posicionamento Global
IOT	Internet Of Things ou Internet das Coisas
JSON	Javascript Object Notion ou Notação de Objetos JavaScript
RFID	Radio Frequency Identification ou Identificação por Rádio Frequência
RPM	Rotações Por Minuto
SINIAV	Sistema Nacional de Identificação Automática de Veículos
SMS	Short Message Service ou Serviço de Mensagens Curtas
XML	Extensible Markup Language ou Linguagem de Marcação Extendida

SUMÁRIO

1. Introdução	9
1.2 Objetivos específicos.....	10
1.3 Estrutura da monografia.....	10
2. Fundamentação Teórica	11
2.1 Internet das Coisas (Internet of Things)	11
2.2 Análise de Dados e Big Data	13
2.3 Tecnologia nas empresas de transporte	16
3. Metodologia	21
3.1 Caracterização da Pesquisa	21
3.2. Descrição do Caso.....	21
3.3 Coleta de dados	22
3.3.2 Mensageria Kafka	25
4 Análise dos Dados	29
4.4 Implementação	36
5. Considerações Finais	38
Referências.....	40

1. Introdução

É notável a revolução causada pela internet nas últimas décadas, a conectividade e o grande volume de dados tem crescido de forma impressionante a cada dia, sendo que estes dados já não seguem mais necessariamente um padrão estruturado, eles podem vir de redes sociais através de compartilhamento de vídeos, dispositivos inteligentes (IoT), sensores dos mais diversos tipos e formatos, veículos disponibilizando seu posicionamento e informações do comportamento do condutor. Todos estes dados fazem parte de um grande banco de dados extremamente rico e disponível para utilizarmos como apoio para tomada de decisão, melhorar produtos e serviços, ou até mesmo criar novas oportunidades.

Segundo estudo da Gristec de 2014, o país já possuía cerca de 5% de veículos fazendo uso de algum dispositivo de monitoramento, o que corresponde a dois milhões e trezentos mil veículos de uma frota total de quarente e seis milhões para o período, evidenciando o potencial do mercado e conseqüentemente o volume de dados de monitoramento que ainda deverão ser explorados nas próximas décadas. De modo que é imprescindível que empresas do setor de monitoramento veicular se preparem para expandir seus serviços e digerir grandes volumes e novos formatos de dados. O cliente já não se satisfaz apenas em saber o posicionamento do seu veículo ou carga, mas está cada vez mais interessado em como ter um maior retorno do seu investimento, através da redução de gastos ou otimização de logística. Agregar novos dispositivos e sensores às soluções, e extrair todo o potencial das informações disponíveis deste grande banco de dados não estruturado passa a ser vital para se manter em um mercado cada vez mais competitivo.

Sabemos que extrair todo este potencial das informações existentes no mercado de monitoramento é uma tarefa árdua, mas também sabemos que esta tarefa deve ser iniciada o quanto antes por empresas deste ramo, que deverão investir nestas soluções nos próximos anos. Este trabalho, de forma direta, procura expor o funcionamento deste meio e ser um facilitador desta tarefa de iniciação a análise dos dados coletados por dispositivos de monitoramento de veículos. Este trabalho também serve como ponto inicial para uma solução técnica, pois propõe uma implementação simples e robusta com as tecnologias existentes atualmente no mercado.

1.1 Objetivo geral

O objetivo central deste trabalho é explorar como a análise dos dados coletados por dispositivos *Internet of Things (IoT)* podem contribuir para as empresas de transportes rodoviários.

1.2 Objetivos específicos

Com o intuito de atingir o objetivo central, temos como objetivos específicos:

- a) Identificar os benefícios proporcionados pela aplicação de big data em empresas do ramo de monitoramento de veículos;
- b) Verificar a análise de dados dos sensores disponíveis em empresas do ramo de monitoramento de veículos;
- c) Propor a implementação dessas tecnologias no mercado das empresas de transporte rodoviários.

1.3 Estrutura da monografia

Esta monografia é composta por quatro capítulos. No primeiro temos a introdução e os objetivos, como forma de mostrar ao leitor o propósito do estudo. O segundo capítulo, a fundamentação teórica, trazendo informações sobre internet das coisas, big data e o mercado de monitoramento de veículos, criando um arcabouço para o estudo de caso.

Na sequência, no terceiro capítulo é descrito os procedimentos metodológicos, por meio da descrição do método de investigação, a estratégia de pesquisa e definição. No capítulo quatro apresentamos o estudo de caso, decisões de arquitetura para a solução, e implementação da melhoria proposta a partir da análise dos dados.

Por fim o capítulo cinco, temos a conclusão do trabalho, com as ponderações e conclusões sobre pesquisa efetuada, descrevendo as limitações bem como sugestões de estudos futuros.

2. Fundamentação Teórica

Nas próximas seções vamos abordar sobre definições e características da Internet das Coisas, Business Intelligence e Big Data, direcionando o foco para a área de monitoramento de veículos, obtendo condições para aplicarmos em um estudo de caso.

2.1 Internet das Coisas (*Internet of Things*)

O surgimento da *internet das coisas* é associado aos termos *computação ubíqua* e *virtualidade incorporada*, propostos e publicados por Mark Weiser em “The Computer in the 21st Century” (1991), defendendo sua visão de que “elementos especializados de hardware e software, conectados por fios, ondas de rádio e infravermelhos, serão tão onipresentes que ninguém notará sua presença”. O autor ainda afirma que “as mais profundas tecnologias são aquelas que desaparecem. Elas dissipam-se nas coisas do dia-a-dia até se tornarem indistinguíveis”.

Na visão de Atzori, a ideia básica desse conceito é a presença generalizada de uma variedade de coisas ou objetos - como a identificação por radiofrequência (RFID) Tags, sensores, atuadores, celulares, etc. - que são capazes de interagir entre si, buscando um objetivo comum, de modo que estes objetos podem ser identificados pelas seguintes características: ser identificável, comunicar-se e interagir entre si, através de redes interconectadas de objetos, com usuários finais ou outras entidades na rede (ATZORI, 2010).

Após a visão inicial de Weiser, gradativamente podemos perceber a evolução dos dispositivos e nossa dependência cada vez maior. Segundo previsão da CISCO em 2013, é esperado que no ano de 2020 sejam atingidos cerca de 50 bilhões de dispositivos conectados na rede, de modo que podemos vislumbrar um grande número de aplicações desta quantidade de dispositivos interconectadas, onde saídas de dados de sensores se tornam entradas para outros dispositivos, que através de uma lógica inteligente, estas informações podem ser utilizadas como estímulos para atuadores, que executam tarefas, ou ainda alimentar outros objetos. Segundo Weiser, neste ponto é quando o real poder do conceito de internet das coisas será conhecido.

A área de saúde é um dos segmentos que tem se beneficiado da internet das coisas, onde podemos destacar o monitoramento das condições vitais de pacientes ou atletas, acompanhando em tempo real informações sobre quilômetros percorridos, ritmo cardíaco ou calorias gastas, possibilitando antecipar suporte médico ou evitar custos e esforços desnecessários (VERMESAN; FRIESS, 2013). Ainda na área de saúde, podemos destacar o benefício da utilização de internet das coisas para o rastreamento de medicamentos, como forma de coibir a falsificação destes produtos. Segundo Metzner, a rastreabilidade pode ser definida como um sistema de identificação que permite resgatar a origem e a história do produto em todas as etapas da cadeia de suprimentos, que vai da produção da matéria-prima até o uso pelo consumidor final. Sendo que existe uma série de tecnologias que podem ser utilizadas para garantir a rastreabilidade do produto, como: códigos de barras, QR codes (Quick Response) e identificação por radiofrequência (RFID, Radio Frequency Identification) (METZNER, 2014).

Outro segmento que obtém vantagens do emprego da internet das coisas é o segmento agrícola, como por exemplo na melhoria de qualidade do vinho, onde através de sensores é possível monitorar as condições do solo, iniciando automaticamente a irrigação caso necessário, ou também acompanhar o diâmetro dos troncos do vinhedo ou identificar níveis de açúcar das uvas (VERMESAN; FRIESS, 2013).

No segmento de transporte inteligente de mobilidade, Vermesan e Friess (2013) destacam o uso de internet das coisas para o gerenciamento do veículo, monitorando parâmetros e serviços executados para permitir manutenções preventivas, diagnósticos remotos e suporte instantâneo, bem como também antecipar a compra de peças que precisarão ser trocadas em curto prazo. Outro item é o gerenciamento e controle de tráfego, onde os próprios veículos seriam inteligentes o suficiente para se organizarem e evitar congestionamentos, otimizando o uso de energia e tempo. Isto poderia ser feito em cooperação com sensores e controle de tráfego de uma cidade inteligente.

Apesar dos benefícios que a aplicação deste conceito pode proporcionar no nosso cotidiano, é fundamental que também exista uma preocupação com a segurança e privacidade das informações coletadas, pois isso pode ser um limitador para implantação em usuários ou organizações privadas (MIORANDI et al., 2012).

Onde segundo Chabridon et al. (2014) a privacidade é geralmente percebida pelos usuários como uma expectativa de permanecer em um estado de proteção sem ter que o perseguir ativamente. Sendo que os usuários apenas se preocupam efetivamente quando percebem que esta foi violada.

