

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

**BRENNO SERRATO
GUSTAVO DE CAMARGO**

MODELO DE MONITORAMENTO DE DADOS VEÍCULARES

CURITIBA

2023

**BRENNO SERRATO
GUSTAVO DE CAMARGO**

MODELO DE MONITORAMENTO DE DADOS VEÍCULARES

Vehicle data monitoring model

Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Sistemas de Informação do Curso de Bacharelado em Sistemas de Informação da Universidade Tecnológica Federal do Paraná.

Orientador: Prof. Dr. Robson Ribeiro Linhares

CURITIBA

2023



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**BRENNO SERRATO
GUSTAVO DE CAMARGO**

MODELO DE MONITORAMENTO DE DADOS VEÍCULARES

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção do
título de Bacharel em Sistemas de Informação
do Curso de Bacharelado em Sistemas de
Informação da Universidade Tecnológica
Federal do Paraná.

Data de aprovação: 26/junho/2023

Robson Ribeiro Linhares
Doutor
Universidade Tecnológica Federal do Paraná

Ana Cristina Barreiras Kochem Vendramin
Doutora
Universidade Tecnológica Federal do Paraná

João Alberto Fabro
Doutor
Universidade Tecnológica Federal do Paraná

**CURITIBA
2023**

Dedicamos este trabalho aos nossos familiares, amigos e todos que colaboraram.

RESUMO

O rápido crescimento da frota veicular mundial faz com que novos desafios sobre a condição do veículo sejam enfrentados pelas populações e organizações. Um dos principais, diz respeito à manutenção e às consequências que isto causa no desempenho de determinadas funções de um veículo. A necessidade de se monitorar o estado, reduzindo a indisponibilidade, encontra na tecnologia uma possibilidade de solução. Por este motivo a Internet de Veículos se torna uma tendência cada vez mais importante para o futuro. Porém, o processo que aplica essa tecnologia pode ser determinante para uma solução com grande impacto. Neste sentido, este estudo aborda os desafios do desenvolvimento de um modelo de sistema e sua prototipação que leve em conta cada etapa do fluxo, com o intuito de reduzir a indisponibilidade através de análises executadas nos dados capturados por sensores e atuadores internos do veículo. Após a realização da captura dos dados em um automóvel, em três diferentes cenários propostos, e uma análise posterior, os resultados obtidos indicaram que o modelo desenvolvido permite auxiliar na detecção de possíveis falhas, possibilitando levantar pontos de atenção no veículo durante a realização de experimento.

Palavras-chave: internet das coisas; big data; monitoramento; diagnóstico a bordo; internet de veículos.

ABSTRACT

The rapid growth of the worldwide vehicle fleet means that new challenges regarding the condition of the vehicle are faced by populations and organizations. One of the main concerns, is the maintenance and the consequences caused in the performance of certain functions. The necessity of monitoring the status, reducing the unavailability, finds in technology a possible solution. For this reason, the Internet-of-Vehicles (IoV) becomes an increasingly important trend for the future. However, the process that applies this technology can be decisive for a solution with a great impact. For that matter, this study addresses the challenges of developing a system model and its prototype that considers each stage in the workflow, with the aim of reducing unavailability through analyses performed on the data captured by internal vehicle sensors and actuators. After completing the data capture, under three different proposed scenarios, and a subsequent analysis, the results obtained indicated that the developed model assists in the detection of possible failures, allowing to raise points of attention in the vehicle during the experiment.

Keywords: internet of things; big data; monitoring; on-board diagnostic; internet of vehicles.

LISTA DE FIGURAS

Figura 1 – Arquitetura IoV em três camadas.	18
Figura 2 – Arquitetura IoV em quatro camadas.	19
Figura 3 – Arquitetura IoV em sete camadas.	20
Figura 4 – Estrutura de um OBD.	22
Figura 5 – Diferentes conectores OBD-I.	22
Figura 6 – Conectores ELM327.	24
Figura 7 – Pinagem conector ELM327.	25
Figura 8 – Relação entre falhas e tempo de vida.	27
Figura 9 – Tipos de manutenção	28
Figura 10 – Diagrama compressão	29
Figura 11 – Filtros da revisão sistemática.	34
Figura 12 – Arquitetura do modelo	44
Figura 13 – Conectores OBD estudados	46
Figura 14 – Fluxograma de seleção de veículo	47
Figura 15 – Rota cenário 1	51
Figura 16 – Rota cenário 2	52
Figura 17 – Rota cenário 3	52
Figura 18 – Esquema do banco local	57
Figura 19 – Relatório 1 gerado através do Apache Superset	70
Figura 20 – Relatório 2 gerado através do Apache Superset	71

LISTA DE GRÁFICOS

Gráfico 1 – Publicações ao longo dos anos	42
Gráfico 2 – Relação entre MAF, posição do acelerador e carga do motor no cenário 3	58
Gráfico 3 – Relação entre a temperatura de arrefecimento e velocidade no cenário 1 (Dia 25/08)	59
Gráfico 4 – Relação entre a temperatura de arrefecimento e velocidade no cenário 1 (Dia 02/09)	59
Gráfico 5 – Evolução no decorrer dos minutos de coleta da temperatura de arrefe- cimento em diferentes coletas realizadas no cenário 2	60

LISTA DE TABELAS

Tabela 1 – Revisão sistemática	35
Tabela 1 – Revisão sistemática	36
Tabela 1 – Revisão sistemática	37
Tabela 1 – Revisão sistemática	38
Tabela 1 – Revisão sistemática	39
Tabela 1 – Revisão sistemática	40
Tabela 1 – Revisão sistemática	41

LISTA DE QUADROS

Quadro 1 – Funções OBD-I	23
Quadro 2 – Funções OBD-II	23
Quadro 3 – Resultado preliminar da revisão sistemática	34
Quadro 4 – Veículos utilizados para coleta de dados	45
Quadro 5 – Compatibilidade de veículos e conectores	47
Quadro 6 – Sensores disponíveis para coleta	49
Quadro 7 – Sensores utilizados na coleta	50
Quadro 8 – Primeira fase do tratamento de dados	53
Quadro 9 – Segunda fase do tratamento de dados	54
Quadro 10 – Comparação entre métodos de compressão	55

LISTA DE ABREVIATURAS E SIGLAS

Siglas

BI	Business Intelligence
DPI	Deep Packet Inspection
DSRC	Dedicated short-range communication
ECU	Engine Control Unit
GB	Gigabyte
IoT	Internet-of-Things
IoV	Internet-of-Vehicles
LZMA	Lempel-Ziv-Markov chain Algorithm
MAF	Mass air flow
MTTF	Mean Time to Failure
OLTP	Online Transaction Processing
ROVDS	Remote On-line Vehicle Diagnostics System
SCP	Secure Copy Protocol
SQL	Structured Query Language
SSH	Secure Shell
UDS	Unified Diagnostic Services
USB	Universal Serial Bus
V2I	Vehicle to Infrastructure
V2P	Vehicle to pedestrian
V2R	Vehicle to Road
V2S	Vehicle to sensor
V2V	Vehicle to Vehicle

VNO Virtual Network Operator

WiFi Wireless Fidelity

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Objetivos	14
1.1.1	Objetivo geral	14
1.1.2	Objetivos específicos	15
1.2	Justificativa	15
1.3	Escopo	16
1.4	Estrutura do trabalho	16
2	LEVANTAMENTO BIBLIOGRÁFICO	17
2.1	Internet de Veículos (IoV)	17
2.1.1	Arquiteturas de IoV	18
2.2	Diagnóstico a bordo (OBD)	21
2.2.1	OBD-I	21
2.2.2	OBD-II	23
2.2.3	Adaptadores ELM327	23
2.3	Big data	25
2.4	Manutenção	26
2.4.1	Manutenção Corretiva	26
2.4.2	Manutenção Preventiva	27
2.4.3	Manutenção Preditiva	28
2.5	Compressão	29
2.6	Tratamento	31
2.7	Armazenamento	31
3	ESTADO DA ARTE	33
3.1	Revisão Sistemática	33
3.1.1	Exploração da Revisão	35
4	METODOLOGIA	44
4.1	Coleta	45
4.1.1	Sensores de coleta	48
4.1.2	Perfil do motorista	51
4.1.3	Cenários	51

4.2	Tratamento	53
4.3	Armazenamento	54
4.4	Compressão	54
4.5	Sincronização	55
4.6	Análise	56
4.7	Operação	56
5	DISCUSSÃO	58
5.1	Posição do acelerador, Massa de fluxo de ar e Carga do motor	58
5.2	Temperatura do líquido de arrefecimento e velocidade	59
5.3	Considerações sobre a análise	60
6	CONCLUSÃO	62
	REFERÊNCIAS	64
	APÊNDICES	68
	APÊNDICE A – RELATÓRIOS	70

1 INTRODUÇÃO

O carro deixou de ser apenas um artigo de luxo para se tornar uma peça necessária na sociedade moderna. O uso de veículos em todo o mundo aumentou drasticamente durante a última década. Segundo a Comissão Europeia teremos um aumento de 50% no transporte nos próximos 20 anos (PRYTZ, 2014). Mais de 84 milhões de automóveis foram fabricados no ano de 2012 segundo Roudier (2013). Esse rápido aumento de veículos e a necessidade das pessoas levou a muitas preocupações em questões como direção perigosa, más condições do veículo, más condições das estradas além de outros fatores contribuintes para potenciais acidentes e despesas. Além disso, a dinâmica de compra e venda do mercado demanda que as informações a respeito das condições do veículo e de sua manutenção estejam atualizadas, disponíveis, íntegras e fidedignas.

Realizar a manutenção de um veículo pode se tornar uma tarefa complicada, pois com a modernização dos automóveis eles acabam tornando-se estruturas complexas que envolvem tanto hardware, quanto software. Pensando nesses desafios, a indústria de veículos adotou alguns tipos de estratégias de manutenção, como: manutenção preditiva, manutenção corretiva e manutenção preventiva. Lie e Chun (1986) define a manutenção preventiva como qualquer manutenção que ocorre quando o sistema ainda não apresentou alguma falha. Segundo Shafi *et al.* (2018), este tipo de manutenção é usada para falhas pouco frequentes mas que possuem um reparo extremamente caro. Já a manutenção corretiva é quando o automóvel apresentou algum problema e ocasionalmente ocorre a reposição das peças defeituosas. Por último, a manutenção preditiva é onde as condições históricas e atuais do veículo são analisadas para procurar prever o que provavelmente irá falhar, sendo um subtipo da manutenção preventiva. A estratégia de manutenção mais comumente utilizada na indústria de veículos é normalmente corretiva, resultando em redução da vida útil do veículo e gastos mais elevados.

A manutenção preditiva, também referenciada como prognóstica, é uma estratégia para tentar solucionar estas questões, antecipando possíveis problemas futuros usando como base dados e situações passadas e atuais somadas com análises. A discussão detalhada sobre manutenção preditiva da indústria de veículos é apresentada por Prytz (2014), estudo este que também discute como as abordagens de aprendizado de máquina estão sendo usadas em uma indústria automotiva para a previsão de falhas. Marcorin e Lima (2003) destacam que com essa característica é possível a programação para intervenções, como aquisição de peças e reparos. Apesar de um aparente alto investimento, o autor afirma que essa atitude é plenamente recompensada pelos seus resultados, devido à vasta gama de possibilidades, ferramentas e métodos a serem aplicados, obtendo uma boa relação custo-benefício.

Este tipo de manutenção lida com os dados a bordo e existem sérios desafios quando se lida com a mesma. Os dados gerados por sensores a bordo não só possuem vários tipos, como também se apresentam em grandes quantidades produzindo assim milhares de sinais ou fluxos. Xu *et al.* (2018) indica que cada automóvel conta com mais de 200 sensores embutidos

desde 2020, com isso a quantidade de dados em um veículo pode chegar a 4 Terabytes de dados gerados a cada dia. Esses números mostram que os sistemas precisam de capacidade de armazenamento enorme, o que resulta em alto custo e possível lentidão. Ainda que tenha acontecido um aumento significativo na pesquisa sobre diagnóstico de veículos durante a última década, a implementação desses sistemas não acompanhou o mesmo ritmo.

Normalmente, qualquer subsistema compreende processos eletromecânicos, atuadores e sensores. No caso dos automóveis modernos, os sensores e atuadores de veículos são associados e controlados por uma *Engine Control Unit (ECU)*, responsável por gerenciar e monitorar o funcionamento do veículo. Dois protocolos utilizados para comunicação com o ECU são *On-Board Diagnostics (OBD)* em sua segunda versão (OBD2) e *Unified Diagnostic Services (UDS)* (Mostafa Anwar Taie; Diab; ElHelw, 2012). As estruturas de diagnóstico a bordo OBD fornecem dados momentâneos de diferentes partes do veículo, enquanto o UDS é um protocolo menos utilizado porém capaz de trazer mais detalhes gerais, sendo que este segundo método pode variar de montadora para montadora. Ambos os protocolos têm como objetivo de viabilizar o compartilhamento de informações entre o veículo e o motorista, as operações dos sensores e os veículos circundantes e a condição da estrada com os interessados (JOHANSON; DAHLE; SÖDERBERG, 2011). Nos últimos anos, a tecnologia Internet-of-Things (IoT) emergiu como uma importante área de pesquisa e aplicação com um enorme potencial de crescimento. Como um dos principais ramos da IoT, a *Internet-of-Vehicles (IoV)* atrai o interesse do mercado em pesquisas e produtos. E recentemente, o IoV baseado no armazenamento em nuvem se beneficiou do rápido desenvolvimento de redes móveis e tecnologias de *big data* (ZHANG *et al.*, 2017). Os veículos conectados e a tecnologia avançada permitem o funcionamento do *Remote On-line Vehicle Diagnostics System (ROVDS)* (BURGESS *et al.*, 2013). Essas tecnologias da informação expandem os desafios do ROVDS que envolve a aquisição de informações através da captura, processamento e armazenamento (Nkenyereye; Jang, 2017). Esta variedade de problemas reforça a necessidade de repensar cada etapa do processo, visto que há espaço para melhoria em cada uma delas.

1.1 Objetivos

Nesta seção são apontados os objetivos desta monografia, divididos em duas categorias distintas. A primeira, de forma a expor uma visão geral do objetivo. A segunda, indicando ponto a ponto os objetivos específicos para realização do trabalho.

1.1.1 Objetivo geral

Dado o cenário descrito, este trabalho de conclusão de curso procura apontar se a especificação de um modelo aplicado na construção, desenvolvimento, prototipação e validação de

um sistema de informação que agregue coleta, tratamento, armazenamento e análise de dados de veículos, poderia diminuir a indisponibilidade e quebra dos automóveis, prevenindo necessidades de reparos.

1.1.2 Objetivos específicos

Os objetivos específicos desse trabalho são:

- Levantar os problemas, dificuldades e os requisitos de captura e armazenamento de dados de veículos encontrados em estudos semelhantes e também os identificados ao decorrer deste estudo.
- Projetar um modelo de arquitetura de sistemas que visa solucionar ou minimizar os problemas levantados para a coleta, armazenamento, transmissão, tratamento e análise dos dados.
- Validar se o modelo proposto cumpre os requisitos levantados.
- Prototipar uma implementação do modelo proposto com integração de componentes que faça a exibição de dados analisados, gere relatórios, realize a comunicação via Bluetooth, se conecte via obd2 com o veículo e armazene os dados localmente ou em nuvem, dependendo da disponibilidade de rede no momento.

1.2 Justificativa

Dado que o veículo faz parte do dia a dia de grande parte da população para a realização da maioria de suas tarefas, a indisponibilidade do mesmo se torna um problema, pois pode acarretar em um custo elevado de tempo e dinheiro. Deixar de fazer a manutenção necessária pode levar a falhas mecânicas e causar perda total do mesmo. Além disso, grande parte dos automóveis produzidos, principalmente os populares, podem não ter alguns indicadores de medições importantes no painel, como temperatura do motor e outros.

Existem aplicações no mercado que realizam a coleta e exibição dos dados, porém a indisponibilidade de rede, de armazenamento dos dados, falta de tratamento dos mesmos e uma arquitetura planejada de forma errada acaba limitando o uso destas soluções. Portanto, uma especificação que leve em consideração cada fase, tanto do desenvolvimento, quanto do problema em si, pode aumentar a probabilidade de sucesso total ou acarretar em melhorias significativas nas resoluções de cada fragmento do problema, comparado às soluções existentes, que em sua maioria visam somente o resultado final.

Soluções plenamente otimizadas em todos os aspectos são de difícil obtenção e requerem uma carga alta de estudo, já que existem tanto limitações técnicas, quanto limitações de

tempo e recursos. Por este motivo, se torna válido o estudo de diversas abordagens de implementação quando buscamos a solução de um problema complexo, com o intuito de se obter um sistema que resolva o problema que foi proposto e apresente justificativas para a tomada de decisões.

1.3 Escopo

Este estudo busca, portanto, verificar se utilizando um modelo estruturado é possível diminuir a indisponibilidade do automóvel por motivos de manutenção corretiva. Ou seja, desenvolvendo uma pesquisa em campo, fazendo uso de coleta de dados e um nível de prototipagem que possibilite ao motorista visualizar as informações de maneira mais intuitiva sobre o comportamento do veículo em situações específicas, identificando possíveis padrões, através exclusivamente da observação dos gráficos e relatórios extraídos, sem utilizar de técnicas de inteligência artificial, além de alterações no comportamento dos valores lidos pelos sensores dentro de um determinado mesmo caso de uso.

1.4 Estrutura do trabalho

No capítulo dois deste trabalho é apresentado um levantamento teórico dos principais conceitos envolvidos neste estudo, expondo algumas definições de autores que trabalham com o tema. No mesmo capítulo é tratado sobre Diagnóstico a bordo, *Big data*, *Business Intelligence (BI)*, Manutenção, Compressão, Internet das Coisas e suas possíveis arquiteturas.

O capítulo três apresenta o estado da arte, juntamente com a revisão sistemática, que busca verificar soluções já existentes para problemas semelhantes ao proposto por este trabalho, assim como um estudo a respeito dos modelos de soluções propostos na literatura.

O capítulo quatro descreve a metodologia e a organização, utilizada no desenvolvimento deste trabalho, assim como os materiais e métodos utilizados para a realização do mesmo.

No capítulo cinco encontra-se a discussão realizada sobre os resultados obtidos após a realização do experimento e em seguida, no sexto e último capítulo, apresenta-se a conclusão obtida do estudo, assim como as considerações finais dos autores em relação à pesquisa detalhada em todos os capítulos.