Em contrapartida, na visão de Weber e Weber (2010), um controle rigoroso da privacidade por parte do proprietário pode colocar em risco a veracidade de algumas atividades, ocultando informações que podem determinar atividades criminosas. Deste modo, é importante que em conjunto à aplicação desta tecnologia, que instituições internacionais busquem a construção de um quadro jurídico de base legal para regulamentar as práticas oriundas da internet das coisas, pois atualmente não são satisfatórias para o potencial de exploração de dados deste conceito (SANTOS; SALES; 2015).

2.2 Análise de Dados e Big Data

Em um mercado cada vez mais competitivo e desafiador, o sucesso das organizações está atrelado a eficiência e inteligência com que os processos são realizados. Segundo McGee e Prusak, “na economia de informação, a concorrência entre as organizações baseia-se na sua capacidade de adquirir, tratar, interpretar e utilizar informação de forma eficaz”. De modo que, para que uma empresa se mantenha competitiva no mercado, ela deve ser efetiva ao coletar, armazenar e utilizar os mais diversos tipos de dados, fundamentando sua tomada de decisão para os mais variados cenários.

Segundo Freitas, “a atividade de tomar decisões é crucial para as organizações. Esta atividade acontece todo o tempo, em todos os níveis, e influencia diretamente a performance da organização” (FREITAS et al.,1997, p. 36). Para Thompson Jr. (1995, p. 2), “o produto das decisões modernas de negócios é uma função da quantidade de informações de que os gestores dispõem sobre planos de ação e os possíveis resultados que podem decorrer deles”. Deste modo, para que as tomadas de decisão sejam eficazes, é muito importante embasa-las utilizando informações consistentes, reduzindo riscos e aumentando o retorno do investimento.

Segundo estudo de Gantz e Reinsel (2012), com dados inicialmente coletados em 2005 e se estendendo até 2020, o universo digital vai crescer 300 vezes, de 130

exabytes para 40.000 exabytes. Sendo que de hoje até 2020, o universo digital vai dobrar a cada dois anos. O estudo também aponta que o investimento em infraestrutura vai crescer 40% entre 2012 e 2020, resultando em um barateamento do custo de gigabyte (GB) de \$2.00 para \$0.20. Considerando que é certo que o investimento em áreas como gerenciamento de armazenamento, segurança, big data e computação em nuvem terão um crescimento substancialmente mais rápido. Ainda segundo o estudo, apenas uma pequena fração deste crescente universo de informações tem sido explorada e realmente aproveitada.

Por este motivo é crucial que as empresas estejam preparadas para manipular esta grande quantidade de dados, chegando a decisões que contribuam para a permanência da empresa no mercado. Sendo que um dos pontos importantes para que isso ocorra é conseguir separar informações relevantes deste emaranhado de dados. Santos (2000, p. 1) afirma que "O tomador de decisões necessita de informações relevantes, mas, antes de tudo, precisa de dispositivos de filtros, pois está exposto a uma massa infinita de informações irrelevantes, muitas delas, que ele mesmo solicitara".

Nesta situação, com o crescente volume de informações disponíveis, associado as necessidades constantes dos gestores pela busca de novas ideias, melhores práticas, novas soluções e dispositivos de filtros, emerge o conceito de Big Data, que tem se tornado cada vez mais relevante no processo de tomada de decisões estratégicas, corroborada por Diebold (2012, p. 4), que afirma que "a necessidade de lidar com Big Data, e a necessidade de desbloquear a informação escondida dentro dele, agora é um tema-chave em todas as ciências - sem dúvida o tema científico fundamental dos nossos tempos".

E para Taurion (2013, p.34): "Big Data ainda está no canto da tela do radar dos executivos, mas tem o potencial de ser um disruptor de competitividade entre empresas. Afinal se uma empresa puder obter insights aprofundados sobre seus clientes, o que eles desejam e mesmo opinam sobre a empresa e seus produtos tem condições de mudar o jogo".

Big Data diz respeito ao processamento e análise de enormes repositórios de dados, tão desproporcionais que seria impossível aplicar ferramentas convencionais de análise de banco de dados (VERMESAN; FRIESS 2013). Apesar do termo Big Data estar mais comumente associado ao enorme volume de dados, esta não é a

única característica deste conceito. Doug Laney (2001) articulou a definição de big data a partir de “3Vs”, que constituem: Volume, Velocidade e Variedade.

A respeito do volume, Laney (2001) evidencia a evolução da tecnologia e o crescimento do mercado eletrônico, que de fato vem apresentando ano após ano crescimentos expressivos, sendo que segundo um estudo da International Data Corporation, estima-se que em 2020 sejam atingidos 35 zettabytes de dados digitais criados por ano. Importante notar que esta tendência não está relacionada apenas ao que Laney (2001) se referiu, mas também devido a disseminação de dispositivos inteligentes, que proporcionam conectividade a rede, armazenamento e compartilhamento de informações.

O termo velocidade diz respeito a taxa em que estes dados são gerados, e conseqüentemente a velocidade necessária que se deve analisar e atuar sobre eles (GANDOMI; HAIDER, 2015). Um exemplo é a ferramenta *Analytics* do Google, que permite que o administrador de um site obtenha informações em tempo real dos visitantes do site, como: de onde os visitantes vieram, o que estão visualizando, quais as páginas mais visualizadas, e uma séria de outras métricas analíticas.

Até mesmo comerciantes de varejo convencionais já tem taxas altas de geração de dados, como por exemplo o Wal-Mart, que processa mais de um milhão de transações por hora (CUKIER, 2010). A variedade se refere a heterogeneidade estrutural do conjunto de dados, onde atualmente temos vários tipos de dados, podendo classifica-los em estruturados, não estruturados e semiestruturados.

Dados estruturados correspondem a apenas 5% de todos os dados existentes (CUKIER, 2010) são estes os dados tabulares, existentes em bancos relacionais ou planilhas. Áudio, vídeo, textos e imagens são classificados em dados não estruturados, pois estes dados não possuem uma estrutura organizacional definida. Já dados semiestruturados são dados que não possuem uma conformidade com padrões rígidos, podemos citar o *Javascript Object Notion (JSON)*, *Extensible Markup Language (XML)* e linguagens textuais para troca de informações em geral.

Marr (2014) ainda cita a existência de mais dois “Vs”, que devemos levar em consideração, são: Veracidade e Valor. Para o autor, o enorme volume de dados pode trazer consigo baixa qualidade, de forma que a tecnologia deve ser capaz em garantir a veracidade destes dados. Com relação ao valor, o autor expõe que é fundamental se extrair valor desta imensidão de dados, e também se deve saber os custos benefícios relacionados à esta análise.

2.3 Tecnologia nas empresas de transporte

Estudos recentes (WE ARE SOCIAL, 2017) revelam que já temos mais da metade da população mundial conectada na internet, sendo 3.773 bilhões de pessoas conectadas de um total de 7.476 bilhões, de modo que a penetração na América do Sul é de 66%, indicando que de cada três pessoas, uma acessa a internet. Sendo que o estudo também revela que o avanço da internet cresce a taxa de 10% por ano desde 2016.

Segundo um estudo da empresa Canalsys de 2012, aponta que desde 2011 os Smartphones ultrapassam os computadores pessoais levando em conta as unidades vendidas em todo o mundo. Um outro estudo, elucidativo e correlacionado com a Internet, foi lançado pela Global Web Index em 2015 e indica que 80% dos usuários da Internet em todo o mundo, entre os 16 e os 64 anos, possuem um Smartphone.

Com um número cada vez maior de dispositivos tecnológicos e conectados à Internet, surgem diversas vantagens que têm um impacto considerável no mercado da Internet das Coisas e conseqüentemente no monitoramento de veículos, dentre elas temos: conectividade em qualquer lugar do mundo em tempo real; informações relevantes, respostas a questões de forma rápida e intuitiva através dos dispositivos; custos reduzidos pelo barateamento da tecnologia; compartilhamento e interação de informações entre utilizadores e dispositivos entre si.

O benchmarking de dados é uma técnica sólida e acessível devido a facilidade de aplicação em comparação com outros métodos, possibilitando medir e comparar o desempenho de informações da operação, descobrindo lacunas e oportunidades de melhoria, que nem sempre são facilmente identificadas.

Combinar o conceito geral de benchmarking com o IoT traz muitos benefícios, principalmente por aproveitar dados detalhados de diversos dispositivos em tempo real. Comparando com o passado, relatórios da indústria e benchmarks eram publicados mensalmente, trimestralmente, ou mesmo anualmente, o que significa que as empresas tinham que esperar por semanas, meses ou até mais para se ter informações de comparação. Hoje, através da tecnologia e dispositivos, podemos de forma imediata obter estas informações, que podem gerar conhecimentos analíticos poderosos para aumentar a produtividade, eficiência e segurança.

Independentemente dos tipos de dispositivos utilizados na operação, o volume de detalhes dos dados permite que as empresas passem a compreender melhor as diferentes facetas dos seus negócios de uma forma que não era possível anteriormente. Sendo que quanto maior o volume de dados e detalhe, maior é a capacidade de extrair valor do benchmarking.

Apesar dos benefícios, é importante ressaltar que o benchmarking é sensível a qualidade dos dados coletados, por este motivo é preciso dar atenção especial a qualidade dos dados recuperados, pois podem acarretar em conclusões pouco precisas e levar a decisões equivocadas na operação. Por isso é fundamental que as equipes estejam preparadas e constantemente estudando anomalias nos dados coletados.

No Brasil o mercado de monitoramento de veículos nasceu com foco principalmente da demanda por segurança do transporte de cargas, que pela popularização e barateamento da tecnologia, acabou também expandindo para veículos leves. Desta forma, podemos dividir o mercado de rastreamento de veículos em dois grandes segmentos: o monitoramento de veículos pesados (grandes caminhões) e o de veículos médios e leves (pequenos caminhões, utilitários e carros de passeio em geral).

Colocando em foco o segmento de veículos pesados, podemos dizer que é indiscutível a importância deste segmento no mercado de transportes de qualquer economia, ele está presente em todas as etapas da produção e do consumo de bens e serviços. E apesar da Confederação Nacional do Transporte no Brasil apresentar informações que a crise econômica tem impactado fortemente o setor de transportes, ele ainda representa por cerca de 60% do transporte de cargas no Brasil, o que evidencia a dependência da economia brasileira em sobre este tipo de transporte, sendo este um dos motivos do grande interesse das empresas do segmento de monitoramento em criar melhores ferramentas para o gerenciamento de frotas pesadas.

Já o mercado de frotas leves tem se apresentado em expansão e pouco explorado por empresas de monitoramento, principalmente devido ao volume maior de veículos, como podemos analisar nos dados disponibilizados pelo DENATRAN, conforme o quadro 1 que revela uma frota superior a 96 milhões de veículos no Brasil, sendo composta em sua grande maioria por automóveis, utilitários e motocicletas.

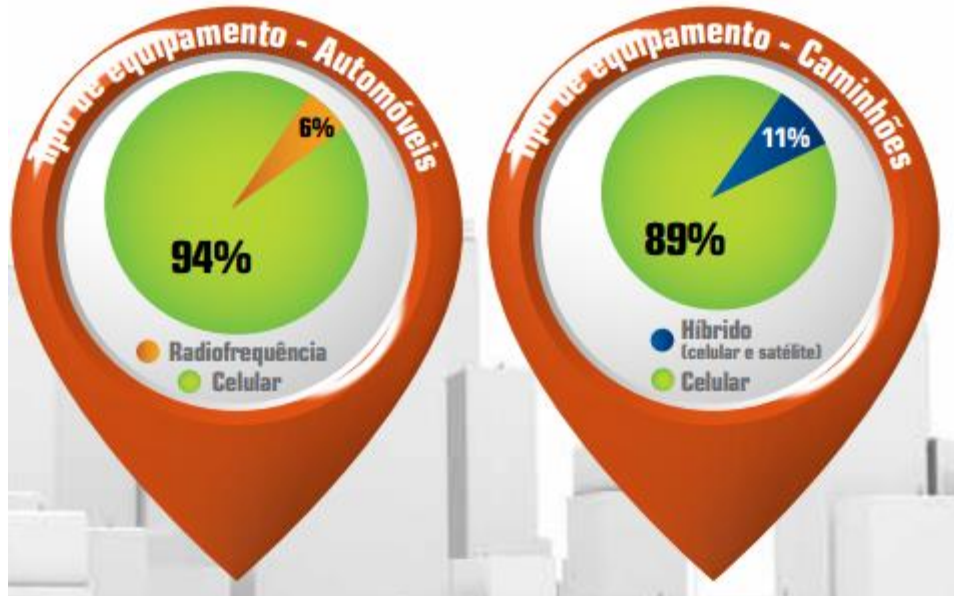
Quadro 1- Frota nacional de veículos

Variável	Brasil
Automóveis	51.296.981
Caminhões	2.684.227
Caminhões-trator	606.679
Caminhonetes	6.880.333
Caminhonetas	3.053.759
Micro-ônibus	383.325
Motocicletas	20.942.633
Motonetas	3.990.558
Ônibus	601.522
Tratores	30.896
Utilitários	707.152

Fonte: Ministério das Cidades (2016)

Segundo o CESVI, no Brasil as empresas que disponibilizam o rastreamento de veículos geralmente fazem uso das tecnologias GPS-GPRS (*Global Positioning System - General Packet Radio Service*), presente em 94% dos automóveis e 89% dos caminhões conforme a figura 1, a rádio frequência, presente em apenas 6% do total de automóveis, e satelital, sendo utilizada em 11% dos caminhões, em conjunto com GPS-GPRS. Sendo que além destas tecnologias de transmissão, também agregam a estes rastreadores um conjunto de sensores e dispositivos para ser possível obter outras informações relevantes do veículo, dependendo da necessidade e do produto contrato pelo cliente.

Figura 1- Tipos de equipamentos utilizados no monitoramento veicular



Fonte: CESVI (2013)

Os rastreadores GPS-GPRS fazem uso da tecnologia de sistema de posicionamento global, originalmente desenvolvida pelo Departamento de Defesa dos EUA, utilizando a navegação via rádio por satélite. O próprio GPS consiste em até 32 satélites que transmitem sinais de micro-ondas para receptores GPS individuais. Estes satélites continuamente orbitam a terra, permitindo que os sistemas de GPS possam calcular com precisão a sua posição na superfície do globo. De forma que o dispositivo receptor GPS em si consiste em uma antena sintonizada para a frequência destes satélites e um relógio de alta precisão. Ao receber sinais de 4 ou mais dos satélites em órbita, o receptor GPS triangula sua posição na superfície terrestre. Em seguida, ele transmite esses dados através da tecnologia de telefonia móvel GPRS, juntamente com qualquer outra informação que recolhe dos sensores acoplados ao veículo, como a velocidade e rotações por minuto do motor.

Estes tipos de rastreadores fazem o envio das informações em tempo real, o que é chamado de rastreamento ativo, entretanto, caso não exista sinal de telefonia na região em que o dispositivo está localizado, o equipamento pode armazenar as informações em memória, e transmiti-las posteriormente quando houver sinal, o que é chamado de rastreamento passivo.

Já o equipamento por radiofrequência ou RF, se baseia na triangulação de sinais captados por redes formadas por grandes antenas receptoras, que permitem

uma ampla cobertura de área, e a localização do objeto rastreado até mesmo em lugares fechados ou cobertos. Este tipo de equipamento geralmente é utilizado em grandes capitais ou regiões metropolitanas, onde existe uma maior concentração destas antenas, outro ponto é que o equipamento é à prova de inibidores de sinais, também conhecido como jammers, e por ser mais simples tem um custo muito inferior ao GPS-GPRS e satelital. Importante citar que atualmente no Brasil temos um grande projeto em debate com esta tecnologia, o Sistema Nacional de Identificação Automática de Veículos (SINIAV). Que prevê que todos os veículos nacionais recebem uma tag RFID, possibilitando a identificação do veículo automaticamente quando o mesmo atravessar um “portal RFID”. Através deste projeto o governo federal busca reduzir o roubo de veículos e cargas, bem como controlar de forma mais efetiva a fiscalização tributária.

Já o rastreamento satelital tem como diferencial garantir a cobertura de sinal mesmo em regiões mais remotas, isso ocorre porque a comunicação das informações é feita através de um terminal satelital instalada no veículo, que se comunica diretamente com a rede de satélites, não havendo dependência com a rede de telefonia móvel, como ocorre na tecnologia GPRS. Mas devido a altos custos dos equipamentos, como também de comunicação, é uma solução utilizada para clientes que tem necessidades específicas, como por exemplo, transportadoras que trafegam em áreas sem cobertura e transportam cargas valiosas.

3. Metodologia

O objetivo geral desta monografia é explorar como a análise dos dados coletados por dispositivos *Internet of Things (IoT)* podem contribuir para as empresas de transportes rodoviário por meio de um Estudo de Caso que expressa a dimensão da análise de dados neste mercado. Nos capítulos a seguir é descrito como foi operacionalizado esta pesquisa bem como maiores informações sobre a coleta e análise dos dados.

3.1 Caracterização da Pesquisa

O método de investigação escolhido para este trabalho foi o Estudo de Caso (YIN, 2013), pois através deste método pode-se examinar uma ou várias entidades em um contexto real, não existindo manipulações ou controle das experiências envolvidas. Também seguindo Gomez, Flores e Jimenez (1996), o objetivo de um Estudo de Caso é explorar, descrever, explicar, avaliar e transformar, conforme foi operacionalizado neste estudo, que é de explorar e mostrar de uma forma prática os benefícios da análise do volume de dados para empresas de monitoramento de veículos.

3.2. Descrição do Caso

Para demonstrar a utilização da Big Data e o valor que ela pode trazer para uma empresa do ramo de monitoramento, é apresentado o processo de coleta de dados, análise e implementação para a melhoria da eficiência de um dos principais ofensores de uma frota de veículos, o consumo de combustível.

Neste estudo de caso utilizou-se de um nome fictício para o caso investigado sendo 'Car Monitor' com o intuito de guardar a integridade da organização pelo fato da pesquisa ser de caráter científico. A Car Monitor é uma empresa de monitoramento de veículos de carga atuante no mercado nacional brasileiro, instalada na região do Sul do Brasil que possui uma frota aproximadamente de 230 mil veículos monitorados e classificados em: veículos de passeio, veículos de cargas leves e veículos de cargas pesadas.

Vale ressaltar que os dados desta pesquisa são reais cedidos pela empresa, sendo que as informações utilizadas na análise dos dados são do período de 01 dia de monitoramento de posicionamento e telemetria de quinze veículos de transporte de carga pesada, de forma que outras informações pertinentes aos veículos e clientes foram omitidas como forma de preservar o sigilo.