2 LEVANTAMENTO BIBLIOGRÁFICO

Neste capítulo são apresentados alguns conceitos que servem como base para o desenvolvimento deste trabalho, tais como: Internet de Veículos (IoV), Diagnóstico a bordo (OBD) e *Big data*. Este estudo é importante para que fiquem claras as definições e diferentes conceitos, em torno dos temas abordados.

2.1 Internet de Veículos (IoV)

Veículos com sensores incorporados geram grandes quantidades de dados a cada instante, segundo Gerla *et al.* (2014), e com o aumento do número de veículos conectados à Internet das Coisas (IoT) as Redes Ad-hoc de Veículos estão mudando para a Internet de Veículos (IoV) (Yang *et al.*, 2014). Como um membro indispensável da IoT, a Internet de Veículos desempenha um papel importante no sistema de transporte inteligente e pode promover o desenvolvimento de cidades inteligentes, pois espera-se que ela ofereça uma ampla variedade de aplicações para melhorar a segurança nas estradas, a eficiência do tráfego e a entrega de informações (Huang; Lu; Choo, 2017).

Projetar comunicações eficientes com abordagens de coleta de dados é uma necessidade extrema neste momento. Alguns desafios para este contexto consistem em: localização de veículos, posição do mesmo, manter e interpretar uma visão exata de toda a rede, um grande número de nós, mudanças rápidas mobilidade dos nós, rápidas mudanças topológicas e frequentes desconexões de rede adicionam desafios potenciais na área de VANETs (GILLANI *et al.*, 2022).

Segundo Xu *et al.* (2018), a relação entre o *big data* e o IoV é de benefícios para os dois lados. O autor afirma que o IoV é necessário para oferecer suporte à aquisição, transmissão, armazenamento e computação de *big data* para proprietários e usuários de dados. Em contrapartida, o IoV pode se beneficiar de *big data*, utilizando a mineração dos dados disponíveis, que incluem grandes informações de escopo do ambiente interno e externo de veículos.

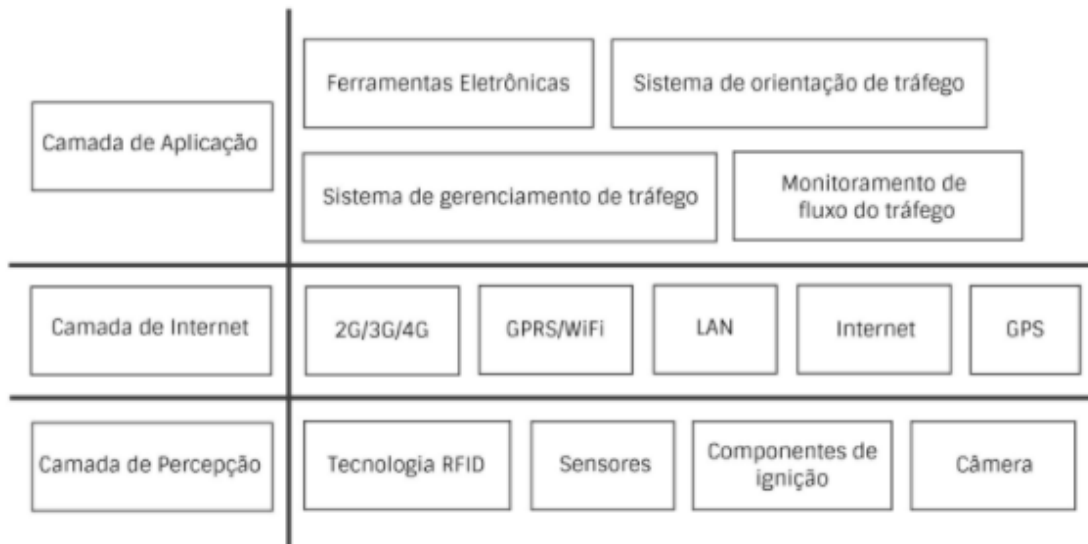
Devido a estes fatos, dentro das definições de IoV, existem divergências nas pesquisas sobre qual arquitetura melhor define a estrutura de uma rede de veículos conectados para manipulação destes dados. Essa arquitetura de rede heterogênea pode ser categorizada em cinco tipos de comunicação veicular: *Vehicle to Vehicle (V2V)*, *Vehicle to Infrastructure (V2I)*, *Vehicle to Road (V2R)*, *Vehicle to sensor (V2S)* e *Vehicle to pedestrian (V2P)* (ISLAM; HOSSAN; JANG, 2018). Em estudos mais recentes, mais dois conceitos foram introduzidos neste escopo, são eles: *Infrastructure-to-Infrastructure (I2I)* e *Vehicle-to-broadband cloud (V2B)* que representam a comunicação da infraestrutura presente na rede de veículos e a comunicação do veículo com a rede móvel em nuvem, como por exemplo 3G ou 4G (YEFERNY; HAMAD, 2020).

2.1.1 Arquiteturas de IoV

Pode haver várias soluções para o projeto de um sistema de IoV. Por exemplo, a arquitetura *Online Transaction Processing (OLTP)* tradicional, vista em Zhang *et al.* (2017), usa um banco de dados maduro e estável como um centro para implantar serviços. Em tal arquitetura, o sistema coleta dados carregados por veículos e os armazena no banco de dados, além de fornecer serviços com base na capacidade analítica do subsistema de banco de dados.

Na figura 1 é possível observar a arquitetura de três níveis, definida por Sadiku, Tembely e Musa (2018), sendo o primeiro nível, a camada de percepção, formada por todos os sensores dentro do veículo, responsáveis pela coleta das condições ambientais e eventos específicos, como, por exemplo, a situação do veículo. O segundo nível seria a camada de comunicação ou internet, onde temos comunicação sem fio, V2I, V2P e V2S. Ela também fornece aos carros o processamento baseado em *big data* e é responsável pelo armazenamento, análise, processamento e tomada de decisões sobre diferentes situações de risco. Sadiku, Tembely e Musa (2018) explicam também que na terceira camada (Camada de Aplicação) estão as ferramentas estatísticas, suporte para armazenamento e infraestrutura de processamento que constitui a inteligência de IoV e fornece aos carros o processamento, armazenamento, análise e tomada de decisões sobre diferentes situações de risco.

Figura 1 – Arquitetura IoV em três camadas.



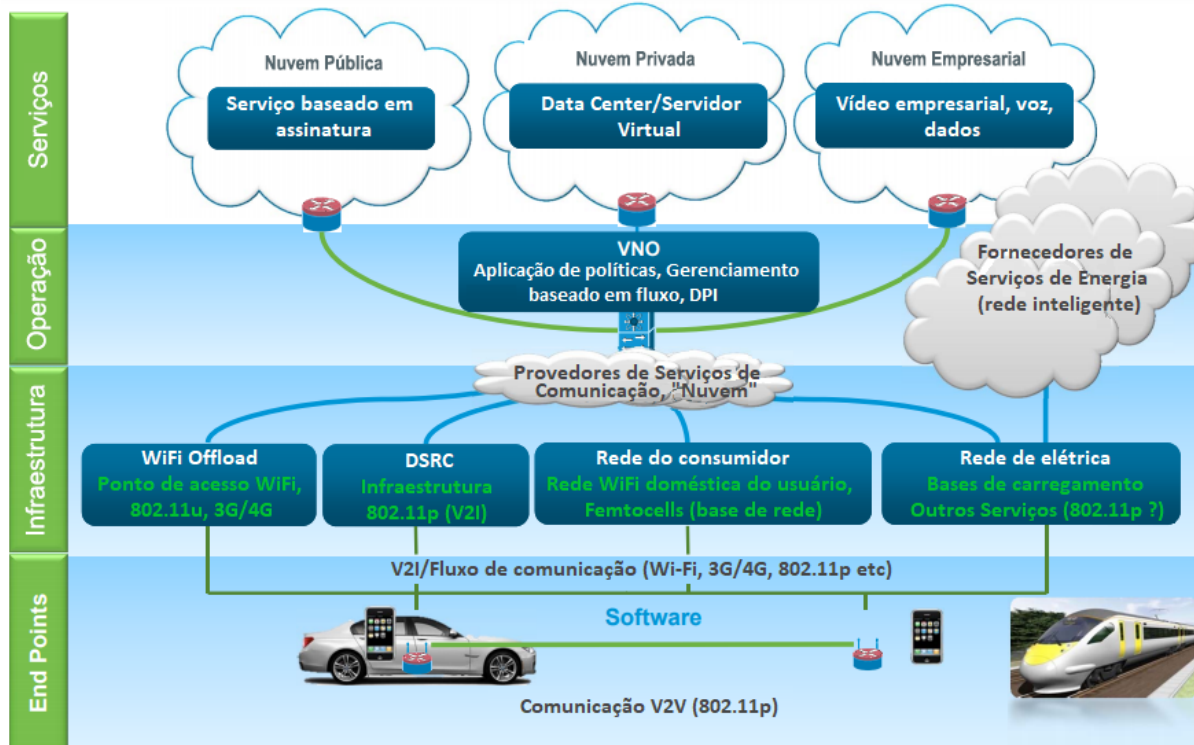
Fonte: Adaptado de (SADIKU; TEMBELY; MUSA, 2018).

Em contraste à arquitetura apresentada anteriormente, Bonomi (2013), representando a Cisco Systems, Inc. (1984), propôs uma arquitetura de IoV baseada em quatro camadas, ilustrada na figura 2.

A arquitetura apresentada, possui foco em três tipos de nuvens diferentes (empresarial, pública, privada), e tem no seu ponto de serviço uma *Virtual Network Operator (VNO)*, respon-

sável por aumentar a segurança utilizando técnicas como *Deep Packet Inspection (DPI)*. Ela se ramifica em diversas infraestruturas, como: pontos de acessos de 3G/4G, *Dedicated short-range communication (DSRC)*, chegando até seus pontos finais, como os veículos ou dispositivos móveis.

Figura 2 – Arquitetura IoV em quatro camadas.



Fonte: Adaptado de (BONOMI, 2013).

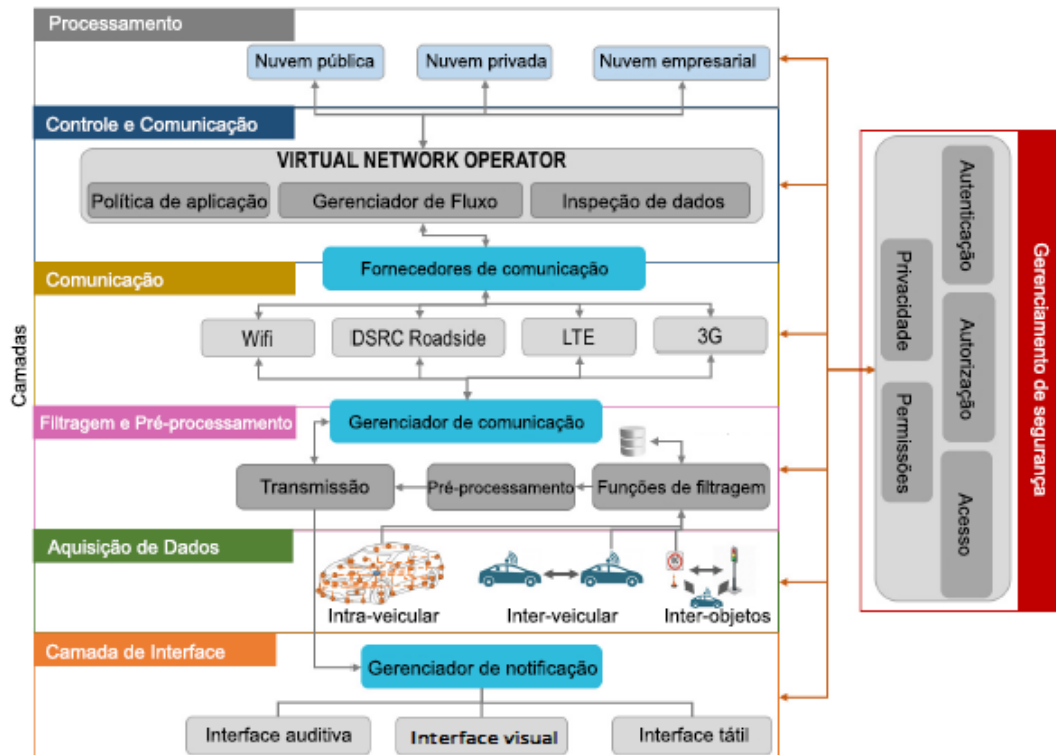
A camada de *end-point* abrange os veículos, o software e a comunicação V2V. A segunda camada é a infraestrutura, que assim como o segundo nível da primeira arquitetura apresentada, define todas as tecnologias que permitem conexões entre todos os atores da IoV (Wireless Fidelity (WiFi), 3G, 4G). E é na terceira camada em que pode-se ver a maior diferença entre as arquiteturas até aqui apresentadas. Segundo Bonomi (2013), a camada de operação monitora a aplicação da política e o gerenciamento baseado em fluxo. E por fim, a camada de serviço especifica os serviços que os diferentes tipos de nuvem (nuvem pública, nuvem privada e nuvem corporativa) oferecem baseados no *data center* ou na demanda.

Foi analisando essas duas arquiteturas que Contreras-Castillo, Zeadally e Guerrero-Ibañez (2018) propuseram uma nova arquitetura para IoV, baseada em uma estrutura de sete camadas e apresentada na figura 3.

Segundo o autor, as arquiteturas apresentadas anteriormente possuem problemas em relação à segurança, como a falta de autenticação e também não apresentam uma camada onde seja possível selecionar a melhor maneira de se transmitir os dados, para realizar a conexão com os serviços. Além disso, os modelos limitam a maneira com a qual as notificações são apresentadas aos condutores e por fim, o último problema diz respeito à falta de pré-

processamento das informações captadas, podendo levar a uma sobrecarga do sistema. As camadas apresentadas na figura 3 são as seguintes:

Figura 3 – Arquitetura IoV em sete camadas.



Fonte: Adaptado de (Contreras-Castillo; Zeadally; Guerrero-Ibañez, 2018).

Primeira camada (interface do usuário): é a responsável pela comunicação direta com o motorista, coordenando as notificações recebidas por ele, selecionando a maneira mais adequada para a situação em que o veículo e motorista se encontram no momento.

Segunda camada (aquisição de dados): sua tarefa é a coleta dos dados das diversas formas disponíveis pelo veículo (sensores, outros veículos e emissores de sinais na rodovia).

Terceira camada (pré-processamento e filtragem de dados): esta camada evita que informações desnecessárias sejam armazenadas, o que tornaria o volume de dados ainda maior e mais complexo.

Quarta camada (comunicação): é a responsável por selecionar a melhor forma de transmissão dos dados, levando em conta por exemplo o alcance das redes (WiFi, 3G) e até mesmo a relevância da informação e o nível de segurança necessário.

Quinta camada (controle e gerenciamento): camada onde as políticas registradas dos diversos atores da IoV são aplicadas. Ou seja, para melhor gerenciamento dos dados coletados, é necessário observar, por exemplo, as regras de gerenciamento de tráfego.

Sexta camada (processamento): onde é realizado o processamento das informações recebidas tanto localmente quanto remotamente, sendo essencial para utilizações futuras em aplicações por exemplo.

Sétima camada (segurança): se comunica diretamente com todas as outras camadas, sendo atribuída a tarefa de garantir a segurança por meio de autenticação e controle de acesso no decorrer de toda arquitetura.

2.2 Diagnóstico a bordo (OBD)

A sigla OBD (*On-Board Diagnostic*) designa um sistema de autodiagnóstico disponível na maioria dos veículos automotores que circulam atualmente. A ligação ao sistema ocorre por meio de um conector padronizado que foi sancionado como obrigatório na Europa e nos Estados Unidos para todos os veículos produzidos a partir de 1996 e no Brasil, a partir de 2010 com o padrão de segunda geração OBD-II (SANTOS, 2016). Esse sistema foi inicialmente desenvolvido para o controle da emissão de poluentes pelos veículos, porém com o passar dos anos começou a surgir o interesse pela padronização dos protocolos de comunicação utilizados pelos veículos e então a redução dos custos de manutenção eletrônica.

Segundo Santos (2016), a leitura dos dados de diagnóstico permite identificar vários tipos de anomalias, desde as mais simples como uma lâmpada queimada até problemas maiores como problema no sistema de injeção eletrônica. Outro fator importantíssimo, explicado com mais detalhes nas seções seguintes, foi a padronização posterior sofrida pelo OBD. Esse acontecimento tornou o acesso aos equipamentos para diagnóstico mais fáceis por existir um padrão nos conectores, mais simples pois poderia ser sincronizado via Bluetooth com computador ou aparelho celular e além disso, possibilitou que o proprietário do veículo, com um menor nível de conhecimento sobre mecânica, pudesse chegar ao mecânico com uma suposição sobre o problema, diminuindo o risco de ser enganado na realização da manutenção.

Um sistema OBD básico ilustrado na figura 4, consiste em uma Unidade de Controle Eletrônico (*Engine Control Unit* - ECU) que utiliza a entrada de vários sensores para manipular atuadores, como sensores de oxigênio controlando injetores de combustível (OBD Solutions, 2017).

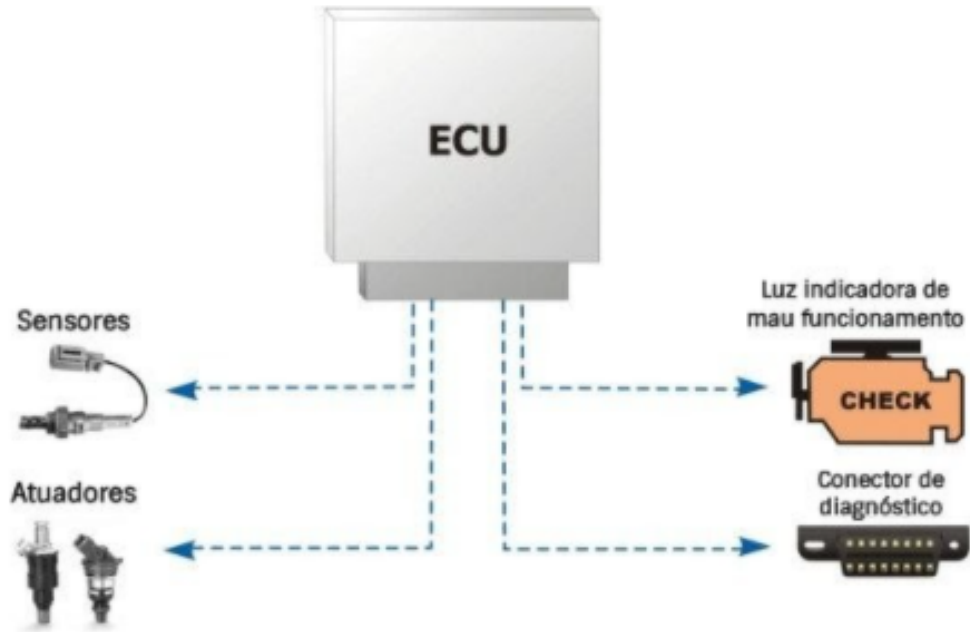
Atualmente o sistema OBD possui duas versões, sendo elas chamadas de OBD-I e OBD-II.