3.3 Coleta de dados

Para o estudo de caso a 'Car Monitor' disponibilizou dados coletados em caminhões através do equipamento MTC700 em conjunto com um expansor de entradas e saídas, denominado MX100. Segundo a Maxtrac, a família MTC700 foi desenvolvida para atender as necessidades do mercado de rastreamento e telemetria. Suas funcionalidades entregam soluções para operações de logística, gerenciamento de risco, inteligência embarcada, sistemas de transporte coletivo, dentre outros (MAXTRAC MTC700, 2017).

O equipamento MTC700 calcula diversas informações sobre o veículo monitorado através do uso do sinal de GPS (*Global Positioning System*), tais como: posicionamento (latitude e longitude), data e hora, hodômetro e velocidade. Sendo que para complemento destas informações obtém dados do equipamento expansor, o MX100, que permite coletar dados do veículo (telemetria), dentre eles: rotações por minuto do motor; distância percorrida do veículo; pressão do óleo normal ou não; embreagem acionada; freio acionado por quantas vezes em intervalo de tempo ou percurso; ativação do limpador de para-brisas com limiars para identificação de situação de chuva; freio motor, temperatura do motor; nível de combustível; ignição analisando o tempo entre o acionamento da ignição e a movimentação do veículo; tensão da bateria; integração de leitor iButton, tornando possível a identificação do motorista do veículo (MAXTRAC MTC100, 2017).

Após a coleta das informações o equipamento faz o envio por meio da tecnologia GPRS (General Packet Radio Service), e mesmo caso este serviço se apresente indisponível no momento do envio, o equipamento é capaz de se comunicar com a central através de um modem satelital externo, caso este esteja integrado, ou ainda gerar pacotes de posição e enviar ao servidor por SMS (*Short Message Service*).

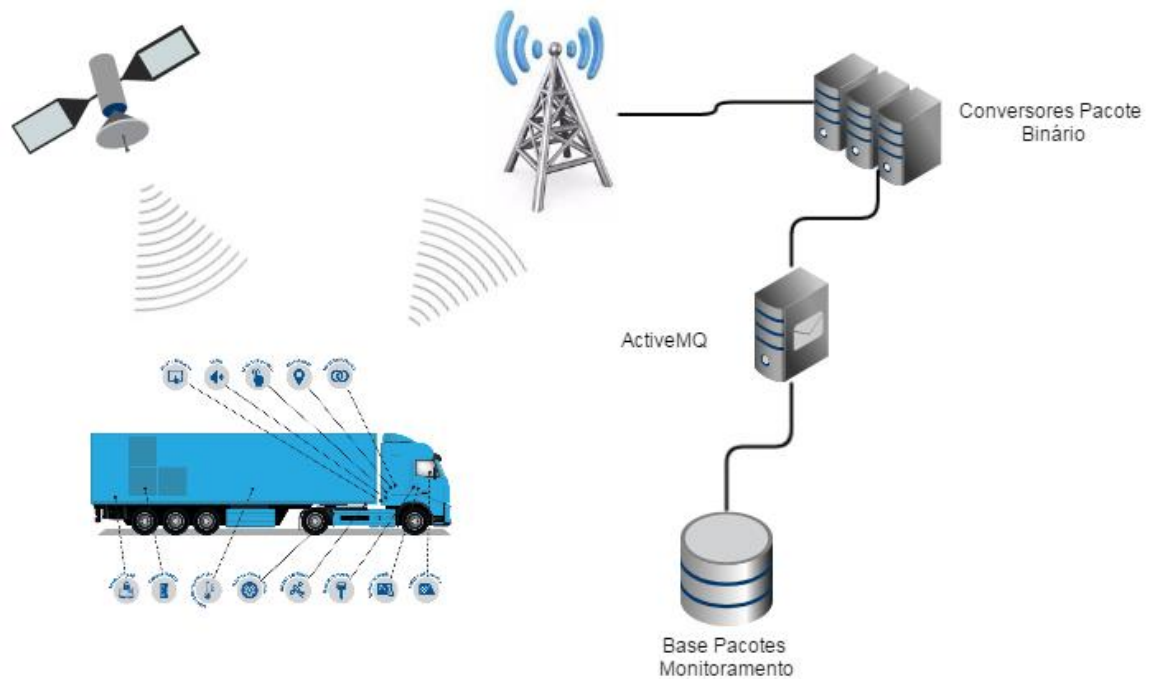
Os dados foram coletados pelos equipamentos MTC700, MX100 e seus sensores, que podem ser caracterizados como um artefato de *Internet of Things*. Os dados deste equipamento são recebidos e divididos por eventos. A partir da análise dos eventos disponibilizados pelos equipamentos de rastreamento e do conjunto de dados dos sensores acoplados a cada veículo, a Car Monitor tem como objetivo reduzir o custo operacional, focando na redução de gasto de combustível da frota, o que levou a análise desses eventos ser relevante para essa pesquisa.

Para este estudo de caso utilizou-se os dados coletados pelos dos rastreadores e os diversos sensores acoplados (IoT) aos veículos para análise do consumo de combustível. Contudo, conforme o modelo evolui e fica mais complexo, se torna necessário integrar diversos outros tipos de dados na *stream* do modelo, como por exemplo informações dos clientes disponibilizadas pelo sistema de *Customer Relationship Management* (CRM) interno, ou ainda, vislumbrando dados não estruturados, imagens de câmeras acopladas aos veículos, traduzindo informações da qualidade da pista trafegada ou ambiente, evoluindo a solução como um todo.

3.3.1 Arquitetura da Coleta de Dados

Atualmente na 'Car Monitor' a obtenção dos dados dos dispositivos instalados nos veículos ocorre por meio da tecnologia GPS-GPRRS citada na seção 2.3 sobre tecnologias nas empresas de transporte de forma que estes pacotes de dados são transferidos em formato binário proprietário do equipamento para um conjunto de servidores, que fazem a conversão destas informações para um formato mais flexível, que em seguida são direcionadas para uma mensageria ActiveMQ, e em seguida são salvas em uma base de dados relacional por um consumidor, conforme a Figura 2.

Figura 2- Arquitetura de funcionamento da solução de monitoramento atual



Fonte: Autoria própria.

O conceito de *Big Data* pode ser utilizado neste estudo porque a ‘Car Monitor’ estima receber em média cinco mil eventos por segundo dos cerca de 230 mil veículos do portfólio de clientes, espalhados pelo Brasil e a arquitetura atual já tem apresentado sinais de sobrecarga e limite de processamento tanto na tecnologia de mensageria atual quanto na base de dados relacional de pacotes de monitoramento, atendendo aos quesitos de Volume e Velocidade que são “Vs” deste conceito (LANEY 2001).

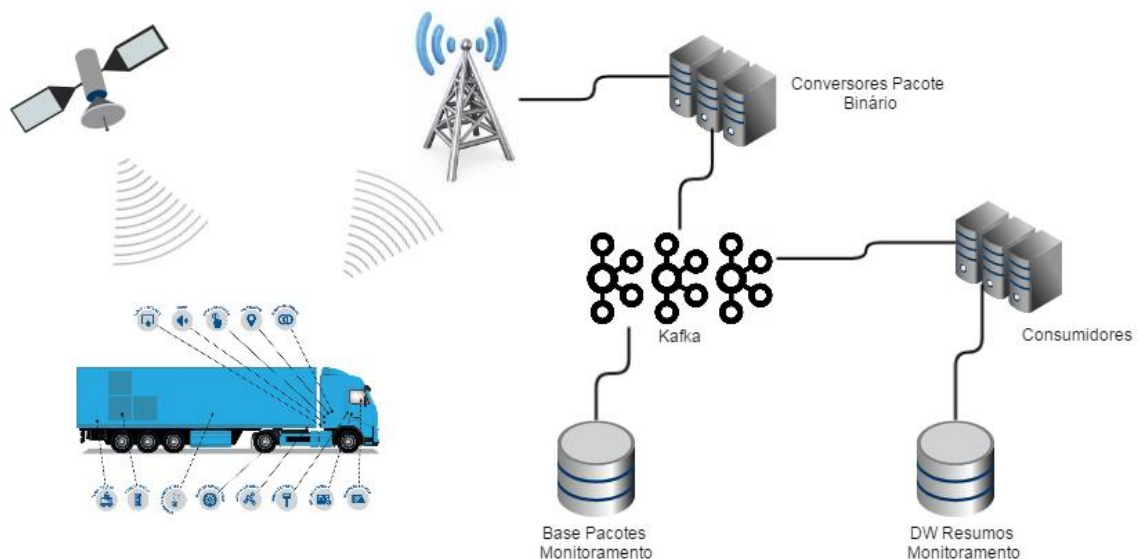
Neste sentido criar uma análise em tempo real de um fluxo massivo de eventos utilizando uma base de dados relacional pode ser um desafio difícil de alcançar. Madden cita que uma das dificuldades desta abordagem é porque primeiro os bancos de dados devem importar dados lentamente em uma representação nativa, antes que eles possam ser consultados, limitando sua capacidade de gerenciar este fluxo de dados.

Madden também comenta que embora os motores de banco de dados relacionais forneçam alguns suportes para análises e modelagem estatísticas, estes esforços não foram amplamente adotados e, como regra geral, não paraleliza efetivamente quando se tem uma quantidade maciça de dados (MADDEN 2012), tornando a atual solução de recepção e armazenamento das informações dos

dispositivos para consultas posteriores insatisfatória para o crescimento de volume de dados esperado.