2.2.1 OBD-I

A primeira versão do sistema possuía um grande problema de incompatibilidade de leitores, devido à falta de padronização das montadoras nos seus conectores. As fabricantes utilizavam diferentes tipos de conectores e códigos de erros (DTC, ou *Diagnostic Trouble Codes*). Na figura 5 podemos observar o grande número de conectores existentes nesta versão.

Através da conexão via OBD-I era possível obter os seguintes parâmetros da ECU do veículo, conforme apresentado no quadro 1:

Figura 4 – Estrutura de um OBD.



Fonte: Adaptado de (OBD Solutions, 2017).

Figura 5 – Diferentes conectores OBD-I.



Fonte: Adaptado de (MCCORD, 2011).

Quadro 1 – Funções OBD-I

OBD I
Sensor de oxigênio
Sistema de recirculação de gases
Sistema de combustível
Componentes elétricos
Sistemas eletrônicos
Informação de diagnóstico
Códigos de erros

2.2.2 OBD-II

Na segunda versão, além da solução destas incompatibilidades citadas acima, foi também proposta a padronização do conector J1962 (SAE, 1992). Desta maneira, poderia ser desenvolvido então um conector padrão para leitura de dados de veículos de diferentes fabricantes.

Além destas alterações de padronização, o número de sensores que possuíam leitura via o sistema OBD mais do que dobrou (de 7 para 15), as novas funções são apresentadas no quadro 2.

Quadro 2 – Funções OBD-II

OBD II
Sensor de oxigênio
Sistema de recirculação de gases
Sistema de combustível
Componentes elétricos
Sistemas eletrônicos
Informação de diagnóstico
Códigos de erros
Eficiência do catalisador
Aquecimento do catalisador
Combustão
Sistema de arrefecimento
Sistema de ar secundário
Parâmetros do motor
Memorização de avarias
Padronização das ligações

2.2.3 Adaptadores ELM327

Os adaptadores ELM327 são utilizados para realizar a conexão da porta OBD-II de um veículo com outros dispositivos, como por exemplo computadores ou celulares. Fabricados

pela empresa Elm Electronics (1998), atualmente existem vários tipos de adaptadores ELM327, como é possível observar na figura 6, podendo eles transmitir dados de diversas formas: via WiFi que tem um custo maior para se adquirir e também um gasto maior de energia, porém uma maior segurança comparado ao Bluetooth; o adaptador Bluetooth, que é o encontrado em maior número no mercado e com preço menor; via porta USB ou Serial.

Figura 6 – Conectores ELM327.



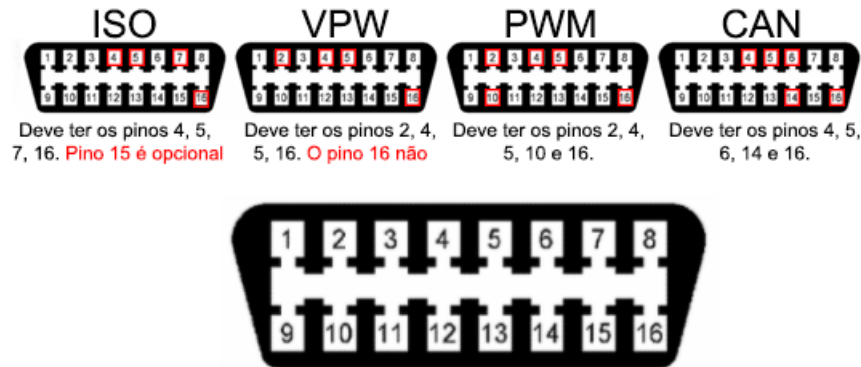
Fonte: (OBD Facilities, 2010).

Conforme Santos (2016), a regulamentação ainda é fraca em relação a esse assunto, obrigando apenas a disponibilização de dados relacionados à emissão de poluentes. A maioria dos fabricantes de veículos não oferecem todos os dados e este é o grande problema envolvendo a obtenção de dados via conector ELM327. Como apresentado na figura 7, os adaptadores possuem suporte aos diversos protocolos existentes, definidos por Bastos (2012) como cinco:

- a) SAE J1850 PWM: possui a taxa de transferência de 41,6 Kbps, utiliza 2 pinos do conector OBD2 e o tamanho da sua mensagem é de 12 bytes;
- b) SAE J1850 VPW: possui um único fio de conexão, por isso é considerado de baixo custo. Taxa de transferência de 10,4 Kbps e tamanho para mensagens de 12 bytes;
- c) ISO 1941-2: a comunicação é assíncrona *Universal Asynchronous Receiver/Transmitter* (UART), taxa de transmissão 10,4 Kbps e mensagem de 5 a 11 bytes;
- d) ISO 14230: popularmente chamado de Keyword 2000 (KW2000), é apenas um link de diagnóstico e não pode ser utilizado para transmitir mensagens. Foi bastante utilizado antes mesmo da OBD2 por diversos fabricantes europeus que pressionavam órgãos regulamentadores para viabilizar a utilização desse link. Sua liberação ocorreu na década de 1990, e só foi aprovada por ter o protocolo e os requerimentos de hardware quase idênticos à ISO 1941-2. Este protocolo possui a velocidade entre 1,2 e 10,4 Kbps e suas mensagens podem conter até 255 bytes;

e) ISO 15765 *Controller Area Network* (CAN): sua comunicação é serial e sua velocidade é de 500 Kbps. O padrão ISO 15765-4 determina os requisitos para a aplicação OBD.

Figura 7 – Pinagem conector ELM327.



Pino	Descrição
1	Fabricante
2	Bus positive Line of SAE-J1850
3	Fabricante
4	Chassis ground
5	Signal ground
6	CAN high (ISO 15765-4 and SAE-J2234)
7	K line of ISO 9141-2 and ISO 14230-4
8	Fabricante
9	Fabricante
10	Bus negative Line of SAE-J1850
11	Fabricante
12	Fabricante
13	Fabricante
14	CAN low (ISO 15765-4 and SAE-J2234)
15	L line of ISO 9141-2 and ISO 14230-4
16	Battery voltage

Fonte: Adaptado de (MACHADO; OLIVEIRA, 2007).

2.3 Big data

Veículos mais modernos, celulares com mais funcionalidades, câmeras mais potentes, tablets, aplicações e serviços são alguns exemplos da evolução tecnológica. Estas conquistas só foram possíveis pela digitalização dos dados por eles suportados. Estas tecnologias também abrem espaço para a geração de dados que no passado ainda não eram possíveis de serem captados. Lycett (2013) nomeia este fenômeno dos novos dados disponíveis graças às tecnologias de “o processo de fazer sentido”.

Para manipular esses dados de diversas fontes e quantidades elevadas, normalmente se utiliza o termo “*big data*”. A definição de *big data* não é nova e no decorrer do tempo sofre diversas alterações de significância como aponta o estudo realizado por Ylijoki e Porras (2016). Uma das definições mais conhecidas é a de Laney (2001) que descreve as três dimensões do

big data: volume, velocidade e variedade. Volume refere-se a quantidades cada vez maiores de dados. Velocidade mostra a necessidade destes dados terem seu fluxo realizado de forma rápida tendendo a tempo real ou o mais próximo possível. Já a variedade indica a capacidade de lidar com diversos tipos de dados, estruturados ou não, textos, imagens, vídeos, sensores são alguns exemplos da diversidade dos tipos de dados.

Sistemas de diagnóstico veicular online remoto (ROVDS) podem armazenar 40 TB de status e condições em um ano. Já no caso de veículos conectados reportarem o status de códigos de diagnóstico é possível gerar 25 Gigabyte (GB) de dados diariamente para cada veículo (Cregger; Brugeman; Wallace, 2012). Além do volume, a necessidade que a coleta e análise sejam executadas no menor tempo possível, com o objetivo de preservar as condições ideais do carro, somada com a pluralidade dos inúmeros tipos de dados, tornam os ROVDS bons exemplos de *big data*.

2.4 Manutenção

Nos mais diversos ramos da indústria, a manutenção é um ponto importante quando se trata do valor de um produto. Dependendo do setor da indústria, as despesas com manutenção, considerando o custo total de produção de um produto, podem representar entre 15 e 60% deste valor (MOBLEY, 2002). Quando falamos do setor automotivo não é diferente, segundo a consultora de pesquisa de mercado Future Market Insights, estima-se que os gastos com manutenção automotiva aumentem de US\$441,3 bilhões estimado atualmente, para US\$750 bilhões até 2026. Atualmente são 3 os principais tipos de manutenção existentes no mercado, relacionado a automóveis, sendo as mesmas descritas detalhadamente nas subseções seguintes, onde apresentamos suas aplicações, vantagens e desvantagens na utilização (INSIGHT, 2021).

2.4.1 Manutenção Corretiva

Este tipo de manutenção tem como propósito corrigir um problema identificado. Entre os três tipos apresentados, este possui o maior grau de dificuldade e também de forma geral os custos mais elevados, já que além do custo de reparo do equipamento, uma falha pode acarretar em outros danos, em outros componentes. Apesar destes fatos, é um tipo de manutenção popularmente adotado até os dias atuais, já que existem sim casos em que sua implementação pode ser mais vantajosa sob as outras apresentadas, desde que seja feita de forma consciente com suas possíveis consequências. É a estratégia conhecida como “*run-to-failure*”, onde só é realizada a troca de alguma peça caso a mesma venha a interromper seu funcionamento (GAMA; RIBEIRO; VELOSO, 2022). Tratando-se de veículos, é uma abordagem menos interes-

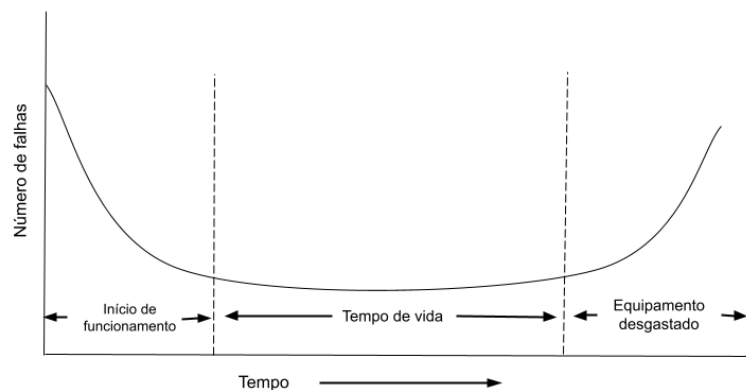
sante pois além dos possíveis transtornos que pode ocasionar com uma possível falha em um momento inesperado, pode até mesmo colocar em risco a integridade física do motorista.

2.4.2 Manutenção Preventiva

A manutenção preventiva, busca realizar o reparo antes que o problema aconteça, a fim de evitar danos mais significativos ao equipamento. De forma geral, a principal diferença entre a preventiva e preditiva está no fato que a preventiva tem como base um tempo limite de operação de um equipamento para realização da sua manutenção e não é baseada em análises de dados e monitoramento.

Um conceito importante, tratando-se de manutenção preventiva, é o *Mean Time to Failure (MTTF)*, que busca relacionar o número de falhas com o tempo de operação de uma máquina. A partir da figura 8, podemos identificar que nos instantes iniciais de operação de um equipamento, a chance de falhas é maior, sendo explicado por possíveis problemas de instalação, depois estabilizando e novamente retornando a possuir números mais altos, perto do seu prazo de validade (MOBLEY, 2002). Nos veículos, esse processo pode acontecer da mesma maneira, onde erros na instalação e manutenção podem ocorrer nos instantes iniciais. Caso não ocorram elas têm seu ciclo normal de desgaste, até o momento de uma nova verificação. As vantagens desse tipo de manutenção são que podem evitar um problema repentino, muitas vezes catastróficos que possam ocorrer sem nenhum tipo de alerta, ou seja, que não seriam identificados através de monitoramento, por exemplo. As principais dificuldades são realizar esse agendamento periódico de forma a possuir segurança e um baixo custo com manutenção, já que muitas vezes pode ocorrer uma falha durante os intervalos de tempo programados, ou a pausa do equipamento e também a troca de um componente de forma desnecessária (HAO *et al.*, 2010).

Figura 8 – Relação entre falhas e tempo de vida.



Fonte: Adaptado de Mobley (2002).

2.4.3 Manutenção Preditiva

A manutenção preditiva consiste na análise de dados, baseada no monitoramento constante das condições mecânicas, da eficiência e também outros indicadores de um equipamento, para procurar apontar o melhor momento para a parada do mesmo para realização da manutenção. Este tipo de manutenção é muito semelhante à manutenção preventiva, porém busca estimar o maior intervalo possível entre uma manutenção e outra. Dessa forma, evitando momentos de indisponibilidade para um possível reparo que possa ser desnecessário, otimizando tempo e custos com manutenção.

A manutenção preditiva também é apontada por muitos autores como uma maneira de otimização, considerando por exemplo uma linha de produção e os gastos com manutenção dentro desta linha para produção de uma mercadoria. A adoção da manutenção preditiva pode diminuir o custo final do produto, já que aumenta a disponibilidade das máquinas.

Uma visão geral entre as três formas de manutenção pode ser observada a partir da figura 9.

Figura 9 – Tipos de manutenção

Corretiva	Preventiva	Preditiva
Utilizar até falhar	Manutenção rotineira antes da falha	Não conserte se não vai ocorrer a falha
<ul style="list-style-type: none"> + Menor número de manutenções + Sem investimento inicial + Baixo custo - Alto risco de falha secundária - Não programado e longo tempo de indisponibilidade - Alto custo de peças de reposição - Alto custo de tempo e mão de obra - Perigoso 	<ul style="list-style-type: none"> + Manutenção controlada + Menores chances de acidentes + Peças e custos podem ser administrados + Menos falhas inesperadas - Manutenção sem ocorrer falhas - Impossibilidade de evitar quebras inesperadas 	<ul style="list-style-type: none"> + Menos quebras inesperadas + Peças são compradas quando há necessidade + Manutenção realizada quando necessário + Maior tempo entre manutenções - Maior investimento inicial - Requer maiores conhecimentos - Maior custo de realização - Necessário mais ferramentas para ser mantido

Fonte: Adaptado de Hao *et al.* (2010).

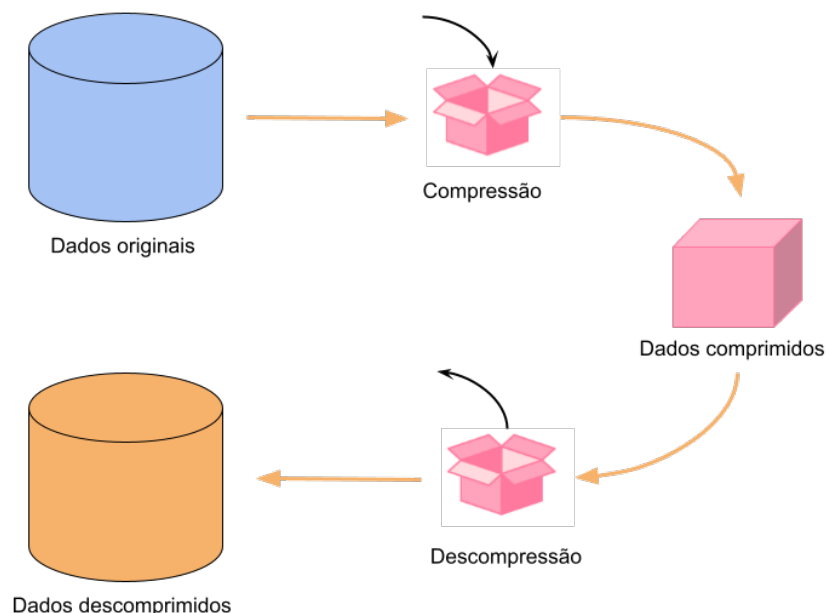
2.5 Compressão

Compressão é o processo de eliminar redundâncias para poupar a capacidade de armazenamento e transmissão de dados (SUARJAYA, 2012). É também interpretado como a capacidade de reduzir um arquivo de maior tamanho em um de tamanho reduzido (??).

Para suportar os constantes avanços da tecnologia e os consequentes saltos na quantidade de dados gerados a cada segundo, a capacidade de armazenamento e transmissão necessitam de melhorias. Na tentativa de superar este desafio, um conceito chamado compressão de dados (CD) foi apresentado na literatura. Introduzida por Samuel Morse em 1838 a estratégia de comprimir as letras em telégrafos serviu de inspiração para mais tarde a implementação do Algoritmo de Huffman (SALOMON; MOTTA, 2009).

Em geral, a compressão é executada a partir de dados redundantes e irrelevantes. Os tipos de compressão são separados em: sem perdas (*lossless*) e com perdas (*lossy*). A compressão sem perdas é o processo de converter os dados originais em uma versão mais concisa sem perder nenhum dado durante seu processo, já a compressão com perdas admite que alguns dados se percam no processo de descompressão, normalmente é tolerável para o utilizador esta mudança no conjunto original (SUARJAYA, 2012). Um fluxo padrão de compressão e descompressão pode ser visto na figura 10.

Figura 10 – Diagrama compressão



Fonte: Autoria própria (2023).

De acordo com um estudo realizado por Correa, Pinto e Montez (2022) sobre compressão de dados em sistemas de sensores de veículos, a compressão eficiente dos dados permite

reduzir o tamanho dos pacotes de dados transmitidos pela rede, resultando em uma economia significativa de largura de banda.

Ao comprimir os dados coletados pelos sensores, é possível reduzir o volume total de informações a serem transmitidas ou armazenadas. Essa redução de tamanho é especialmente valiosa quando se lida com grandes volumes de dados provenientes de sensores em veículos, pois reduz a sobrecarga da rede e otimiza a eficiência do armazenamento. O estudo de Kimura e Latifi (2005) sobre compressão de dados em sistemas de sensores sem fio ressalta que a compressão eficiente dos dados permite armazenar mais informações em um espaço limitado, maximizando a capacidade de armazenamento disponível. A utilização da compressão de dados não apenas economiza largura de banda e espaço de armazenamento, mas também contribui para a eficiência geral do sistema de sensores veiculares. Ao reduzir o tamanho dos pacotes de dados, também diminui o tempo necessário para transmitir ou processar as informações, melhorando a latência e o desempenho do sistema como um todo.

O método de compressão BZip2 é um algoritmo amplamente utilizado para compactar dados, especialmente aqueles que possuem padrões repetitivos e longas sequências de zeros. Ele utiliza uma combinação de técnicas de codificação de Huffman e transformação de Burrows-Wheeler para alcançar altas taxas de compressão (BURROWS; WHEELER, 1994).

LZMA2 busca identificar e eliminar redundâncias nos dados, resultando em uma redução significativa no tamanho dos arquivos. O método de compressão LZMA2 é uma variação aprimorada do algoritmo Lempel-Ziv-Markov chain Algorithm (LZMA). O funcionamento do LZMA2 envolve duas etapas principais: a fase de codificação e a fase de decodificação. Na fase de codificação, o algoritmo utiliza um dicionário para criar um modelo de contexto adaptativo, que é atualizado à medida que os dados são processados. O modelo de contexto é usado para prever o próximo símbolo a ser codificado com base no contexto atual e nas informações armazenadas no dicionário.

Durante a codificação, o LZMA2 faz uso de uma combinação de técnicas, incluindo a compressão LZ77 (que identifica sequências repetitivas de dados) e a codificação de Huffman (que atribui códigos binários de tamanho variável aos símbolos mais frequentes). Além disso, o LZMA2 utiliza uma etapa adicional chamada "reordenamento de bytes" para melhorar ainda mais a compressão, rearranjando os dados de forma mais eficiente (SCHOLL; LAERHOVEN, 2016).

O método de compressão Gzip é um algoritmo comumente utilizado para compactar dados em formato de arquivo, com o objetivo de reduzir seu tamanho e economizar espaço de armazenamento. O Gzip utiliza uma combinação de compressão de dicionário e codificação de Huffman para atingir altas taxas de compressão. Durante o processo de compressão, o Gzip cria um dicionário de padrões de repetição encontrados nos dados e substitui esses padrões por códigos mais curtos, resultando em uma representação mais compacta dos dados. Em seguida, a codificação de Huffman é aplicada para atribuir códigos binários de tamanho variável

aos símbolos mais frequentes, o que contribui para uma compressão adicional dos dados (BERZ *et al.*, 2015).

2.6 Tratamento

Para tornar os dados coletados dos sensores de um carro mais adequados para análise, muitas vezes é necessário passar por várias técnicas de pré-processamento. Uma das abordagens mais utilizadas para este problema é a limpeza de dados, mais conhecida pelo termo em inglês Data Cleaning, que possui como objetivo principal detectar e remover erros e inconsistências dos dados com o intuito de aumentar a qualidade dos dados. Rahm e Do (2000) apresentam um extenso trabalho a respeito da crescente necessidade de limpeza dos dados com o intuito de obter uma base de dados precisa e consistente, tornando possível através de múltiplas fontes de dados, consolidar uma representação fidedigna e eliminar adversidades provenientes de dados não correspondentes a situações analisadas. Por exemplo, de acordo para a pesquisa “Price Waterhouse Coopers” realizada em 2001, 75% das 599 empresas sofreram perdas devido a problemas de qualidade, visto que os processos de análises de dados confiam na integridade das suas coletas, para que suas investigações através de aprendizado de máquina ou métodos estatísticos, gerem resultados sólidos e confiáveis. (LI, 2012). Este ponto vai de encontro com o estudo realizado por Rao *et al.* (2012), onde afirmam que dados de qualidade só podem ser produzidos através da limpeza dos dados coletados a partir de diversas fontes, já que no momento da extração se encontram sujos e não íntegros.

2.7 Armazenamento

O armazenamento de dados é um aspecto fundamental no contexto dos sensores veiculares, sendo crucial garantir que as informações coletadas sejam armazenadas de maneira eficiente e escalável. De acordo com um estudo realizado por Sharma (2021) sobre sistemas de armazenamento para veículos autônomos, é essencial ter uma base de dados escalável para lidar com o grande volume de dados gerados pelos sensores. A capacidade de escalabilidade permite que o sistema de armazenamento se adapte ao aumento da quantidade de dados ao longo do tempo, garantindo que a informação seja armazenada de forma confiável e sem comprometer o desempenho do sistema.

Além disso, a importância de uma base de dados escalável para sensores de veículos é destacada por Abdelkader, Elgazzar e Khamis (2021) em seu estudo sobre a análise de dados de sensores automotivos. Eles enfatizam que a escalabilidade é essencial para a viabilidade e eficiência das aplicações de análise de dados em tempo real. Com um número crescente de veículos equipados com sensores e a expansão contínua da Internet das Coisas (IoT), a

capacidade de armazenar grandes volumes de dados de forma escalável é crucial para garantir que as análises sejam realizadas de maneira eficaz.

O estudo de (CHEN; HAN; WANG, 2010) traz como o uso do SQLite em aplicações IoT torna-se uma escolha adequada para dispositivos com recursos limitados. Em virtude de ser um sistema de gerenciamento de banco de dados relacional que se destaca por sua portabilidade e simplicidade, o SQLite utiliza um formato de arquivo de banco de dados compacto, que permite uma alta taxa de compressão dos dados. Isso significa que os dados coletados pelos sensores veiculares podem ser armazenados de forma eficiente, ocupando menos espaço de armazenamento. Além disso, o SQLite é conhecido por sua eficiência em relação ao processamento de consultas Structured Query Language (SQL). Ele oferece suporte completo a recursos como índices, junções e transações, permitindo consultas rápidas e eficientes nos dados armazenados. Isso é especialmente importante para aplicações de sensores veiculares, onde a análise em tempo real e a tomada de decisões rápidas são essenciais para a segurança e o desempenho do veículo.

3 ESTADO DA ARTE

Nesta seção, é apresentado como o processo de revisão sistemática foi executado. Este procedimento possui como principal objetivo verificar soluções já existentes para uma ou mais etapas de coleta, tratamento, transmissão, armazenamento, análise, diagnóstico ou monitoramento de dados de bordo de veículos.

3.1 Revisão Sistemática

As palavras-chave utilizadas para abordar as diversas etapas que constituem o fluxo total dos dados de bordo de um veículo, formam a expressão de busca: 'OBD' e 'IoT'. Foram considerados válidos para a busca inicial os estudos que possuíam estas palavras em qualquer parte do documento. Foram utilizadas as bases Biblioteca Digital IEEE Xplore, Biblioteca Digital ACM, Scopus, Science Direct, Springer, Periódicos Capes e Web Science como fontes de coleta de estudos para apoiar a revisão sistemática da literatura. Essa pesquisa teve como fundamentação teórica e segue os procedimentos da publicação de revisão sistemática apresentada por (KITCHENHAM, 2004).

A partir da expressão de busca e as bases, são reunidos estudos em torno de todas as etapas e suas variações de técnicas e estratégias. Consequentemente, torna-se possível verificar o estado da arte neste contexto. Na revisão, usou-se apenas conceitos abrangentes diretamente ligados ao contexto de veículos para identificar diferentes perspectivas de como a literatura soluciona os problemas, quais são as ferramentas e técnicas empregadas e seus processos evolutivos.

Ao desenvolver a revisão sistemática desta maneira, é possível obter acesso a uma rica base de informações e conhecimentos diversos sobre o tema discorrido em nosso trabalho. Isso possibilitou identificar diferentes perspectivas e construções já existentes de propostas semelhantes a deste estudo, o que traz uma boa perspectiva de problemas e desafios solucionados e a serem resolvidos também. Essa pesquisa foi realizada em setembro de 2020 e devido à situação pandêmica de covid-19, ocorrida no início de 2021, não obteve sequência, tendo os resultados limitados até o ano de 2020. Portanto, seguindo as especificações anteriormente explicadas, foram retornados um total de 497 documentos, como é possível observar com mais detalhes no quadro 3.

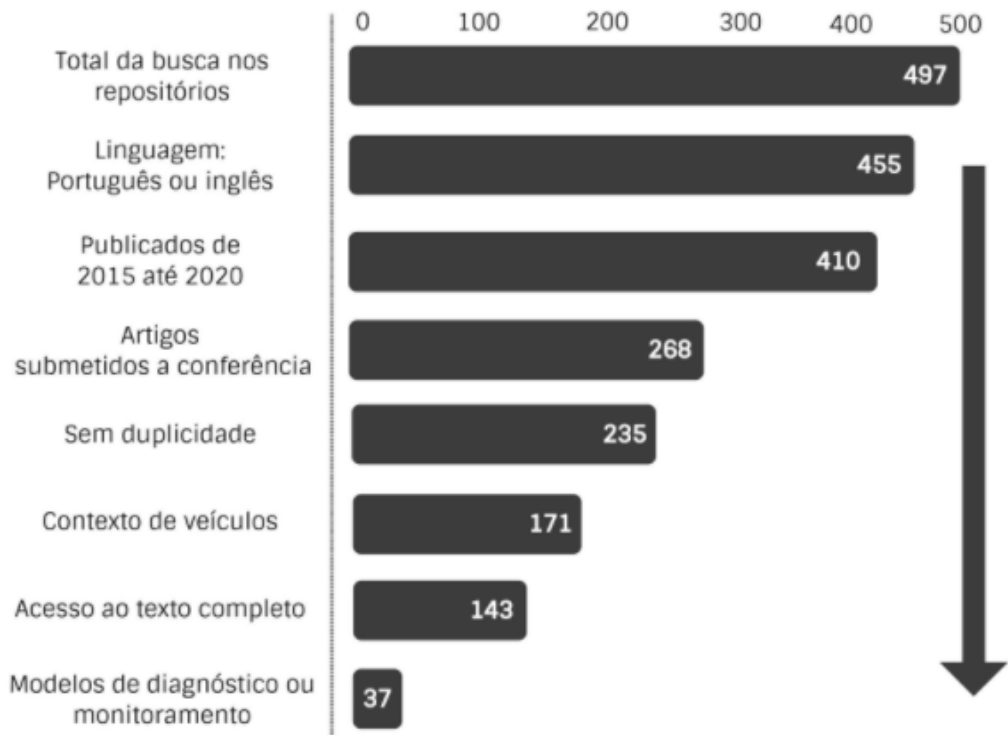
Inicialmente, de maneira a refinar estes resultados foram aplicados critérios, o primeiro critério (1) diz respeito à língua na qual foram escritos estes documentos. Portanto, do total de 497 documentos anteriores incluímos apenas os escritos em inglês ou português e desta maneira 42 estudos foram excluídos. Para o segundo critério (2) foram filtrados 45 estudos com publicações anteriores a 2015 com o intuito de manter as soluções mais adequadas aos problemas atuais. O critério subsequente (3) trata do formato em que a publicação foi feita, foram mantidas pesquisas no formato de artigos, visto que existiam ainda documentos de registros

de patentes e atas de reuniões que não forneceriam o detalhamento esperado de proposta de projeto, neste filtro foram removidos 142 documentos. Restando 268 artigos foi aplicado o critério de duplicidade (4) que retirou 33 documentos. A partir desta seleção, obteve-se 235 artigos restantes. Como quinto critério (5) o título ou resumo deveriam estar relacionados ao contexto de veículos, nesta fase 171 pesquisas foram selecionadas para a próxima fase. O próximo critério (6) foi verificar se era possível ter acesso ao documento completo pela fonte encontrada e 31 trabalhos foram descartados. No último critério (7) foram consideradas apenas pesquisas que proponham um modelo de diagnóstico ou monitoramento de dados de carros com foco no veículo o que fez com que 34 artigos fossem selecionados ao final. A figura 11 ilustra os filtros utilizados na revisão sistemática e os números de artigos obtidos em cada etapa.

Quadro 3 – Resultado preliminar da revisão sistemática

Repositório	Número de Resultados
Biblioteca Digital IEEE Xplore	26
Biblioteca Digital ACM	44
Scopus	38
Science Direct	95
Springer	153
Periódicos Capes	132
Web of Science	9

Figura 11 – Filtros da revisão sistemática.



Fonte: Autoria própria.

Com base nesta filtragem e seleção destes 34 artigos, são discutidas essas pesquisas no desenvolvimento da segunda etapa deste estudo. Todos os documentos conforme último critério citado possuem um modelo de diagnóstico ou monitoramento de dados de bordo, estes modelos são investigados na etapa de exploração de modelos que é apresentada na seção a seguir.

3.1.1 Exploração da Revisão

A qualidade de pesquisa é diretamente proporcional à qualidade do processo de condução do levantamento dos estudos em sua área de pesquisa, argumentam Boote e Beile (2005). De forma a realizar a exploração da revisão, foi desenvolvido um quadro organizado em 7 colunas, buscando entender e segmentar cada um dos artigos, possibilitando o agrupamento dos mesmos para compreender algumas informações relevantes para o estudo. Conforme apresentado na tabela 1, foram realizadas algumas pesquisas dentro dos trabalhos selecionados, onde procurou-se responder se apresentavam a arquitetura proposta de forma explícita, a quantidade de camadas apresentadas, quais etapas do processo estão envolvidas no modelo, quais as tecnologias utilizadas, os diferenciais e técnicas e hipóteses utilizadas e quais foram as melhorias propostas pelos autores em cada um dos artigos.

Tabela 1 – Revisão sistemática

Título	Ano	Apresenta a arquitetura?	Nº de camadas	Etapas	Tecnologias utilizadas	Possíveis contribuições futuras citadas
Fuzzy decision method to improve the information exchange in a vehicle sensor tracking system	2015	Sim	3	coleta, armazenamento e transmissão	OBD, Android, Fuzzy, LowCode, Sistemas Inteligentes	otimizar análises e armazenamento do processo com melhor refinamento do processo
Memory optimized lifetime vehicle data acquisition framework	2015	Sim	3	coleta, tratamento, armazenamento, transmissão, análise	OBD, CAN, GSM, HTTP	qualidade dos dados, mais modelos de transformações de dados
lot based automated car maintenance assist	2017	Sim	4	coleta, armazenamento, transmissão e análise	OBD, raspberry, firebase, android	expansão da análise e melhoria de alarmes

Tabela 1 – Revisão sistemática

Título	Ano	Apresenta a arquitetura?	Nº de camadas	Etapas	Tecnologias utilizadas	Possíveis contribuições futuras citadas
Smart Car Use Case: Dynamic Reconfigurable IoT Convergence with BigData	2015	Sim	3	coleta, armazenamento, transmissão, análise	OBD, CAN, MQTT, HDFS, MongoDB, Rest	transmissão de dados e otimizações de algoritmos de imagens
Applied Internet of Things (IoT): Car monitoring system using IBM BlueMix	2016	Sim	3	coleta, armazenamento, transmissão	OBD, IBM BlueMix	melhoria no monitoramento e análises
Comparing OBD and CAN Sampling on the go with the SensorHUB Framework	2016	Sim	3	coleta, tratamento, transmissão, armazenamento e análise	OBD, CAN, Apache BigData, Hadoop, PentahoBI, OBDCan-Compare	realizar aprendizado de máquinas, interface com mais sensores
A Study on Big Data Thinking of the Internet of Things-Based Smart-Connected Car in Conjunction with Controller Area Network Bus and 4G-Long Term Evolution	2017	Sim	3	coleta, transmissão e análise	OBD, 4G-LTE, Distributed File System	armazenamento relacional
An App Visualization design based on IoT self-diagnosis Micro Control Unit for car accident prevention	2017	Sim	5	coleta, tratamento, transmissão e análise	OBD, Zigbee, wifi, bluetooth, Sistemas Inteligentes	-

Tabela 1 – Revisão sistemática

Título	Ano	Apresenta a arquitetura?	Nº de camadas	Etapas	Tecnologias utilizadas	Possíveis contribuições futuras citadas
CarStream: an industrial system of big data processing for internet-of-vehicles	2017	Sim	4	coleta, tratamento, armazenamento, transmissão e análise	OBD, CAN, Apache Kafka, JS-torm Cluster, Hadoop, NOSQL, SQL	desenvolver arquitetura de microsistema e aplicar o carstream em uma frota maior
Connected car: Engines diagnostic via Internet of Things (IoT)	2017	Sim	3	coleta, tratamento, transmissão	CAN, C, REST, Arduino	armazenamento, segurança e casos com big data
OBD-II and raspberry Pi technology to diagnose car's machine current condition: study literature	2017	Sim	3	coleta, armazenamento e transmissão	OBD, raspberry, firebase, android	melhoria na coleta trocando o padrão do conector (wifi para bluetooth)
A customer feedback platform for vehicle manufacturing compliant with industry 4.0 vision	2018	Sim	3	coleta, transmissão, armazenamento, análise	OBD, BT, 3G/4G, Wifi, API-Rest, Java, NGINX, mongoDB, MYSQL, NodeJS	novas análises, melhoria na transmissão e aumento da variedade de sensores captados
A Design of a Lightweight In-Vehicle Edge Gateway for the Self-Diagnosis of an Autonomous Vehicle	2018	Não	-	coleta, transmissão	OBD, CAN, Most, Ethernet, Sistemas Inteligentes	testes em veículos reais

Tabela 1 – Revisão sistemática

Título	Ano	Apresenta a arquitetura?	Nº de camadas	Etapas	Tecnologias utilizadas	Possíveis contribuições futuras citadas
An Integrated Self-Diagnosis System for an Autonomous Vehicle Based on an IoT Gateway and Deep Learning	2018	Sim	4	coleta, tratamento, armazenamento, transmissão, análise	CAN, Gateway, MOST, TDS, FlexRay, Sistemas Inteligentes	utilização de mais sensores e veículos reais
Enabling cyber-physical systems for 5G networking: a case study on the automotive vertical domain	2018	Sim	4	coleta, transmissão	OBD, 5G, Open Source MANO (OSM)	avaliação de soluções fim a fim
IoT Cloud Based Real Time Automobile Monitoring System	2018	Sim	3	coleta, transmissão e análise	OBD, GPU, Machine Learning, Raspberry PI, Sistemas Inteligentes	escalabilidade, aumento no número de sensores coletados e novas análises com diferentes algoritmos
IoT-Based System Health Management Infrastructure as a Service	2018	Sim	4	coleta, transmissão, armazenamento e análise	OBD, Wifi, BT, Arduino, Node, Raspberry Pi, MySQL	aplicação em outros modelos industriais
ProCMotive: Bringing Programmability and Connectivity into Isolated Vehicles	2018	Sim	4	coleta, tratamento, armazenamento, transmissão, análise	OBD, Smart-Core, Raspberry PI	envio de comandos e reavaliação da coleta
Towards the implementation of the Social Internet of Vehicles	2018	Sim	4	coleta, transmissão, armazenamento e processamento	OBD, Smart-Core, Wifi, BT, Raspbian, Docker	-