Devido aos desafios da arquitetura atual, é proposto a utilização de uma mensageria distribuída, em um modelo que comporte uma expansão futura com produtores e consumidores de mensagens em tempo real, explicada na Figura 3, de modo que o processamento e agregação dos eventos seja feito em tempo real através dos consumidores, que armazenam estas informações consolidadas em uma base de dados de resumo.

Figura 3- Arquitetura de funcionamento da solução de monitoramento após mensageria distribuída



Fonte: Autoria própria.

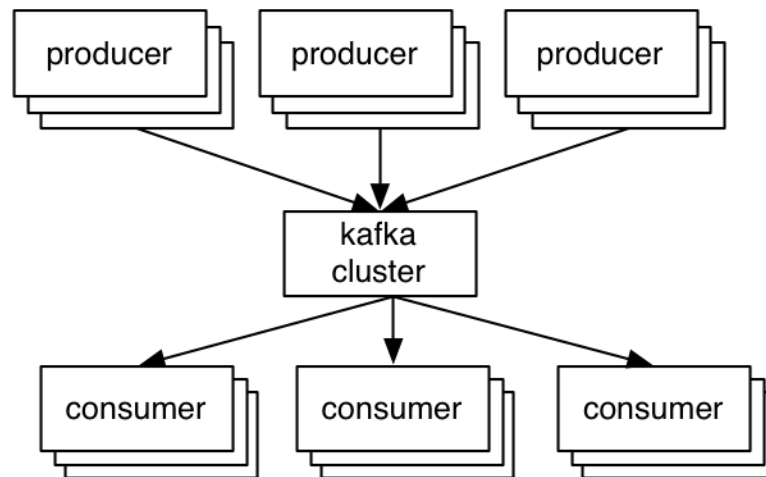
3.3.2 Mensageria Kafka

A *Apache Kafka* é uma mensageria distribuída do tipo *publish-subscribe*, caracterizada por ser rápida, escalável e tolerante a falhas. Sendo que foi originada pela *LinkedIn* devido a necessidade de processamento de grandes quantidades de eventos em tempo real. A *LinkedIn* procurou outras soluções de mensageria antes da criação da *Kafka*, entretanto chegou a conclusão que os produtos existentes no mercado não satisfaziam os requisitos de processamento de grandes volumes de dados e escalabilidade nos padrões que a empresa havia se tornado, pois as

mensagerias estavam focadas em resolver o problema de baixa latência (Goodhope *et al.* 2012). A solução foi a criação da mensageria denominada *Kafka*, que em 2011 foi liberada para a comunidade open source sob licença *Apache License 2.0*.

Na mensageria os produtores enviam mensagens para o cluster de servidores *Kafka*, que mantêm esses registros e cada consumidor interessado se conecta no cluster e obtém os registros conforme a Figura 4.

Figura 4 - Visão geral consumidores e produtores *Kafka*

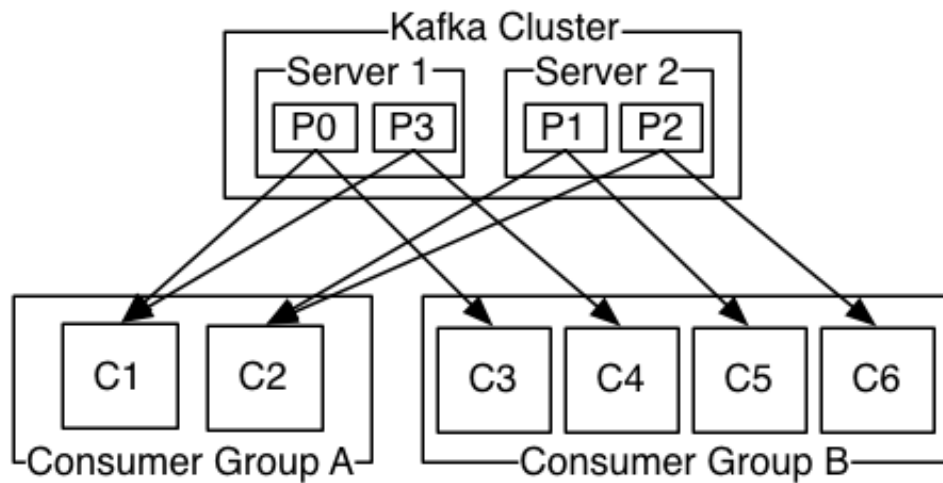


Fonte: Kreps, 2014

A principal abstração em *Kafka* é o *tópico*. Os produtores publicam seus registros (logs) em um tópico, e os consumidores se inscrevem em um ou mais tópicos. Um tópico da *Kafka* é apenas um arquivo onde os registros são escritos de forma sequencial. Os produtores adicionam estes registros e os consumidores se inscrevem para receber as mudanças destes arquivos.

Conforme a Figura 5 cada registro é um par de chave/valor, sendo que a chave é usada apenas para atribuir a gravação a uma partição de log. Este particionamento dos tópicos é importante pois proporciona a escalabilidade da mensageria, sendo possível acomodar estas partições em máquinas diferentes, conectadas por um número maior de consumidores:

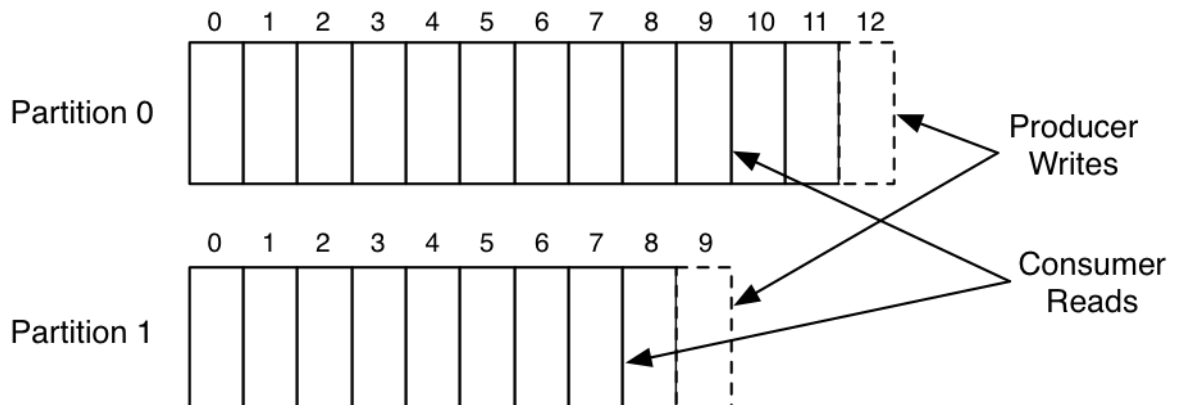
Figura 5- Comunicação entre Cluster Kafka e Consumidores



Fonte: Apache Kafka (2017)

Na Figura 6 é possível observar um exemplo simples de um único produtor e consumidor de leitura e escrita a partir de um tópico de duas partições.

Figura 6- Funcionamento do processo de escrita das mensagens na mensageria Kafka



Fonte: Kreps (2014)

A Figura 6 mostra um produtor adicionando *logs* para as duas partições e um consumidor efetuando a leitura dos mesmos logs. Sendo que cada registro tem um número de entrada associado, chamado de *offset*. O *offset* é utilizado para o consumidor saber em qual posição está efetuando a leitura dos registros.

Um tópico pode estar dividido em diversas partições, sendo que estas partições podem estar espalhadas por um conjunto de máquinas, permitindo que um tópico mantenha um volume muito maior de dados que caberiam em qualquer máquina.

Observe que, ao contrário da maioria das mensagerias, o log sempre é persistente. As mensagens são gravadas imediatamente no sistema de arquivos quando são recebidas e não são excluídas quando são lidas, mas são mantidas por um tempo configurável. Permitindo o uso em situações em que o consumidor de dados pode precisar reprocessar os dados.

Também permite suportar publicação eficiente de economia de espaço, já que há um único log compartilhado, independentemente de quantos consumidores; nos sistemas de mensagens tradicionais geralmente há uma fila por consumidor, então adicionar um consumidor duplica seu tamanho de dados. Isso torna o Kafka um bom ajuste para coisas fora dos limites dos sistemas de mensagens normais. Kafka também replica seus logs em vários servidores para tolerância a falhas.

Portanto, o intuito da utilização da mensageria Kafka é aproveitar a sua natureza de alta escalabilidade, processamento de grandes volumes de mensagens e baixo requisito de hardware, vislumbrando o futuro aumento de processamento de eventos da 'Car Monitor', sendo tanto no monitoramento de novos veículos como na captação de novas informações de outros dispositivos ou dados da internet.

Por fim, segue o quadro 2 que tem como objetivo sumarizar como foram coletados os dados e a realização da análise para operacionalização desta pesquisa.

Quadro 2- Resumo dos dados de pesquisa

Estudo de Caso Único	Coleta de Dados	Fontes de Dados	Análise de Dados
Caso 01 Empresa Car Monitor	Levantamento de Informações em Sistemas	Sistema de Rastreamento Veicular	Análise de Conteúdo
		Sistema de Telemetria Veicular	
		Registros dos Sensores	
	Dados Secundários	Documentos, manuais, procedimentos técnicos	

Fonte: Autoria própria.