Tabela 1 – Revisão sistemática

Título	Ano	Apresenta a arquitetura?	Nº de camadas	Etapas	Tecnologias utilizadas	Possíveis contribuições futuras citadas
Vehicle Remote Health Monitoring and Prognostic Maintenance System	2018	Sim	3	coleta, transmissão, tratamento, análise	OBD, BT, PCA, Sistemas Inteligentes	mais veículos, mais dados e refinar algoritmos de classificação
A Blockchain-based Secure Internet of Vehicles Management Framework	2019	Sim	4	coleta, transmissão, tratamento, processamento, armazenamento	OBD, Dapp, Wifi, BT, 5G, Arduino	utilizar um framework de blockchain com estrutura não ortogonal
Design of an IoT-based Vehicle State Monitoring System Using Raspberry Pi	2019	Sim	3	coleta, armazenamento, transmissão, análise	OBD, CAN, Raspberry Pi, Python, Node-red	detecção em tempo real das falhas
Improving the Fleet Monitoring Management, through a Software Platform with IoT	2019	Sim	4	coleta, tratamento, transmissão, armazenamento	OBD, IoT, Cloud, REST	maior quantidade de testes e validações
Instantaneous Fuel Consumption Estimation Using Smartphones	2019	Sim	3	coleta, tratamento, transmissão e análise	OBD, FAST-Sim, USGS TNM Server	calibrar os algoritmos propostos e aplicar em veículos híbridos
IoT-based predictive maintenance for fleet management	2019	Sim	3	coleta, tratamento, armazenamento, transmissão e análise	OBD, BT, Android, LTE, Wifi, COSMO, VSRDB, MQTT, DDB, Raspberry Pi, Sistemas Inteligentes	utilização em maior escala para melhorar o treinamento e desenvolver uma MVP

Tabela 1 – Revisão sistemática

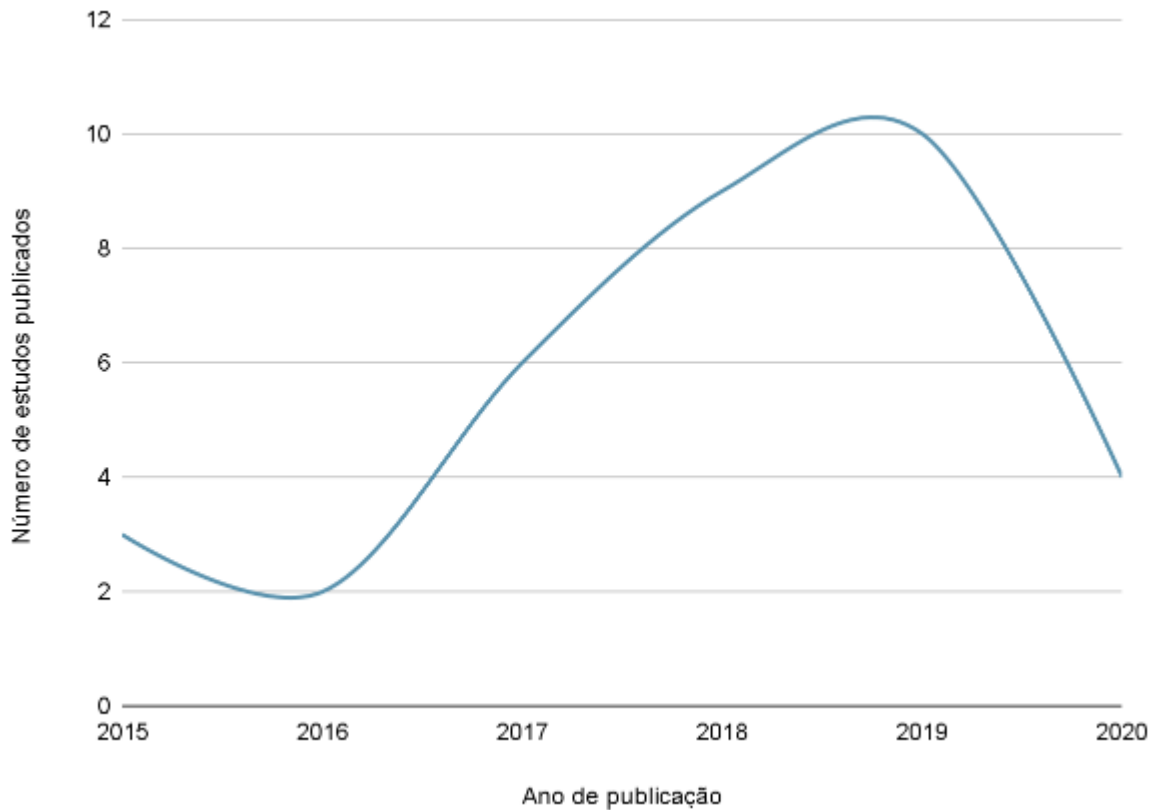
Título	Ano	Apresenta a arquitetura?	Nº de camadas	Etapas	Tecnologias utilizadas	Possíveis contribuições futuras citadas
LPWAN-Based Vehicular Monitoring Platform with a Generic IP Network Interface	2019	Sim	5	coleta, tratamento, transmissão, armazenamento e análise	OBD, LoRaWan, Wifi, SCHC, UNIX	testar em diferentes ambientes, como rodovia e com maior número de veículos conectados
Machine learning based real-time vehicle data analysis for safe driving modeling	2019	Sim	3	coleta, tratamento, transmissão, armazenamento e análise	OBD, Dashcam, Machine Learning, Sistemas Inteligentes	utilização do modelo com maior base de dados e utilização para entender modo de condução
Performance Evaluation of an Edge OBD-II Device for Industry 4.0	2019	Sim	3	coleta, armazenamento, tratamento, transmissão	OBD, GSM, BT, Wifi, AWS, Freematics One+,	gravar dados localmente quando não possuir internet, realizar compressão
The Lightweight Autonomous Vehicle Self-Diagnosis (LAVS) Using Machine Learning Based on Sensors and Multi-Protocol IoT Gateway	2019	Sim	4	coleta, tratamento, transmissão, armazenamento e análise	OBD, Machine Learning, MIGM, Sistemas Inteligentes	utilização de dados em tempo real
Smart vehicle monitoring and analysis system with IOT technology	2019	Sim	3	coleta, análise, transmissão, armazenamento	OBD, MySQL, DreamWeaver	melhoria da performance da coleta e transmissão
Deep Learning (DL) Based On-board Indirect Useful Life Prediction	2020	Não	-	coleta, transmissão, armazenamento e análise	OBD, GPS, MobileNet, DNN, Flask	analisar as condições do tráfego para melhorar o consumo de combustível

Tabela 1 – Revisão sistemática

Título	Ano	Apresenta a arquitetura?	Nº de camadas	Etapas	Tecnologias utilizadas	Possíveis contribuições futuras citadas
Experimental vs. simulation analysis of LoRa for vehicular communications	2020	Não	-	transmissão	Arduino, LoRaWan, GPS	investigar o impacto da alta velocidade e densidade na estrutura estudada
Fuzzy control-based system feed-back by OBD-II data acquisition for complementary injection of hydrogen into internal combustion engines	2020	Não	-	coleta, transmissão	OBD, BT, Raspberry Pi, Python, Fuzzy, Sistemas Inteligentes	aumentar os parâmetros de controle para melhorar os resultados
Performance Evaluation of an evolving data compression algorithm embedded into an OBD-II edge device	2020	Sim	3	coleta, tratamento, armazenamento, transmissão, análise	OBD, Swinging Door, Freematics One+, 3G/4G	expandir o algoritmo de compressão pra usar mais variáveis simultaneamente

Fonte: Autoria própria (2023).

Do total dos 34 estudos, a apresentação da arquitetura em forma de diagramas e/ou descrições detalhadas que abordam diretamente as suas especificações foram apresentadas em 30 estudos. Ou seja, apenas 4 pesquisas não apresentam de forma objetiva e clara os modelos de arquitetura utilizados. O gráfico 1 mostra o número de pesquisas na área de modelos de diagnósticos, utilizados na revisão, publicados por ano. Nos 3 primeiros anos estudados (de 2015 a 2017) apenas 11 trabalhos foram publicados, quase se igualando ao ano seguinte (2018), que teve 9 modelos publicados. Em 2019 seguinte, o maior número foi apresentado, ao todo 10 estudos. Já dentro do primeiro semestre de 2020, os modelos selecionados totalizaram 4.

Gráfico 1 – Publicações ao longo dos anos

Fonte: Autoria própria (2023).

A respeito do número de camadas de arquitetura, notamos que o desenvolvimento de arquiteturas mais simples, em 3 níveis, prevalece com 18 pesquisas contra 10 trabalhos em 4 níveis e apenas 2 apresentam uma arquitetura mais complexa, separada em 5 níveis. Dos estudos em 3 níveis, 100% especificam de forma clara as etapas de coleta e transmissão. Já na etapa de análise, tratamento e armazenamento temos 13, 9 e 12 estudos abordando respectivamente em cada uma destas etapas. Nota-se que esta preocupação nos primeiros passos do processo, reflete diretamente a importância na obtenção de um resultados satisfatório, visto que estas etapas tornam-se primordiais para a manutenção de dados íntegros e confiáveis.

Dentre os que possuem 4 níveis de arquitetura, 10 estudos apresentam coleta e transmissão, apenas 6 dizem respeito à tratamento e análise abordados em uma forma de melhoria na arquitetura, 90% dos estudos dessa categoria possuem o armazenamento especificado, demonstrando uma crescente proporcional entre o número de camadas e a importância do armazenamento de forma não abstraída nas demais etapas.

Tratando-se dos estudos que possuem mais camadas, novamente coleta e transmissão se mostram imprescindíveis, sendo o foco destes estudos a análise, mesmo com o baixo número de estudos dessa categoria (dois). Levando em consideração as tecnologias utilizadas, a padronização OBD é adotada em 31 estudos contra apenas 3, dentre estes, os quais são

focados em investigar o impacto da substituição deste protocolo, seja na etapa de transmissão quanto na de comunicação dos dispositivos.

Outro ponto relevante neste escopo de pesquisa é a aproximação de tecnologias de hardware com os modelos propostos, onde 14 estudos apresentam com clareza as tecnologias e como a utilização de um hardware em específico auxilia na obtenção de resultados. Na divisão de hardware, 9 estudos utilizam dispositivos Raspberry, contra 4 estudos utilizando Arduino e 1 utilizando um hardware gateway de tecnologia privada.

Em relação aos diferenciais encontrados nos métodos apresentados, 10 estudos utilizaram de sistemas inteligentes, entre eles 80% foram conduzidos com o intuito de classificar os dados coletados do veículo em perfis de utilizador, seja ele o motorista ou o automóvel. Outro ponto que se destacou foi a presença do ramo de IoV, sugerindo a importância da conexão dos veículos com o cenário em que eles são submetidos. Por fim, uma das maiores dificuldades levantadas no grupo de pesquisas selecionadas é o desempenho, já que quando trabalhado com dados gerados de veículos e com os recursos presentes nesse contexto, os parâmetros podem rapidamente apresentar desempenho de maneira não satisfatória.

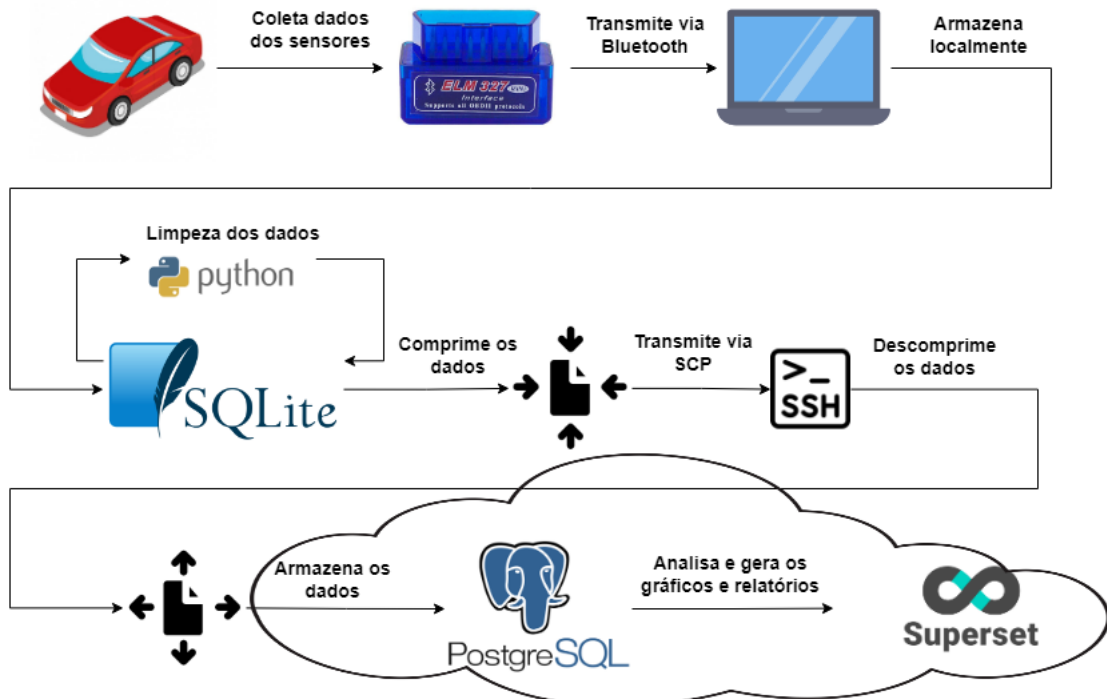
Uma das questões realizadas na revisão é a respeito das melhorias propostas pelos autores em seus respectivos estudos. Deste questionamento, diversas respostas foram coletadas, mas algumas aparecem com maior frequência. Dentre essas respostas, algumas abordam as mudanças nos parâmetros de testes, sendo eles, escopo, número de amostras e até condições de testes, por exemplo a utilização em veículos reais. Um aspecto faltante, que apenas 2 estudos apresentam, é a compressão dos dados, mas os que apresentam em alguma etapa, confirmam os bons resultados dessa adoção, principalmente devido à limitação de Wi-Fi e Bluetooth nesse contexto. Outro ponto ressaltado em melhorias de trabalhos futuros desta área é a questão da escalabilidade e desempenho, visto que cada vez mais o número de sensores tem aumentado nos veículos conforme apresentado anteriormente.

A prevalência de OBD domina amplamente entre os estudos, variando apenas quando o estudo busca substituir esta ferramenta em específico. Portanto, quando considerado este critério tem-se 8,8% de ocorrências da não adoção do OBD.

4 METODOLOGIA

Neste capítulo é apresentada a metodologia proposta para a realização deste projeto com a finalidade de alcançar os resultados e objetivos propostos na seção 1.1, além de uma visão geral da arquitetura final apresentada na figura 12.

Figura 12 – Arquitetura do modelo



Fonte: Autoria própria (2023).

Para atingir as propostas desta pesquisa, foi utilizada uma abordagem multimetodológica, utilizando-se da pesquisa exploratória, onde especificamos um estudo de caso, composto por revisão de literatura, análise de estudo de caso e implementação prática. A revisão da literatura se concentrou no estado atual da arte em modelos de captura e armazenamento de dados de veículos, bem como na identificação dos principais desafios e requisitos em cada etapa prevista em sistemas semelhantes. A análise do estudo de caso envolve uma investigação e desenvolvimento aprofundado de um sistema que incorpora todas as fases do fluxo de dados do veículo, buscando uma abordagem voltada para o lado qualitativo das informações, onde procuramos através dos resultados obtidos, entender se o modelo proposto satisfaz os objetivos especificados.

Nesta fase de desenvolvimento do experimento destacam-se a construção, desenvolvimento e validação do modelo de arquitetura de sistemas proposto, bem como a prototipagem de um aplicativo para coleta e armazenamento de dados. O sistema é desenvolvido para se comunicar com veículos via Bluetooth, utilizando o protocolo OBD2 para coleta de dados. Inicialmente, os dados coletados são armazenados localmente, em um notebook onde esta sendo

executado o script de conexão e coleta, ou na nuvem, dependendo da disponibilidade da rede e contexto da rede. Além disso, também existe um serviço de análise de dados e geração de relatórios com base no modelo proposto para alertar sobre necessidades de reparo.

4.1 Coleta

O desenvolvimento do sistema iniciou-se na fase de coleta. O foco desta atividade é retirar os dados do veículo de forma que seja possível extrair informações conforme frequência desejada. Conforme análises da revisão literária, a utilização do OBD2 como conector é uma tendência muito forte, contando com 91,17% dos estudos revisados, sendo que nos demais os objetivos das pesquisas são focados em investigar os impactos da substituição do protocolo de transmissão e comunicação. Esse efeito é justificado graças à padronização realizada na indústria, assim como a popularização dos conectores em lojas do ramo. Seguindo as considerações, foram selecionados diferentes conectores OBD2/ELM327 para a sequência do protótipo. Dos 4 dispositivos utilizados para o estudo, 2 possuem a marca Cbmmaker, sendo um da versão 1.5 e outro 2.1, tendo como principais diferenças da versão mais antiga para a mais atual: melhorias na velocidade e confiabilidade da coleta, um monitor ativo de economia de bateria e adição de novos parâmetros programáveis na coleta. Ambos os modelos não fornecem um software nativo. O dispositivo da fabricante Kuulaa se assemelha muito aos demais, porém oferece um software nativo, ou seja, existe um sistema proprietário para acompanhar os dados coletados pelo dispositivo, além de permitir conexão através de rede sem fio. O diferencial do dispositivo PCI-25k80, dentro dos citados, diz respeito ao tipo de conexão, já que o mesmo realiza a transmissão de dados através de conexão Universal Serial Bus (USB). A figura 13 representa os dispositivos e suas características:

O veículo é um componente crucial na coleta de dados, como fonte dos dados. Esses dados podem incluir informações sobre o desempenho, localização e uso do veículo, bem como o ambiente em que ele está operando. Ao coletar dados de vários veículos e analisá-los, é possível obter informações sobre padrões e tendências, assim como peculiaridades. Para este experimento utilizou-se os seguintes veículos para a testagem da conexão OBD, conforme quadro 4.

Quadro 4 – Veículos utilizados para coleta de dados

Modelo	Ano	Motor
Citroen C3	2008	1.4 16v Flex
Honda Fit	2012	1.4 16v Flex
Audi A3 Sportback	2008	2.0 Turbo 16v Gasolina

Fonte: Autoria própria (2023).

Figura 13 – Conectores OBD estudados

Nome do fabricante	Cbmmaker	Cbmmaker	Kuulaa	Desconhecido (PCI-25k80)
Imagem do dispositivo				
Versão	v1.5	v2.1	v1.5	v1.5
Tipo de conexão	Bluetooth	Bluetooth	Bluetooth/Wifi	USB
Software nativo	Não	Não	Sim	Não

Fonte: Autoria própria (2023).

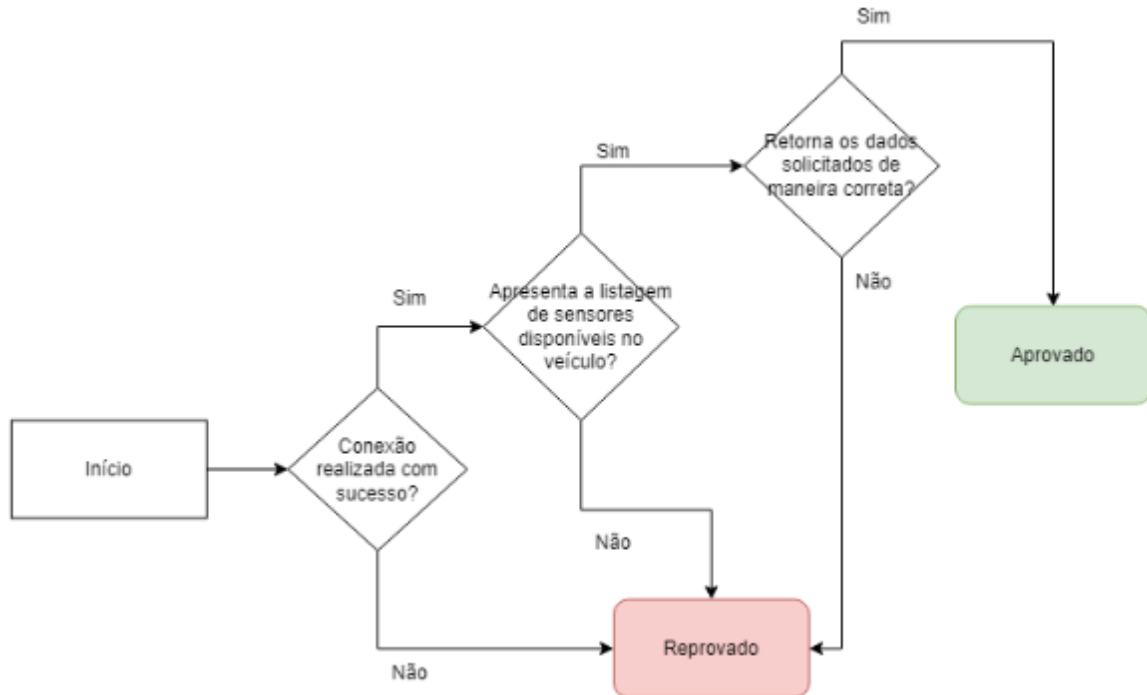
Um ponto essencial é a integração entre hardware e software no processo de coleta, devido ao fato de que se não for possível realizar a comunicação entre ambos, torna-se inviável a extração dos dados dos sensores. Para validar o cumprimento do objetivo, houve alternância de frequência da solicitação de dados, variando em intervalos de 1 até 60 segundos. Na tentativa de maximizar os resultados, foram criadas duas versões de coletores de dados em diferentes linguagens de programação. A primeira delas foi escrita em Python 3.7, utilizando a biblioteca python-OBD disponibilizada sob a licença de GPL-2.0. A versão remanescente escrita em Java, utiliza a biblioteca nomeada de obd-java-api sob licença Apache-2.0.

Diante do cenário descrito, foram realizados diversos testes com o intuito de verificar a possibilidade da extração dos dados assim como a quantidade de sensores disponíveis, que variam de veículo para veículo como de biblioteca para biblioteca. Notou-se que algumas combinações não atingiram o objetivo proposto da etapa, principalmente as versões desatualizadas de coletores. Também houve casos em que o objetivo foi cumprido porém o número de sensores disponíveis era minimizado, limitando a possibilidades nas etapas futuras. Este processo é ilustrado pela figura 14, bem como os resultados das validações dos testes são apresentados no quadro 5.

Com base nos resultados obtidos nos testes de extração de dados, concluiu-se que a escolha do conector CBmmaker v2.1 Bluetooth (Python) é dada graças ao sucesso na coleta em todos os modelos de veículos selecionados. Além de ter funcionado em todos os casos, o mesmo apresentou um número superior de sensores disponíveis, o que permitiu obter um maior número de dados significativos e úteis para o desenvolvimento do modelo deste estudo. Essa escolha foi determinante para o resultado final, já que a coleta de dados é uma das mais

importantes etapas, sendo que todas as demais etapas subsequentes dependem do sucesso da coleta de dados.

Figura 14 – Fluxograma de seleção de veículo



Fonte: Autoria própria (2023).

Quadro 5 – Compatibilidade de veículos e conectores

	Citroen C3	Honda Fit	Audi A3	Numero de sensores disponíveis
CBmmaker v1.5 BT (Python)	✓	✓	X	12
CBmmaker v1.5 BT (Java)	X	X	X	0
CBmmaker v2.1 BT (Python)	✓	✓	✓	12
CBmmaker v2.1 BT (Java)	✓	✓	X	7
Kuulaa v1.5 BT (Python)	✓	X	✓	9
Kuulaa v1.5 BT (Java)	✓	X	✓	5
PCI-25k80 USB v1.5 (Python)	X	X	X	0
PCI-25k80 USB v1.5 (Java)	X	X	X	0

Fonte: Autoria própria (2023).

4.1.1 Sensores de coleta

Sensores automotivos são componentes cruciais para o monitoramento e controle de sistemas em veículos. Eles permitem que a central eletrônica do carro monitore e colete informações sobre o funcionamento de diversos sistemas, desde o motor e transmissão até os sistemas de segurança, como freios e airbags. Os dados coletados pelos sensores são processados em tempo real e usados para ajustar e otimizar o desempenho do veículo, garantindo sua segurança e eficiência.

Sensores automotivos são essenciais para o IoV, permitindo que os veículos sejam monitorados remotamente e forneçam informações em tempo real sobre o seu estado, como sua localização, condição mecânica e padrões de direção. Isso permite que as empresas de automóveis ofereçam serviços personalizados aos seus clientes, como diagnóstico remoto, manutenção preventiva e até mesmo recursos de direção autônoma.

No quadro 6 podemos visualizar uma lista dos sensores disponíveis retornados pela central do veículo Honda Fit. No total obtivemos 23 sensores, sendo que 11 deles ao ser realizada a requisição do dado, era retornado o valor 'None' impossibilitando a sua utilização, resultando então em um total de 12 disponíveis para coleta e com dados significativos.

Observando a lista de sensores disponíveis para coleta, percebemos que em sua grande maioria se tratava de dados específicos, relacionados a componentes isolados do motor do veículo, porém fazendo um cruzamento dos dados é possível obter informações relevantes em relação ao funcionamento da combustão e também do refrigeração do sistema, que fazem parte de alguns itens essenciais para o seu funcionamento adequado. Dada a lista de sensores disponíveis para monitoramento, optamos por realizar a coleta em um conjunto de 8 sensores, sendo eles exibidos no quadro 7, juntamente com sua representação e o tempo selecionado para intervalo da requisição do dado.

Buscou-se através da coleta, encontrar dados que possuíssem uma relação direta em sua grande maioria e que pudessem produzir informações relevantes para o objetivo proposto por este estudo. Por exemplo, através do cruzamento de dados capturados das rotações por minuto, posição do acelerador e fluxo de ar, nota-se que a medida que o motorista acelera mais o seu veículo o fluxo de ar, que passa pela admissão aumenta para realizar a combustão e se o número de rotações por minuto aumenta, acompanhando as outras informações, expondo que a combustão dentro do motor está acontecendo normalmente e que o veículo não possui falha. Outro exemplo poderia ser o monitoramento da temperatura do líquido de arrefecimento, que em condições normais opera entre 80°C e 100°C após o veículo chegar a sua temperatura ideal de funcionamento, desta forma é possível monitorar alterações no sistema de refrigeração do automóvel. Como alguns dados possuem uma mudança mais frequente de seus valores, como por exemplo o número de rotações por minuto, velocidade e posição do acelerador, este fato impossibilita uma coleta em grandes intervalos de tempo, já que são dados que se alteram a cada instante. Portanto, por meio de testes de conexão e capacidade de armazenamento

e processamento, definimos um intervalo de tempo menor buscando maior entendimento do funcionamento do veículo, entretanto mantendo-se atento ao fato de evitar uma sobrecarga do sistema. A variável de temperatura do ar ambiente, possui um intervalo maior de coleta, já que se trata de uma informação que possui menor frequência de variação.

Quadro 6 – Sensores disponíveis para coleta

Nome	Disponibilidade	Descrição
PIDS_A	✓	Pinos aceitos para receber mensagem (01-20)
STATUS	✓	Status desde a ultima leitura de código de erro (texto)
FREEZE_DTC	X	Momento específico de uma falha (texto)
FUEL_STATUS	X	Status do sistema de combustível (texto)
ENGINE_LOAD	✓	Carga aplicada no motor (%)
COOLANT_TEMP	✓	Temperatura do líquido de arrefecimento (°C)
SHORT_FUEL_TRIM_1	X	Mistura ar/combustível jato curto bancada 1 (%)
LONG_FUEL_TRIM_1	X	Mistura ar/combustível jato longo bancada 1 (%)
SHORT_FUEL_TRIM_2	X	Mistura ar/combustível jato curto bancada 2 (%)
LONG_FUEL_TRIM_2	X	Mistura ar/combustível jato longo bancada 2 (%)
FUEL_PRESSURE	X	Pressão de combustível (kPa)
INTAKE_PRESSURE	X	Pressão na admissão (kPa)
RPM	✓	Rotações por minuto (rpm)
SPEED	✓	Velocidade (km/h)
TIMING_ADVANCE	X	Motor fora de ponto (graus)
INTAKE_TEMP	X	Temperatura da admissão (°C)
MAF	✓	Massa de fluxo de ar (g/s)
AMBIENT_AIR_TEMP	✓	Temperatura do ar ambiente (°C)
THROTTLE_POS	✓	Posição do acelerador (%)
AIR_STATUS	X	Status da entrada de ar (texto)
O2_SENSORS	X	Voltagem da sonda lambda (V)
RUN_TIME	✓	Tempo de execução (seg)
FUEL_LEVEL	✓	Nível de combustível (%)
RELATIVE_THROTTLE_POS	✓	Posição relativa do acelerador (%)

Fonte: Autoria própria (2023).

Quadro 7 – Sensores utilizados na coleta

Sensor/Variável	Representação	Intervalo para coleta
Consumo de ar pelo motor por minuto (engine_load)	Quantidade de ar que o motor está consumindo por minuto. Possui relação direta com as rotações por minuto e o fluxo de ar, para calcular em porcentagem o esforço realizado pelo motor. É medido em porcentagem e varia de 0 a 100%.	5 segundos
Rotações por minuto (rpm)	Número de rotações/ciclo (movimento completo) realizados pelo pistão dentro do cilindro em um minuto. Possível relacionar diretamente com consumo de ar pelo motor e velocidade.	5 segundos
Velocidade (speed)	Velocidade do veículo em km/h.	5 segundos.
Tempo ligado (run_time)	Tempo em segundos em que o veículo se encontra em funcionamento.	5 segundos
Temperatura do líquido de arrefecimento (coolant_temp)	Temperatura em graus Célsius do líquido de arrefecimento do motor. Permite visualizar se existe um sobreaquecimento do sistema, onde a faixa de temperatura ideal para atuação é de 80 a 95°C após alguns instantes de motor ligado.	5 segundos
Fluxo de ar (maf)	Massa de ar que está entrando na admissão do veículo, medida em g/s ou então kg/s. Em um funcionamento normal seu valor deve acompanhar rotações por minuto ou posição do acelerador.	5 segundos
Posição do acelerador (throttle_position)	Mede em porcentagem a posição do acelerador (0 a 100). Possível relacionar diretamente com as rotações por minuto e fluxo de ar.	5 segundos
Temperatura do ar ambiente (ambient_air_temp)	Temperatura ambiente externa medida de graus Célsius.	1 minuto

Fonte: Autoria própria (2023).

4.1.2 Perfil do motorista

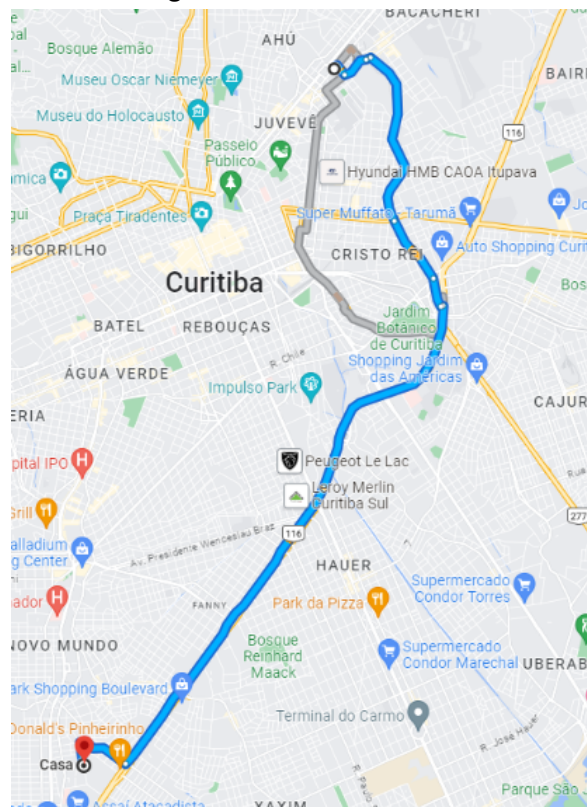
Motorista com 7 anos de carteira, acostumado a dirigir todos os dias em diversas situações e habituado a realizar os trajetos descritos posteriormente. O motorista procurou sempre antecipar os movimentos, evitando freadas e acelerações bruscas, seguindo os princípios da direção defensiva.

4.1.3 Cenários

Foram utilizados para o estudo três cenários possíveis para realização dos experimentos práticos, buscando abranger algumas situações distintas para o veículo e condutor, sendo eles:

Cenário 1 - Trânsito intenso, onde ocorre a maior exaustão do motor do veículo, já que encontra-se com pouca ventilação natural na parte frontal do veículo e por consequência pouco resfriamento do sistema de refrigeração, sendo necessário até mesmo a utilização da ventoinha auxiliar para gerar a ventilação de maneira forçada, conforme figura 15.

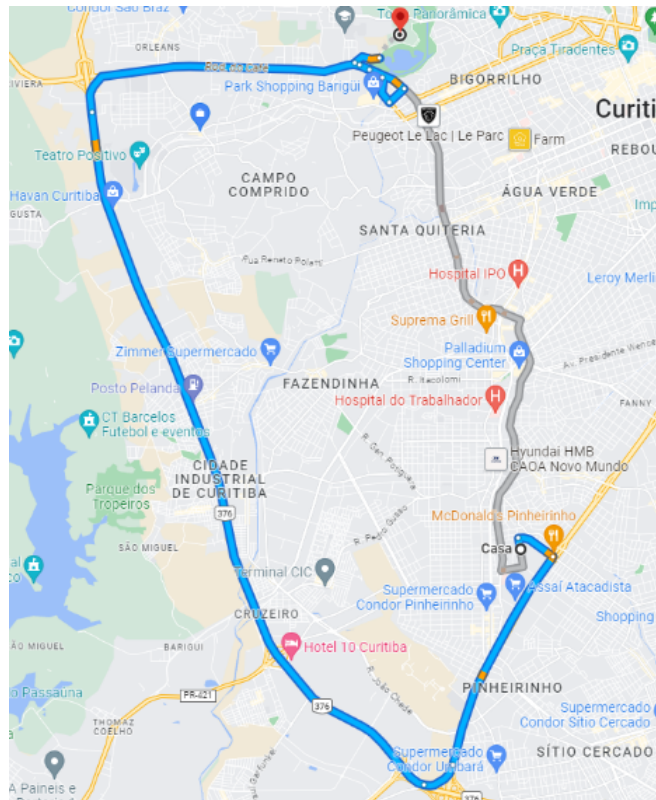
Figura 15 – Rota cenário 1



Fonte: Adaptado de Google Maps.

Cenário 2 - Movimentação mais constante, em uma rodovia e de maneira geral sem realizar paradas, ou seja, velocidade praticamente constante com ventilação natural, conforme figura 16.

Figura 16 – Rota cenário 2

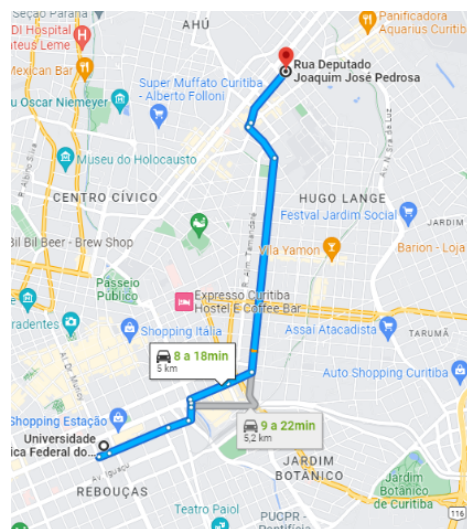


Fonte: Adaptado de Google Maps.

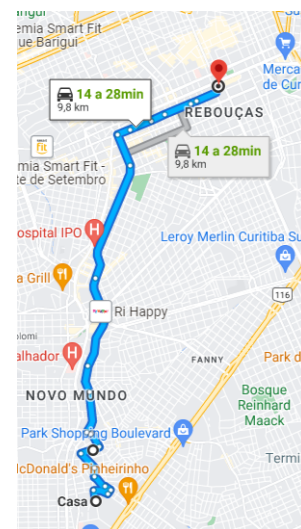
Cenário 3 - Misto de trânsito urbano e vias rápidas, que possui condições semelhantes aos dois cenários anteriores, possibilitando momentos de pouca refrigeração e movimentação do veículo e posteriormente uma maior refrigeração com um maior deslocamento do veículo, conforme figura 17.

Figura 17 – Rota cenário 3

(a) Rota Principal



(b) Rota Alternativa



Fonte: Adaptado de Google Maps.