Conforme o quadro 2 realizou-se um estudo de caso único na empresa “Car monitor” coletando dados nos sistemas de Rastreamento e de Telemetria veicular. Para comparar essas informações utilizou-se de dados secundários como manuais, documentos e procedimentos técnicos. Contudo juntaram se todas as informações e foram analisadas em planilhas, gráficos e descritivas com o objetivo de prover uma explanação acerca do assunto (YIN, 2015; GRAEBNER; EISENHARDT, 2007).

4 Análise dos Dados

Como podemos notar, o processo de coleta dos dados é crucial para o sucesso das fases seguintes, por este motivo que neste ponto é desejável utilizar a experiência de um cientista de dados aliada aos analistas de negócio, de forma que se consiga aproveitar ao máximo as informações, correlacionando-as de forma correta para que a pergunta seja respondida.

Nosso estudo de caso, olhando para a perspectiva de economia de combustível, evidenciou-se que além dos dados de nível de combustível do tanque, também se necessita das distancias percorridas pelos veículos, para que seja feita uma média de consumo. Outras informações que também podem ser levadas em consideração são as dos sensores do funcionamento do veículo, como o sensor de RPM (Rotações Por Minuto) e velocidade conforme informação coletada no Quadro 3, que além de informar se o veículo está atuando em uma faixa de rotação e velocidades adequadas, também pode informar quanto tempo o veículo fica parado em funcionamento, sendo um ofensor direto para o consumo de combustível. Para esta análise também interessa saber o uso de marcha-lenta mais conhecida vulgarmente como “banguela”, que também é um ofensor para a análise de consumo de combustível.

Quadro 3- Informações do pacote enviado pelos equipamentos e sensores de monitoramento

LAT	LONG	DT_PACT	VEIROID	MOTOID	NVL_COMB	RPM	ODOMETRO	VELOCIDADE
-23,46728667	-46,47145833	24/05/2017 18:38	358113	668167	193	600	1123902	0
-23,46728833	-46,47145833	24/05/2017 18:38	358113	668167	193	602	1123902	0
-23,46729	-46,47146	24/05/2017 18:36	358113	668167	194	602	1123902	0
-23,46729	-46,47146	24/05/2017 18:36	358113	668167	194	597	1123902	0
-23,46730167	-46,471465	24/05/2017 18:35	358113	668167	194	608	1123902	1
-23,467305	-46,47146833	24/05/2017 18:35	358113	668167	194	611	1123902	0
-23,46737	-46,47146167	24/05/2017 18:35	358113	668167	195	591	1123902	0
-23,467425	-46,47145167	24/05/2017 18:34	358113	668167	195	636	1123902	0
-23,467425	-46,47145167	24/05/2017 18:34	358113	668167	195	600	1123902	0
-23,46743667	-46,47144833	24/05/2017 18:32	358113	668167	196	600	1123902	1
-23,46745167	-46,47146	24/05/2017 18:32	358113	668167	196	611	1123902	0
-23,46745167	-46,47146	24/05/2017 18:32	358113	668167	196	600	1123902	0
-23,46741333	-46,471465	24/05/2017 18:31	358113	668167	196	619	1123902	0
-23,46729833	-46,47126167	24/05/2017 18:30	358113	668167	197	676	1123901	0
-23,46724167	-46,471075	24/05/2017 18:30	358113	668167	197	1284	1123901	10
-23,467955	-46,47167333	24/05/2017 18:29	358113	668167	197	1081	1123900	5
-23,46802667	-46,47163667	24/05/2017 18:28	358113	668167	197	696	1123899	0
-23,46787167	-46,47026667	24/05/2017 18:28	358113	668167	198	1432	1123898	19
-23,46727333	-46,469885	24/05/2017 18:28	358113	668167	198	973	1123897	5
-23,46527167	-46,47013	24/05/2017 18:27	358113	668167	198	1086	1123895	14
-23,46527167	-46,47013	24/05/2017 18:27	358113	668167	198	1086	1123895	14

Fonte: Autoria própria.

Para análise dos dados coletados foram criados consumidores para os eventos recebidos na mensageria Kafka, e neles implementou-se regras para agregação dos dados, agrupados em veículo e seu condutor, visto que um mesmo veículo da frota pode ser utilizado por diversos motoristas em uma rota. Conforme cada evento chega na mensageria, o consumidor agrega os resultados em memória, produzindo o seguinte resumo descrito no Quadro 4.

Quadro 4- Resumo após processamento consumidores Kafka dos pacotes de monitoramento

idVeiculo	idMotorista	consumo(km/l)	tpParLig(s)	tpParDesl(s)	velMedia(km/h)	rpmMedia	somaCombGasto
522934	87652	0,28	1320	60	7,24	959	36
830148	3956	0,34	11640	0	6,64	911	117
830148	2658	0,36	104220	900	16,76	972	131
830148	4480	0,51	12000	0	13,63	1031	114
830136	3778	0,69	15300	3240	12,81	992	274
830136	2860	0,72	10680	22920	13,12	944	232
603271	94482	0,81	2880	12720	10,35	813	29
812511	24542	2,63	780	0	47,18	882	26
522934	56712	3,26	720	840	40,65	1296	42
380473	916289	3,41	4920	35700	36,32	1172	151
682064	9623	3,54	3120	147060	38,33	1022	110
852986	39464	3,55	3000	13560	52,59	1061	153
830136	1953	3,72	180	0	39,94	1155	5
812511	67332	3,98	2220	8460	120,73	1123	664
951322	43531	4,06	4380	29040	59,33	1151	161
478213	6	4,23	3720	18540	47,93	1213	216
1107288	5259	4,33	1080	3780	40,06	1181	42
849763	77308	4,34	2700	107640	39,81	1192	142
979756	34825	4,34	360	0	55,5	1283	55
767771	80102	4,44	16980	13380	60,55	1081	84
1138740	630	4,44	3240	128880	46,61	1068	160
812511	14282	4,45	1260	720	114,63	1142	122
1123796	3604	4,56	660	0	63,56	1179	106
754946	190557	4,87	1500	60	58,59	1144	56
358113	668167	5,98	88140	30480	61,86	1375	206
613668	2929	6,43	1500	2520	41,76	1069	79

Fonte: Autoria própria.

A partir do resumo em tempo real descrito no Quadro 4 podemos digerir mais facilmente as informações e responder os questionamentos relacionados ao consumo. Questiona-se quais os veículos e seus respectivos motoristas que consomem mais combustível na frota.

Respondendo esta pergunta pode-se focar nos veículos que têm maiores problemas de consumo, que conseqüentemente apresentarão melhor efetividade em ações para redução do custo operacional. Seguindo esta linha de pensamento, chegamos a estes veículos e motoristas, apresentados no Quadro 5.

Quadro 5- Análise consumo combustível a partir do resumo dos pacotes de monitoramento

idVeiculo	idMotorista	consumo(km/l)	tpParLig(s)	tpParDesl(s)	velMedia(km/h)	rpmMedia	somaCombGasto
812511	67332	3,98	2220	8460	120,73	1123	664
830136	3778	0,69	15300	3240	12,81	992	274
830136	2860	0,72	10680	22920	13,12	944	232
478213	6	4,23	3720	18540	47,93	1213	216
358113	668167	5,98	88140	30480	61,86	1375	206
951322	43531	4,06	4380	29040	59,33	1151	161
1138740	630	4,44	3240	128880	46,61	1068	160
852986	39464	3,55	3000	13560	52,59	1061	153
380473	916289	3,41	4920	35700	36,32	1172	151
849763	77308	4,34	2700	107640	39,81	1192	142
830148	2658	0,36	104220	900	16,76	972	131
812511	14282	4,45	1260	720	114,63	1142	122

Fonte: Autoria própria.

Questiona-se sobre o Quadro 5 dos veículos de mesma categoria e que mais consomem combustível, se a média de consumo está dentro da média da frota. Para julgar se o consumo de um determinado tipo de veículo podemos utilizar como os valores médios divulgados pelo fabricante do veículo, que são dados secundários desta pesquisa, entretanto é desejável uma comparação mais real, por este motivo passamos a utilizar a média dos outros veículos do mesmo tipo da frota no Quadro 6 chegando ao resultado abaixo.

Quadro 6- Análise média de consumo combustível a partir do resumo dos pacotes de monitoramento

idVeiculo	idMotorista	consumo(km/l)	tpParLig(s)	tpParDesl(s)	velMedia(km/h)	rpmMedia	somaCombGasto
812511	67332	3,98	2220	8460	120,73	1123	664
830136	3778	0,69	15300	3240	12,81	992	274
830136	2860	0,72	10680	22920	13,12	944	232
478213	6	4,23	3720	18540	47,93	1213	216
358113	668167	5,98	88140	30480	61,86	1375	206
951322	43531	4,06	4380	29040	59,33	1151	161
1138740	630	4,44	3240	128880	46,61	1068	160
852986	39464	3,55	3000	13560	52,59	1061	153
380473	916289	3,41	4920	35700	36,32	1172	151
849763	77308	4,34	2700	107640	39,81	1192	142

830148	2658	0,36	104220	900	16,76	972	131
812511	14282	4,45	1260	720	114,63	1142	122
830148	3956	0,34	11640	0	6,64	911	117
830148	4480	0,51	12000	0	13,63	1031	114
682064	9623	3,54	3120	147060	38,33	1022	110
1123796	3604	4,56	660	0	63,56	1179	106
767771	80102	4,44	16980	13380	60,55	1081	84
613668	2929	6,43	1500	2520	41,76	1069	79
754946	190557	4,87	1500	60	58,59	1144	56
979756	34825	4,34	360	0	55,5	1283	55
522934	56712	3,26	720	840	40,65	1296	42
1107288	5259	4,33	1080	3780	40,06	1181	42
522934	87652	0,28	1320	60	7,24	959	36
603271	94482	0,81	2880	12720	10,35	813	29
812511	24542	2,63	780	0	47,18	882	26
830136	1953	3,72	180	0	39,94	1155	5
	Média Geral	3,24					

Fonte: Autoria própria.