4.2 Tratamento

Neste estudo, a limpeza dos dados é considerada como uma operação executada para remover as anomalias dos dados extraídos das etapas anteriores buscando aproximar o máximo possível da representação real do veículo em suas condições de utilização. Processo este que envolve a eliminação dos erros e resolvendo inconsistências apresentadas, utilizando uma transformação dos dados para um formato uniforme (GU, 2011).

Com um imenso número de dados coletados na primeira etapa, tornou-se quase impossível a limpeza manual devido ao longo tempo que seria necessário para uma pessoa realizar esta tarefa, além de estar mais suscetível a erros. Baseando-se nesta questão e nos recursos computacionais disponíveis, optou-se por seguir uma investigação inicial manual, seguida do desenvolvimento de uma metodologia em conjunto com scripts para realização de cada fase deste conjunto de métodos.

1. A primeira fase consistiu na observação humana manual com o objetivo de identificar padrões de erros e inconsistências que ocorrem nos dados produzidos e coletados dos sensores. A listagem completa pode ser observada no quadro 8.

Quadro 8 – Primeira fase do tratamento de dados

Dados vazios	Em vários ciclos de coletas foi possível analisar que os dados solicitados através do sistema de coleta retornaram um dado vazio, que não representa o estado natural do veículo)
Unidade de medida	Os dados são apresentados sempre com suas respectivas unidades a seguir como padrão do coletor utilizado. Isto pode gerar problemas em aplicações de técnicas que não levam em consideração estes detalhes.
Padrões de intervalo	Nem todos os dados são apresentados no mesmo intervalo, apesar da configuração aplicada no sistema de coleta, os mesmos são coletados assim que possível e estão sujeitos a atrasos e falhas nas suas capturas.
Estrutura utilizada	Seguindo padrões do fabricante, alguns dados apresentaram uma forma pura da coleta, contendo diretamente textos, caracteres ou separadores inválidos.
Arredondamento	Em alguns casos os sensores não possuem um padrão de arredondamento, ou seja, seus dados podem conter excesso de casas decimais, que não agregam valor a nenhuma pesquisa, podendo até a motivar erros de arredondamento futuros.

Fonte: Autoria própria (2023).

2. Em sequência, a segunda fase consiste na definição das transformações, executadas através de algum conjunto de operações, que buscam detectar e eliminar as anomalias nos dados. O quadro 9 representa cada inconsistência detectada na fase anterior e sua respectiva transformação em linguagem natural.

Quadro 9 – Segunda fase do tratamento de dados

Dados vazios	Remoção dos instantes nulos ou vazios.
Unidade de medida	Separação da unidade de medida em outro dicionário e remoção da unidade de medida do valor pontual.
Padrões de intervalo	Normalização dos intervalos através de agregação da média.
Estrutura utilizada	Remoção dos caracteres inválidos e textos.
Arredondamento	Padronização dos números em casas decimais.

Fonte: Autoria própria (2023).

3. Na terceira fase, considerada como a de verificação, a correção e eficácia da transformação são avaliadas. Esta fase consiste em várias iterações para verificar se todos os erros estão sendo corrigidos ou se novos passos são necessários. Como alguns erros só são visíveis após a transformação, é necessário um novo ciclo de processo de análise, projeto e verificação.

4. Após todas as fases e iterações, a substituição dos dados sujos e não íntegros, pela nova coleção de dados limpos e íntegros foi obtida.

4.3 Armazenamento

O modelo deste estudo se ampara em uma separação de duas bases de dados, uma localizada internamente no ponto de coleta e outra com mais recursos onde são agregados todos os dados coletados de todos os veículos para uma base de comparação. Esta decisão possibilita que independentemente da quantidade de veículos que participem do ecossistema do projeto, a integridade e desempenho do mesmo não sejam comprometidos. Levando em consideração a limitação de recursos presentes nos veículos, avaliou-se que não existia a possibilidade de instalar uma ferramenta de banco de dados de grande porte junto com o hardware de coleta, e neste mesmo sentido procurou-se alternativas que desempenham funcionalidades semelhantes porém com um consumo de recursos mais adequado.

Desta maneira, temos o SQLite como frente de armazenamento na primeira camada do modelo e para a camada final de análise utilizou-se Postgresql visto que possui uma ampla gama de recursos e é amplamente reconhecido por sua confiabilidade e escalabilidade. (CHEN; HAN; WANG, 2010). Isso também permitiu manter a compatibilidade entre bancos de dados que possibilita migrar facilmente um conjunto de dados gerados na borda para o sistema de processamento central que será responsável pelas demais etapas.

4.4 Compressão

Algumas técnicas foram avaliadas com o intuito de serem aplicadas no modelo. Um dos critérios adotados para verificação e escolha do melhor método para compressão é a partir da

exposição a um caso real, onde dado um conjunto de dados situados em um arquivo, a compressão sem perdas deve ocorrer reduzindo para o menor tamanho em disco possível. Neste experimento, foi recolhido um intervalo de tempo de dados de sensores que totalizaram 2023424 bytes. Foi considerado o vencedor neste critério, o método que atingiu a maior porcentagem de compressão, conforme o quadro 10.

Quadro 10 – Comparação entre métodos de compressão

Método	Tamanho final (bytes)	% de redução
BZip2	371161	81,66
LZMA2	338248	83,29
Gzip	504412	75,07

Fonte: Autoria própria (2023).

O LZMA2 apresentou a melhor redução no caso de teste anterior e devido a este fato, foi o método selecionado para utilização no modelo.

4.5 Sincronização

Esta atividade consiste na transferência dos dados do local onde foram coletados para onde serão processados e analisados adequadamente. Nesse processo leva-se em consideração principalmente a capacidade e eficiência, visto que o conjunto de dados já encontra-se preparado para ser transferido.

A sincronização é realizada a partir do momento em que uma conexão entre o dispositivo e um ponto de acesso a internet é detectada. Quando isso ocorre, e existe um conjunto de dados preparado a partir das etapas anteriores, este pacote é direcionado ao fluxo de transmissão.

Para a transferência do conjunto de dados, este modelo utiliza o *Secure Copy Protocol (SCP)*, que é uma ferramenta amplamente utilizada para transferir arquivos de forma segura entre sistemas remotos. As justificativas para utilização, incluem a segurança dos dados, a facilidade de uso, a eficiência na transferência e a compatibilidade com diversos sistemas operacionais.

Em termos de segurança, o SCP utiliza criptografia durante a transferência, protegendo contra interceptações indesejadas. Além da utilização do protocolo *Secure Shell (SSH)* para autenticação e estabelecimento de uma conexão segura entre os sistemas, garantindo a integridade e confiabilidade dos arquivos transferidos.

Como processo final da sincronização, com o intuito de manter um espaço de memória suficiente no dispositivo e evitar duplicidade dos dados, é executado um processo de remoção total dos pacotes já sincronizados, ou seja, se o fluxo já foi executado para determinado pacote com sucesso, o mesmo é removido permanentemente.

4.6 Análise

A fim de relacionar os mesmos cenários e obter informações alinhadas com o objetivo proposto neste estudo, foram empreendidos esforços ao longo de um período de 2 meses e 16 dias. Durante esse período, foram extraídos um total de 32.779 dados dos sensores do veículo. A coleta de dados se estendeu por 481 minutos e envolveu a utilização de diversas ferramentas de análise, incluindo o Apache Superset¹, SQL, DBeaver² e ferramentas de tabulação de dados. A principal finalidade dessa análise consistiu em identificar padrões e tendências nos dados, com o intuito de obter informações relevantes para alcançar as propostas do estudo.

Para realizar a análise, adotou-se uma abordagem que envolveu a exibição dos dados extraídos dos mesmos cenários em diferentes gráficos, com o auxílio da ferramenta Apache Superset. Dessa forma, o comportamento das variáveis ao longo do tempo específico foi considerado. Além disso, foram conduzidas análises exploratórias com o objetivo de visualizar os dados e identificar tendências preliminares. A utilização das ferramentas mencionadas proporcionou uma visualização clara dos dados coletados, revelou relações e padrões emergentes nos cenários. Essa abordagem exploratória foi crucial para identificar possíveis correlações e informações relevantes que poderiam contribuir para os objetivos do estudo. Ao adotar uma metodologia, visou-se garantir a confiabilidade e a validade dos resultados obtidos a partir das análises realizadas.

Em resumo, a análise dos dados coletados dos mesmos cenários ao longo de um período de tempo significativo proporcionou uma base para as investigações.

4.7 Operação

A orquestração de sequência dos *scripts* e sistemas utilizados para atingir os objetivos propostos inicialmente, é dada pela seguinte ordem: o primeiro passo é o registro de componente Bluetooth partindo do conector para o notebook, através da biblioteca selecionada, que registra a porta equivalente a esta conexão. Após este evento, é possível solicitar o primeiro conjunto de sensores, ainda no mesmo código fonte, chamado aqui de coleta. Para esta coleta continuar de forma indeterminada, a atividade é executada em forma de laço, com intervalos de tempo pré configurados, baseado no sensor.

O segundo passo envolve o armazenamento direto do dado não processado na estrutura designada, localizada dentro do veículo, com o esquema representado pela figura 18.

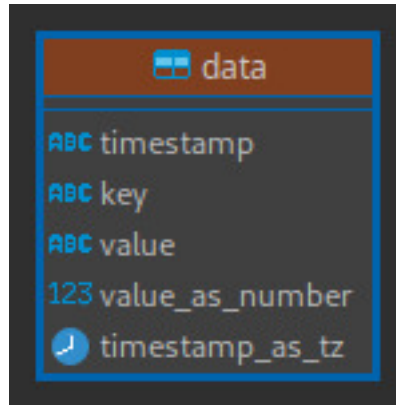
Em paralelo, com a execução deste evento, é executado um *script* repetidamente, responsável por monitorar a conexão com a internet. Caso obtenha sucesso, os mesmos dados são comprimidos e transferidos para outro dispositivo. Este último servidor possui um serviço semelhante, responsável por descomprimir um arquivo quando este é recebido e utilizar uma

¹ <https://superset.apache.org/>

² <https://dbeaver.io/>

ferramenta para transferir os dados para um banco de dados PostgreSQL³. A partir disso, já existe a possibilidade de uma análise, visto que as ferramentas para tal tarefa são diretamente conectadas com estes dados, proporcionando os gráficos e relatórios pré configurados.

Figura 18 – Esquema do banco local



Os códigos fontes responsáveis pela execução deste fluxo, podem ser encontrados através da ferramenta de controle de versionamento de código GitHub⁴, no repositório obd-collector que é mantido pelos autores deste estudo.

³ <https://www.postgresql.org/>

⁴ <https://github.com/>

5 DISCUSSÃO

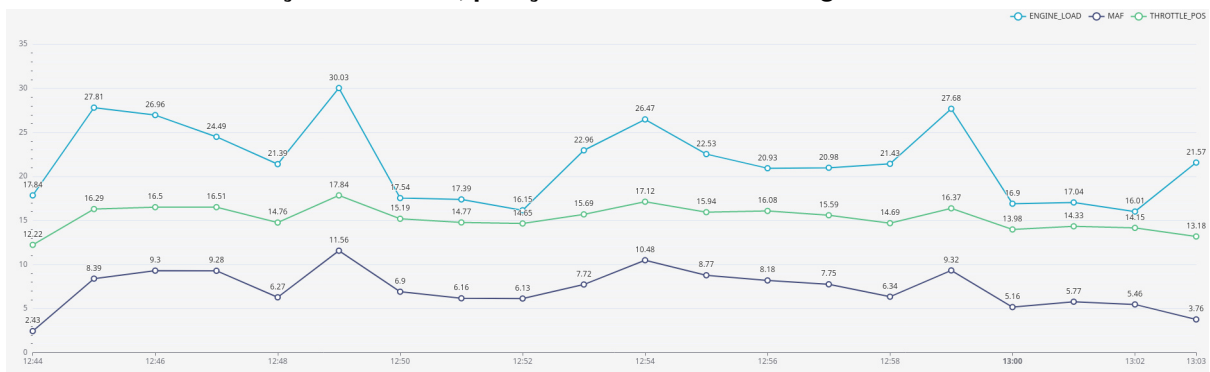
Nesta seção são apresentados análises, tendências e resultados em relação aos dados capturados e processados.

Este estudo abrangeu a comparação de diferentes cenários, a fim de identificar possíveis variações e tendências que pudessem influenciar o desempenho dos veículos. Além disso, realizamos a correlação de variáveis para entender as relações existentes entre os diferentes parâmetros coletados pelos sensores dos veículos. Por meio dessa análise, buscamos identificar possíveis padrões e dependências entre as variáveis, visando compreender como determinados fatores podem impactar no comportamento dos sensores e, conseqüentemente, na manutenção dos carros.

5.1 Posição do acelerador, Massa de fluxo de ar e Carga do motor

No gráfico 2 podemos visualizar o comportamento de três variáveis dentro do cenário 3. Os valores de massa de fluxo de ar, carga aplicada sobre o motor e posição do acelerador tendem a acompanhar uns aos outros. Ou seja, quanto mais o motorista pisar no acelerador, maior será a massa de ar entrando no motor para realizar a combustão e por consequência o valor da carga aplicada sob o motor também tende a aumentar. É possível observar que de maneira geral os valores possuem comportamento muito parecidos, o que está dentro do esperado. A leitura da Mass air flow (MAF) está diretamente relacionada à posição do acelerador e das rotações por minuto do motor (JANKOVIC; MAGNER, 2001). Caso por exemplo o valor da posição do acelerador fosse maior, ou seja, o motorista estivesse acelerando mais o veículo e o valor do sensor MAF não tivesse um comportamento semelhante, seria possível observar que o carro possui uma entrada de ar falsa na parte de admissão de ar, ou até mesmo o sensor MAF poderia estar com defeito, ocasionando falha de combustão no veículo.

Gráfico 2 – Relação entre MAF, posição do acelerador e carga do motor no cenário 3

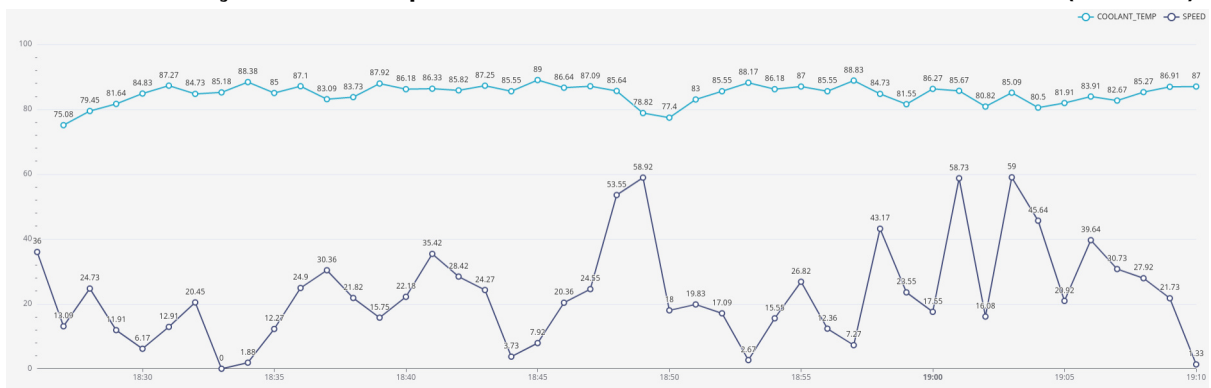


Fonte: Autoria própria (2023).

5.2 Temperatura do líquido de arrefecimento e velocidade

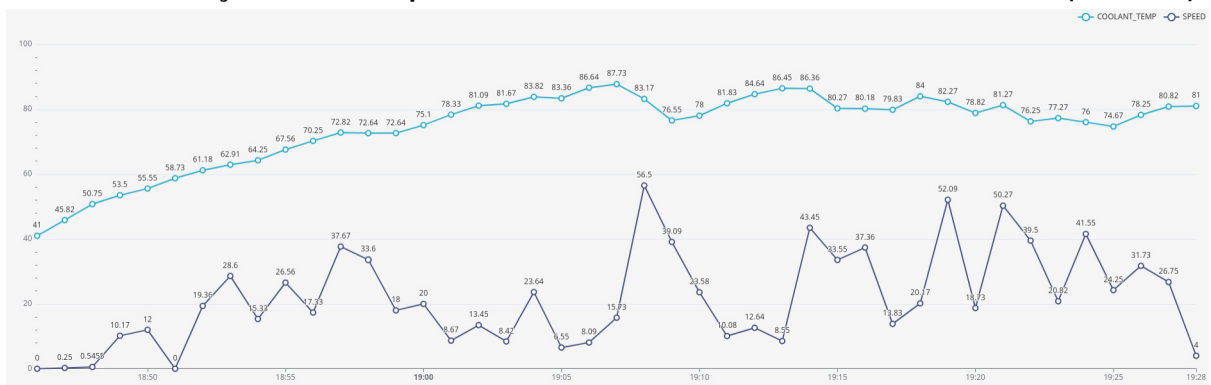
Nos gráficos 3 e 4, nota-se a coleta dentro do cenário 1 onde o veículo foi submetido a um trânsito intenso durante um período de tempo de aproximadamente 40 minutos em dois dias distintos. É possível observar o comportamento de duas variáveis, sendo elas a velocidade do veículo e também a temperatura do líquido de arrefecimento. O sistema de arrefecimento deve manter o carro na temperatura ideal de funcionamento, sendo geralmente a 90°C, pois o motor esquenta muito com as explosões da queima do combustível. Esse sistema tem como responsabilidade manter o propulsor frio para que as peças não derretam, mas não a ponto que as baixas temperaturas interfiram no funcionamento do veículo (SIQUEIRA; ALBUQUERQUE, 2020).

Gráfico 3 – Relação entre a temperatura de arrefecimento e velocidade no cenário 1 (Dia 25/08)



Fonte: Autoria própria (2023).

Gráfico 4 – Relação entre a temperatura de arrefecimento e velocidade no cenário 1 (Dia 02/09)



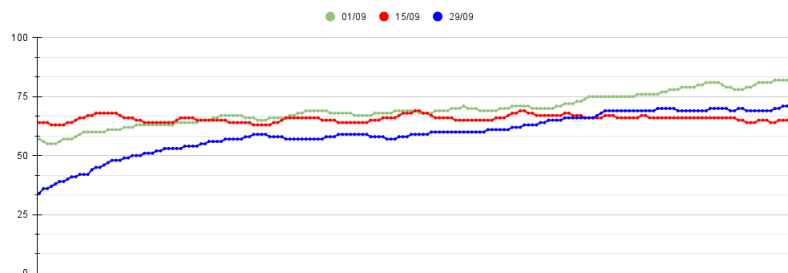
Fonte: Autoria própria (2023).