Por meio desta comparação pode-se perceber que existem veículos nas mesmas condições da frota apresentando consumo adicional de quase 80% em comparação a média total da frota explicado no Quadro 7, o que se questiona se é uma oportunidade de economia para estes veículo e motoristas.

Quadro 7- Análise divergência de consumo combustível a partir do resumo dos pacotes de monitoramento

idVeiculo	idMotorista	consumo(km/l)	consumo adicional
812511	67332	3,98	
830136	3778	0,69	78,7%
830136	2860	0,72	77,8%
478213	6	4,23	
358113	668167	5,98	
951322	43531	4,06	
1138740	630	4,44	
852986	39464	3,55	
380473	916289	3,41	
849763	77308	4,34	
830148	2658	0,36	88,9%
812511	14282	4,45	

Fonte: Autoria própria.

Assumindo o preço do diesel orçado em 29/06/2017 de R\$ 2.60 o litro na região de Curitiba-PR (Preço dos Combustíveis Curitiba, 2017), e o gasto diário de combustível destes veículos de 212 litros, seriam valores então aproximados em R\$ 551,20 gastos diariamente, de forma que se for implementado uma solução para redução do gasto considerando como objetivo o percentual médio da frota, obteríamos uma economia de R\$ 454,00 diariamente, sendo 18%.

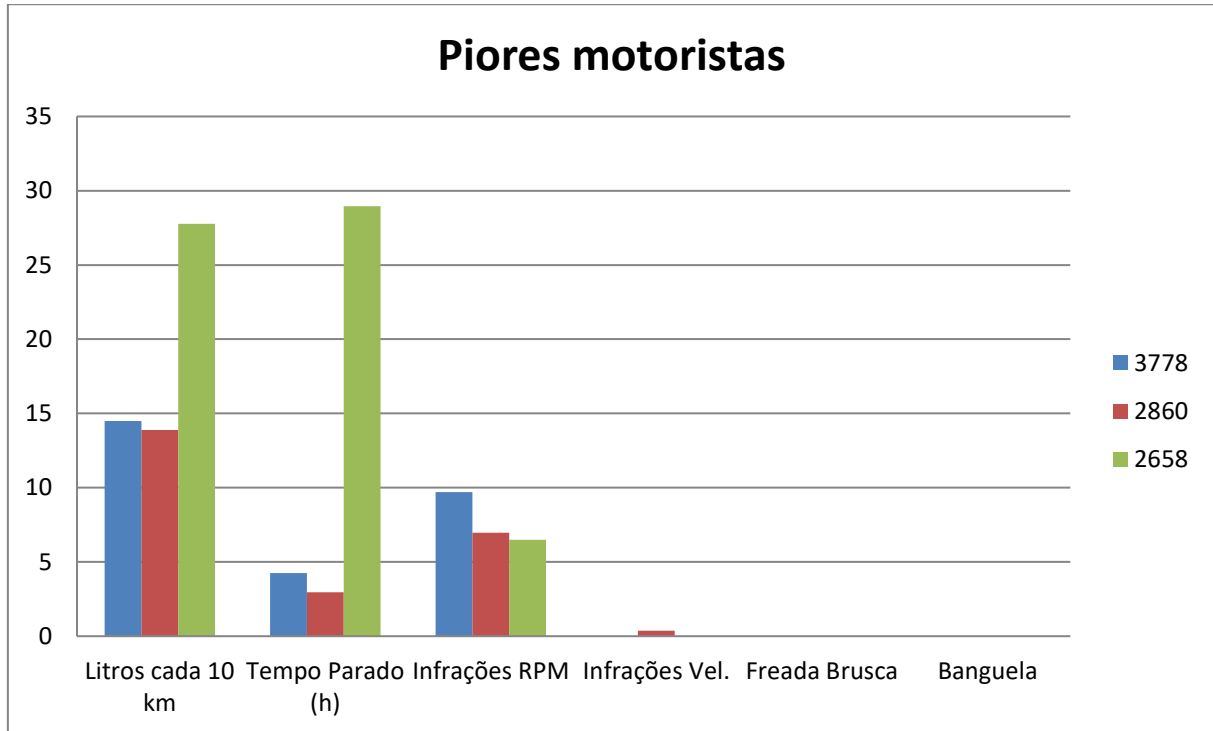
Na sequência questiona-se porque a média de consumo é tão baixa para estes veículos. Foi possível identificar que muitas variáveis podem influenciar no consumo de combustível, como peso transportado, tipo do terreno trafegado, clima. Mas para este estudo vamos nos aprofundar no comportamento do motorista, pois temos estas informações no pacote de dados obtidos.

Ainda por meio do processamento em tempo real e resumo criado, pode-se definir motoristas com comportamento inadequado contabilizando infrações por km rodado através dos eventos, infrações estas que em geral tem impacto no consumo de combustível, como:

- Giro excessivo do motor, considerando a faixa ideal de rotação do motor para um bom de consumo de combustível de um caminhão ser entre 1500 a 2000 RPMs (Rotações Por Minuto);
- Utilização da marcha-lenta (banguela). Pois sabemos que atualmente os motores com injeção eletrônica consomem menos combustível ao estarem engatados do que utilizando o ponto morto, fora que a marcha-lenta, além de ser proibida segundo o código de Trânsito Brasileiro, artigo 231, também implica em um desgaste maior de freio, ao deixar de usar o freio-motor;
- Aceleração brusca, que pode ser obtida analisando a diferença de velocidade e RPM (Rotações Por Minuto) entre eventos, um bom motorista efetua reduções e acelerações de forma suave;
- Tempo de veículo parado com motor ligado, apesar de parecer um dado inofensivo para o consumo, segundo o Escritório de Eficiência Energética e Energia Renovável, deixar o carro ligado quando parado pode representar um gasto de um a dois litros de combustível por hora, dependendo do tamanho do motor e do uso do ar-condicionado.

Aplicando as definições de infrações citadas anteriormente foi possível obter os motoristas com piores comportamentos explicados na Figura 7, o que pode estar contribuindo para a redução da média de combustível.

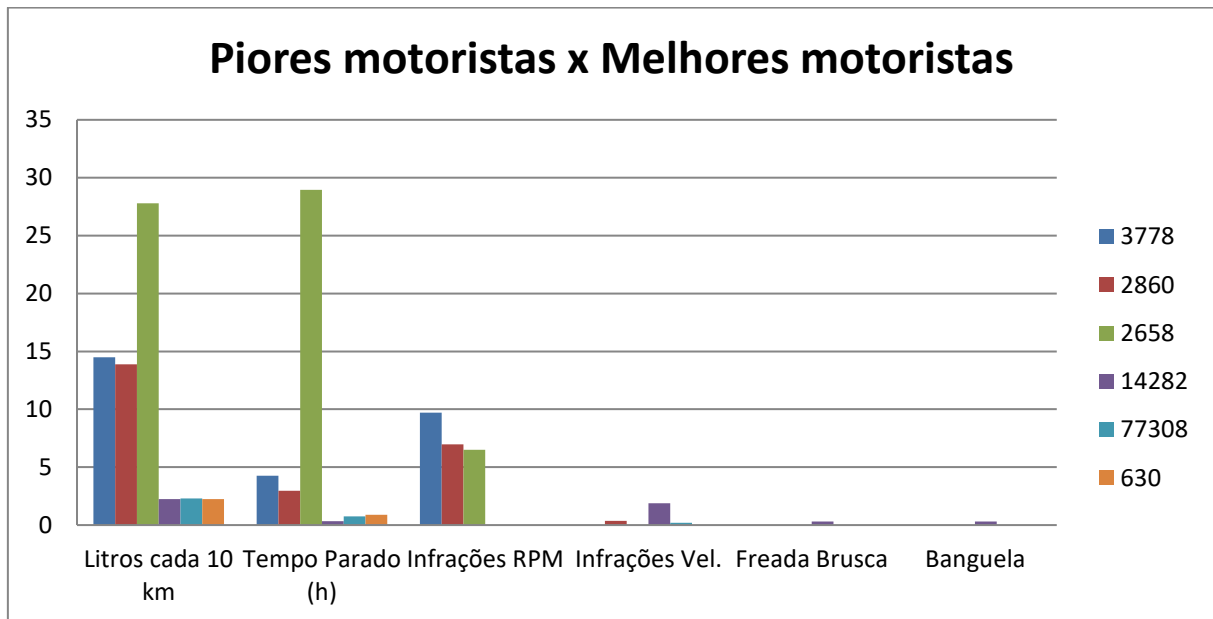
Figura 7- Gráfico apresentando piores motoristas da frota



Fonte: Autoria própria.

Em seguida foi realizada uma comparação com comportamentos que apresentam melhores médias de consumo de combustível na Figura 8.

Figura 8- Gráfico apresentando comparação entre piores e melhores motoristas da frota



Fonte: Autoria própria.

A partir desta análise evidenciou-se que o comportamento inadequado dos motoristas implica em um consumo maior de combustível e conseqüentemente um custo maior para o dono da frota.

Neste caso, identificou-se um consumo superior de 80% comparado a média da frota, onde os fatores que provavelmente mais contribuem com este consumo exacerbado e que devem ter maior atenção são: o grande número de infrações de RPM (Rotações Por Minuto), onde os motoristas aceleram mais do que o necessário, fugindo da faixa de economia de combustível; o tempo excessivo de veículo ligado parado; e as violações de velocidade, que apesar de evidenciarmos um número pequeno, influenciam tanto no consumo de combustível quanto no risco de acidentes.

4.4 Implementação

Após a análise dos dados nesta pesquisa, foram obtidos *insights* para contribuir com a empresa “Car Monitor” oferecendo melhorias para que seus clientes corrijam o comportamento dos motoristas, direcionando para um melhor consumo de combustível.

Algumas estratégias podem ser implementadas, como:

- Monitorar o comportamento do motorista diariamente através da telemetria e instruir sobre a importância de se adequar as melhores práticas ao volante, evitando infrações, principalmente relacionadas ao consumo de combustível;
- Criar um ranking de motoristas e incentivar financeiramente os melhores em comportamento ao volante, gerando uma competitividade saudável entre os motoristas;
- Notificar o motorista de forma sonora quando uma infração ocorrer, instruindo em tempo real para um melhor comportamento ao volante;

Contudo esta pesquisa conclui-se que existe um grande potencial a ser explorado nas empresas de monitoramento de veículos, podendo proporcionar para seus clientes um gerenciamento de frotas mais eficiente e até impactando o meio ambiente por meio da diminuição emissão de gases, promovendo a sustentabilidade ambiental das empresas e econômica com a redução de custos operacionais, evidenciando as vantagens que o uso da tecnologia, leia-se neste caso IoT e Big Data, proporcionam para empresas deste ramo de negócio.

5. Considerações Finais

A partir do estudo de caso foi possível observar o potencial que empresas de monitoramento de veículos tem para proporcionar para seus clientes um gerenciamento de frotas mais eficiente e conseqüentemente lucrativa a partir da coleta e extração de valor de grandes quantidades de informação. Evidentemente que os percentuais de lucratividade podem contribuir para os clientes com frotas maiores e uma operação mais complexa, mas o resultado em clientes menores tem impacto significativo no resultado das empresas.

Apesar da análise apresentada neste estudo de caso se estender a apenas um ofensor do custo operacional de uma frota, através dos dados obtidos é possível vislumbrar diversos outros insights, como:

- Identificar motoristas que não fazem o período de descanso adequado na sua jornada, através da análise dos tempos de paradas através do funcionamento do motor;
- Criar rotas mais inteligentes utilizando informações de postos de combustíveis com preços melhores, rotas com melhor fluxo de tráfego e melhor consumo médio de combustível;
- Otimizar a gestão da frota evitando ociosidade de veículos;
- Antecipar problemas ou manutenções de veículos analisando o comportamento do motorista;

Esta pesquisa visa contribuir para o meio aplicado de uma forma que a análise de dados coletadas por meio de dispositivo e sensores IoT armazenados em um grande volume de dados conceituado *Big Data* possam agregar valor competitivo para as organizações de monitoramento de veículos, respondendo então o objetivo proposto inicialmente na pesquisa.

A análise de dados pode contribuir por meio da redução de custos para as frotas de veículos, e até impactando o meio ambiente tendo uma redução na emissão de gases promovendo a sustentabilidade ambiental das empresas e econômica com a redução de custos que a tecnologia, leia-se neste caso IoT e Big Data podem auxiliar neste sentido.

As limitações da pesquisa foram o volume restrito de dados cedidos pela empresa, principalmente dados relacionados aos clientes e a quantidade de fontes de dados, de modo que se sugerem pesquisas futuras como comparar outras fontes de dados, diferenciar os tipos de frotas e agregar novos dados acerca deste tema.

Referências

Amir Gandomi, Murtaza Haider, **Beyond the hype: Big data concepts, methods, and analytics**, International Journal of Information Management, Volume 35, Issue 2, 2015, Pages 137-144, ISSN 0268-4012, <http://dx.doi.org/10.1016/j.ijinfomgt.2014.10.007>.

APACHE KAFKA (2017), **Documentação oficial Apache Kafka**. Disponível em <https://kafka.apache.org/documentation/>. Acesso em 17 julho de 2017.

Atzori, L., Iera, A. and Morabito, G. (2010) **The Internet of Things: A Survey**. In Computer Networks. Disponível em <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.719.9916&rep=rep1&type=pdf>. Acesso em 27 agosto de 2017.

Carlos Cesar Santos, Jefferson David de Araújo Sales (2015). **O Desafio da Privacidade na Internet das Coisas**.

CHABRIDON, Sophie et al. **A survey on addressing privacy together with quality of context for context management in the Internet of Things**. annals of telecommunications-Annales des télécommunications, v. 69, n. 1-2, p. 47-62, 2014.

Cukier K. (2010), The Economist, **Data, data everywhere: A special report on managing information**. Disponível em <http://www.economist.com/node/15557443>. Acesso em 15 de julho de 2017.

DIEBOLD, Francis. **On the Origin (s) and Development of the Term “Big Data”**. Pennsylvania: Penn Institute for Economic Research, Department of Economics, University of Pennsylvania, 2012.

FREITAS, H. et al. **Informação e Decisão: sistemas de apoio e seu impacto**. Porto Alegre: Ortiz, 1997.

GANTZ, John; REINSEL, David. **The digital universe in 2020: Big Data, nigger digital shadows, and biggest growth in the far east**. IDC, New York, v. 3, Dec. 2012.

GRABNER, M. E.; EISENHARDT, K. M. **Theory Building from cases: opportunities and challenges**. Academy of Management, v. 50, N. 1, p. 25-32, 2007

Jay Kreps (2014), **Benchmarking Apache Kafka: 2 Million Writes Per Second**. Disponível em <https://engineering.linkedin.com/kafka/benchmarking-apache-kafka-2-million-writes-second-three-cheap-machines>. Acesso em 17 julho de 2017.

Ken Goodhope, Joel Koshy, Jay Kreps, Neha Narkhede, Richard Park, Jun Rao, Victor Yang Ye (2012), **Building LinkedIn's Real-time Activity Data Pipeline**. Disponível em <<http://sites.computer.org/debull/A12june/pipeline.pdf>>. Acesso em 18 julho de 2017.

Sam Madden (2012), **From Databases to Big Data**. Disponível em <<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6188576>>. Acesso em 10 de setembro de 2017.

Marr. B (2014). **Big Data, The 5Vs**. Disponível em <<https://pt.slideshare.net/BernardMarr/140228-big-data-volume-velocity-variety-varacity-value>>. Acesso em 15 julho de 2017.

MINISTÉRIO DAS CIDADES. Departamento Nacional de Trânsito. DENATRAN (2016). Disponível em: <<http://cidades.ibge.gov.br/painel/frota.php>>. Acesso em 15 de junho de 2017.

MAXTRAC MTC700 (2017). Disponível em <<https://maxtrack.com.br/produtos/rastreadores-automotivos/mtc-700/>>. Acesso em 15 de junho de 2017

MAXTRAC MX100 (2017). Disponível em <<https://maxtrack.com.br/produtos/acessorios-e-perifericos/mx-100/>>. Acesso em 15 de junho de 2017

MIORANDI, Daniele et al. **Internet of things: Vision, applications and research challenges**. Ad Hoc Networks, v. 10, n. 7, p. 1497-1516, 2012.

Vivian Cristina Velloso Metzner, Roberto Fray da Silva, Carlos Eduardo Cugnasca (2014), **Modelo De Rastreabilidade De Medicamentos Utilizando Identificação Por Radiofrequência, Redes De Sensores Sem Fio E O Conceito De Internet Das Coisas**.

MINISTÉRIO DAS CIDADES. Departamento Nacional de Trânsito. DENATRAN (2016). Disponível em: <<http://cidades.ibge.gov.br/painel/frota.php>>. Acesso em 15 de junho de 2017.

Preço dos Combustíveis Curitiba (2017). Disponível em <<http://www.precodoscombustiveis.com.br/postos/cidade/4005/pr/curitiba>>. Acesso em 29/06/2017.

SANTOS, Raimundo N. **Métodos e ferramentas para gestão de inteligência e do conhecimento**. Perspectivas em Ciência da informação, v. 5, n. 2, p. 205-215, 2000.

TAURION, Cezar. **Big Data**. Rio de Janeiro: Brasport, 2013.

THOMPSON Jr., Arthur. **Tomada de Decisão sob Condições de Certeza, Risco e Incerteza**. Porto Alegre: UFRGS, 1995.

Vermessan, O., & Friess, P. (2013). **Internet Of Things – Conversing Technologies for Smart Environments and Integrated Ecosystems**. River Publishers.

WE ARE SOCIAL. **Digital in 2017 Global Overview**. Disponível em: <<https://wearesocial.com/special-reports/digital-in-2017-global-overview>>. Acesso em 10 de julho de 2017.

WEBER, Rolf H.; WEBER, Romana. **Internet of Things**. New York: Springer, 2010.

Yin, R. K. (2013). **Case Study Research: Design and Methods**. (5th ed.). London: Sage Publications.