Durante estes períodos é possível perceber diversos momentos de velocidade baixa, devido ao intenso trânsito e por consequência a falta da ventilação natural do ar indo de encontro a frente do veículo e consequentemente com o radiador resfriando o veículo, portanto na maioria dos instantes a temperatura permanece perto dos 80 a 90°C. Porém observa-se alguns instantes uma queda brusca da temperatura do veículo, nos momentos em que o veículo consegue atingir uma maior velocidade, consequência de um fluxo de veículos menor. Chegando

perto dos 70°C, ou seja, aproximadamente 20°C abaixo da temperatura ideal para o funcionamento do motor, ponto ideal de máxima eficiência. Desta forma, levanta-se a hipótese de algum dos componentes do motor, responsáveis por manter a temperatura operando perto da ideal e indicada pelo fabricante, possuir algum tipo de defeito.

Em uma análise de comparação entre a variável de temperatura do líquido de arrefecimento, representada pelo gráfico 5, durante 3 dias diferentes, no mesmo cenário (segundo cenário), é possível observar que de maneira geral as temperaturas ficam longe do ideal. Apenas em um dos casos ela chega próximo aos 80°C, porém dentro do intervalo de tempo em que foi realizada a captura, aproximadamente 25 minutos, a temperatura deveria chegar próximo aos 90°C e se manter nessa faixa, sendo controlado pela válvula termostática, ventoinha e demais componentes. Mesmo que este cenário observado seja um trajeto de trânsito moderado/leve, onde o veículo não possui muitos momentos parados, é normal que receba maior ventilação natural o que tende a diminuir a temperatura do veículo.

Gráfico 5 – Evolução no decorrer dos minutos de coleta da temperatura de arrefecimento em diferentes coletas realizadas no cenário 2



Fonte: Autoria própria (2023).

Após realizada a etapa de análise, foi necessário aplicar uma validação nas investigações realizadas. Desta forma, buscou-se uma consultoria especializada em manutenção de veículos, para constatar se o levantamento feito era de fato verdadeiro ou somente uma falha de análise. Após o diagnóstico e a desmontagem da válvula termostática, constatou-se que a mesma estava travada na posição de passagem contínua de líquido de arrefecimento. Esse problema causava um refrigeração constante do motor do veículo, resultando em seu funcionamento fora da temperatura ideal nos intervalos observados. O diagnóstico foi efetuado por meio de um procedimento manual de inspeção, que permitiu identificar a causa do mau desempenho do sistema de arrefecimento. Devido ao custo elevado para a realização da manutenção, não foi realizada a substituição da peça.

5.3 Considerações sobre a análise

As situações às quais os veículos estão expostos diariamente são diversas, abrangendo desde condições climáticas até interações com outros veículos, além das condições das vias e estradas. Por essa razão, mesmo ao criar cenários que buscam manter a maior similaridade

possível, é possível obter resultados diferentes. Essa variação ressalta a importância de coletar uma ampla gama de amostras e registros.

Além das análises apresentadas, é possível obter indícios adicionais para sugerir manutenções no veículo ao utilizar outros sensores, variáveis e estabelecer correlações entre eles. Alguns veículos permitem a coleta de dados como pressão de óleo e pressão de combustível, que são pontos críticos para o funcionamento adequado do motor. Em veículos convencionais, os alertas geralmente são gerados apenas quando problemas críticos são detectados. Por exemplo, a falta de pressão de óleo pode resultar em uma má lubrificação do motor, gerando atrito excessivo e, conseqüentemente, levando a uma quebra quase imediata.

Durante a realização deste estudo de caso, a biblioteca utilizada ofereceu uma ampla lista de sensores disponíveis para coleta, como mencionado anteriormente. No entanto, foram identificadas limitações nas leituras desses sensores devido às variáveis disponibilizadas pela ECU do veículo. Essas limitações impactaram diretamente as análises realizadas, que se restringiram aos dados fornecidos pelo estudo de caso mencionado. Essa restrição foi resultado de limitações de tempo e recursos disponíveis. No entanto, é importante destacar que, mesmo com a quantidade de dados analisados, ainda existem outras situações que poderiam ser investigadas. Portanto, ao interpretar os resultados, é fundamental considerar essas limitações e, se necessário, explorar outras abordagens para obter um conjunto mais abrangente de dados relevantes.

6 CONCLUSÃO

Através do desenvolvimento deste estudo foi possível entender as diversas arquiteturas propostas por outros autores, quando se trata de problemas referentes à coleta de dados em todas as etapas e a indisponibilidade de veículos por motivos de manutenção. A proposição do modelo descrito e desenvolvido, através da prototipação e análise dos resultados obtidos pelo sistema de informação criado, evidenciou que o mesmo é capaz de cumprir o objetivo proposto de diminuir a indisponibilidade e quebra dos automóveis, prevendo necessidades de reparos. Além disso, foram evidenciados diversos desafios acerca do escopo de IoV, como diversidade de dados, quantidade de sensores, disposição dos objetos de estudo e complexidade das ferramentas disponíveis.

Este estudo traz à tona que a definição da arquitetura mais adequada para cada proposta, varia muito de acordo com os recursos e as particularidades de cada desenvolvimento e situação, seja na parte de conexão, coleta, tratamento dos dados, transmissão e todas diversas etapas incluídas em cada um dos processos. Verificamos a partir da revisão sistemática, outros trabalhos já propostos por outros autores reunindo palavras chave que são pontos fundamentais na proposição de sistemas IoV. Com isso, levantou-se as principais soluções propostas, assim como as dificuldades encontradas e suas relações entre os diferentes estudos. Esta atividade mostrou-se fundamental para um melhor desenvolvimento do modelo apresentado neste trabalho. Também, foi possível levantar uma grande variedade de temas sendo abordados pelos autores analisados.

Este estudo apresentou questões fundamentais para a continuidade do fluxo de dados, como a fragilidade em relação à instabilidade de conexão e a importância do tratamento dos dados após a coleta. Além disso, foram identificadas questões relativas ao versionamento, tanto de software quanto de hardware, que exigiram um tempo e esforço maiores para sua solução. Ao longo deste trabalho acadêmico, foi desenvolvido um modelo de arquitetura composto por várias etapas. Na etapa de coleta de dados, utilizou-se o protocolo OBD2 para extrair informações dos sensores do automóvel. O protocolo Bluetooth foi empregado para enviar os dados à etapa subsequente, na qual foram tratados, comprimidos usando o algoritmo LZMA2 e armazenados inicialmente localmente. Esses dados aguardaram uma conexão de internet para serem transmitidos via SCP para um servidor responsável pela análise dos dados. Para o armazenamento, foram adotadas duas tecnologias: SQLite no ponto inicial e Postgresql no ponto final. Na etapa de análise, várias ferramentas foram utilizadas, com destaque para Apache SuperSet e DBeaver, que foram empregadas na geração de gráficos e relatórios. A escolha dessas ferramentas e tecnologias foi baseada em critérios específicos e considerações relacionadas aos requisitos do projeto, através de construção de um protótipo que possibilitasse a exibição dos dados e suas análises, assim como os relatórios.

Observando portanto os resultados obtidos e utilizados na análise dos dados que compõem o banco de dados deste estudo, foi possível apontar pontos em que as variáveis ob-

servadas não tiveram o comportamento esperado para o funcionamento adequado do veículo monitorado, possibilitando portanto, indicar momentos de atenção e até mesmo de substituição de um componente do sistema de refrigeração do veículo utilizado para o caso de uso do estudo.

Mesmo com o caso de sucesso evidenciado neste estudo, necessita-se de uma amostragem maior para de fato garantir que sempre pode-se diminuir a indisponibilidade e quebra do veículo. Entretanto, estudos semelhantes apresentados neste documento em questão, também apontam um bom progresso no diagnóstico de sensores por meio de sistemas que contemplam diversas etapas.

Devido a limitações de recursos e tempo, aplicamos as observações dentro de alguns cenários possíveis de coleta dos dados. Estudos futuros podem contemplar um maior período de coleta, além de maior número de variáveis de sensores observados do veículo em questão e de um monitoramento em tempo real por meio de um painel digital que permita até mesmo exibir alarmes que indiquem pontos de atenção imediatos diretamente no veículo.

REFERÊNCIAS

- ABDELKADER, G.; ELGAZZAR, K.; KHAMIS, A. Connected vehicles: Technology review, state of the art, challenges and opportunities. **Sensors**, v. 21, n. 22, 2021. ISSN 1424-8220. Disponível em: <https://www.mdpi.com/1424-8220/21/22/7712>.
- BASTOS, E. **Estudo das diferenças dos requerimentos das principais legislações de on board diagnostics para padronização de testes de desenvolvimento e validação de transmissão automática de automóveis**. 2012. Trabalho de Pós-Graduação – Centro Universitário do Instituto Mauá de Tecnologia. Curso de especialização em Engenharia Automotiva.
- BERZ, D. *et al.* Comparison of lossless data compression methods. *In: . [S.l.: s.n.]*, 2015.
- BONOMI, F. **The Smart and Connected Vehicle and the Internet of Things**. 2013. WSTS, San Jose, CA, USA.
- BOOTE, D. N.; BEILE, P. Scholars before researchers: On the centrality of the dissertation literature review in research preparation. **Educational Researcher**, v. 34, n. 6, p. 3–15, 2005. Disponível em: <https://doi.org/10.3102/0013189X034006003>.
- BURGESS, L. *et al.* **Traffic Management Centers in a Connected Vehicle Environment**. [S.l.], 2013.
- BURROWS, M.; WHEELER, D. J. A block-sorting lossless data compression algorithm. *In: . [S.l.: s.n.]*, 1994.
- CHEN, D.; HAN, X.; WANG, W. Use of sqlite on embedded system. *In: 2010 International Conference on Intelligent Computing and Cognitive Informatics*. [S.l.: s.n.], 2010. p. 210–213.
- Cisco Systems, Inc. **Cisco - Global Home Page**. 1984. Online; Acessado em: 02-Outubro-2020. Disponível em: <https://www.cisco.com/>.
- Contreras-Castillo, J.; Zeadally, S.; Guerrero-Ibañez, J. A. Internet of vehicles: Architecture, protocols, and security. **IEEE Internet of Things Journal**, v. 5, n. 5, p. 3701–3709, 2018.
- CORREA, J. D. A.; PINTO, A. S. R.; MONTEZ, C. Lossy data compression for iot sensors: A review. **Internet of Things**, v. 19, p. 100516, 2022. ISSN 2542-6605. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2542660522000208>.
- Cregger, J.; Brugeman, V.; Wallace, R. **Connected Vehicle Technology Industry Delphi Survey**. [S.l.], 2012.
- Elm Electronics. **Elm Electronics**. 1998. Online; Acessado em: 02-Outubro-2020. Disponível em: <https://www.elmelectronics.com/>.
- GAMA, J.; RIBEIRO, R. P.; VELOSO, B. Data-driven predictive maintenance. **IEEE Intelligent Systems**, v. 37, n. 4, p. 27–29, 2022.
- Gerla, M. *et al.* Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds. *In: 2014 IEEE World Forum on Internet of Things (WF-IoT)*. [S.l.: s.n.], 2014. p. 241–246.

- GILLANI, M. *et al.* Data collection protocols for vanets: a survey. **Complex and Intelligent Systems**, v. 8, 01 2022.
- GU, R. Data cleaning framework: An extensible approach to data cleaning. 01 2011.
- HAO, Q. *et al.* A decision support system for integrating corrective maintenance, preventive maintenance, and condition-based maintenance. 05 2010.
- Huang, C.; Lu, R.; Choo, K. R. Vehicular fog computing: Architecture, use case, and security and forensic challenges. **IEEE Communications Magazine**, v. 55, n. 11, p. 105–111, 2017.
- INSIGHT, F. M. **Automotive Repair & Maintenance Service Market**. [S.l.], 2021.
- ISLAM, A.; HOSSAN, M. T.; JANG, Y. M. Convolutional neural network scheme-based optical camera communication system for intelligent internet of vehicles. **International Journal of Distributed Sensor Networks**, v. 14, 02 2018.
- JANKOVIC, M.; MAGNER, S. Power output monitoring for vehicles equipped with electronic throttle. *In: Proceedings of the 2001 American Control Conference. (Cat. No.01CH37148)*. [S.l.: s.n.], 2001. v. 2, p. 1310–1314 vol.2.
- JOHANSON, M.; DAHLE, P.; SÖDERBERG, A. **Remote Vehicle Diagnostics over the Internet using the DoIP Protocol**. 2011.
- KIMURA, N.; LATIFI, S. A survey on data compression in wireless sensor networks. *In: International Conference on Information Technology: Coding and Computing (ITCC'05) - Volume II*. [S.l.: s.n.], 2005. v. 2, p. 8–13 Vol. 2.
- KITCHENHAM, B. Procedures for performing systematic reviews. **Keele, UK, Keele Univ.**, v. 33, 08 2004.
- LANEY, D. **3D Data Management: Controlling Data Volume, Velocity, and Variety**. [S.l.], 2001. Disponível em: <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>.
- LI, L. Data quality and data cleaning in database applications. *In: .* [S.l.: s.n.], 2012.
- Lie, C. H.; Chun, Y. H. An algorithm for preventive maintenance policy. **IEEE Transactions on Reliability**, v. 35, n. 1, p. 71–75, 1986.
- LYCETT, M. 'datafication': Making sense of (big) data in a complex world. **European Journal of Information Systems**, v. 22, 07 2013.
- MACHADO, A.; OLIVEIRA, B. O sistema obd (on-board diagnosis). Mestrado em Automação e Sistemas no Instituto Superior de Engenharia do Porto. 2007.
- MARCORIN, W. R.; LIMA, R. C. Análise dos custos de manutenção e de não-manutenção de equipamentos produtivos. *In: .* [S.l.: s.n.], 2003.
- MCCORD, K. **Automotive Diagnostic Systems**. [S.l.]: Cartech, 2011. ISBN 1613255640.
- Contents. *In: MOBLEY, R. K. (Ed.). An Introduction to Predictive Maintenance (Second Edition)*. Second edition. Burlington: Butterworth-Heinemann, 2002, (Plant Engineering). p. v–xii. ISBN 978-0-7506-7531-4. Disponível em: <https://www.sciencedirect.com/science/article/pii/B9780750675314500002>.
- Mostafa Anwar Taie; Diab, M.; ElHelw, M. Remote prognosis, diagnosis and maintenance for automotive architecture based on least squares support vector machine and multiple classifiers.

In: 2012 IV International Congress on Ultra Modern Telecommunications and Control Systems. [S.l.: s.n.], 2012. p. 128–134.

Nkenyereye, L.; Jang, J. Integration of big data for querying can bus data from connected car. *In: 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN).* [S.l.: s.n.], 2017. p. 946–950.

OBD Facilities. **Car diagnostic scanner ELM327 OBD2 - Outils OBD Facile.** 2010. Online; Acessado em: 02-Outubro-2020. Disponível em: <https://www.obdsol.com/knowledgebase/on-board-diagnostics/what-is-obd/>.

OBD Solutions. **OBD Solutions.** 2017. Online; Acessado em: 02-Outubro-2020. Disponível em: <https://www.obdsol.com/knowledgebase/on-board-diagnostics/what-is-obd/>.

PRYTZ, R. **Machine learning methods for vehicle predictive maintenance using off-board and on-board data.** 09 2014. Tese (Doutorado) — Halmstad University, 09 2014.

RAHM, E.; DO, H. Data cleaning: Problems and current approaches. **IEEE Data Eng. Bull.**, v. 23, p. 3–13, 01 2000.

RAO, C. V. S. *et al.* Data cleaning: A framework for robust data quality in enterprise data warehouse. *In: .* [S.l.: s.n.], 2012.

Roudier, Y. van der Straatenand F. **World vehicle production reached 84.1 million in 2012.** 4 rue de Berri, F 75008 Paris, 2013.

SADIKU, M.; TEMBELY, M.; MUSA, S. Internet of vehicles: An introduction. **International Journal of Advanced Research in Computer Science and Software Engineering**, v. 8, p. 11, 02 2018.

SAE. **Society of Automotive Engineers.** 1992. Online; Acessado em: 02-Outubro-2020. Disponível em: <https://www.sae.org/>.

SALOMON, D.; MOTTA, G. **Handbook of Data Compression.** 5th. ed. [S.l.]: Springer Publishing Company, Incorporated, 2009. ISBN 1848829027.

SANTOS, A. L. V. d. **Economia de combustível com o uso de telemetria para veículos de passeio.** 2016. Trabalho de Conclusão de Curso – Universidade Federal do Rio Grande do Sul. Instituto de Informática. Curso de Engenharia de Computação.

SCHOLL, P. M.; LAERHOVEN, K. V. A multi-media exchange format for time-series dataset curation. *In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct.* New York, NY, USA: Association for Computing Machinery, 2016. (UbiComp '16), p. 715–721. ISBN 9781450344623. Disponível em: <https://doi.org/10.1145/2968219.2968278>.

SHAFI, U. *et al.* Vehicle remote health monitoring and prognostic maintenance system. **Journal of Advanced Transportation**, v. 2018, p. 1–10, 01 2018.

SHARMA, R. Big data for autonomous vehicles. *In: ____.* [S.l.: s.n.], 2021. p. 21–47. ISBN 978-3-030-65660-7.

SIQUEIRA, B.; ALBUQUERQUE, T. Avaliação de custos de manutenção em automóveis- estudo de caso. *In: .* [S.l.: s.n.], 2020. p. 2333–2353.

SUARJAYA, I. M. A. D. A new algorithm for data compression optimization. 2012.

Xu, W. *et al.* Internet of vehicles in big data era. **IEEE/CAA Journal of Automatica Sinica**, v. 5, n. 1, p. 19–35, 2018.

Yang, F. *et al.* An overview of internet of vehicles. **China Communications**, v. 11, n. 10, p. 1–15, 2014.

YEFERNY, T.; HAMAD, S. Vehicular ad-hoc networks: Architecture, applications and challenges (international journal of computer science and network security). 03 2020.

YLIJOKI, O.; PORRAS, J. Perspectives to definition of big data: A mapping study and discussion. **Journal of Innovation Management**, v. 4, p. 69–91, 05 2016.

ZHANG, M. *et al.* Carstream: an industrial system of big data processing for internet-of-vehicles. **Proceedings of the VLDB Endowment**, v. 10, p. 1766–1777, 08 2017.

APÊNDICES

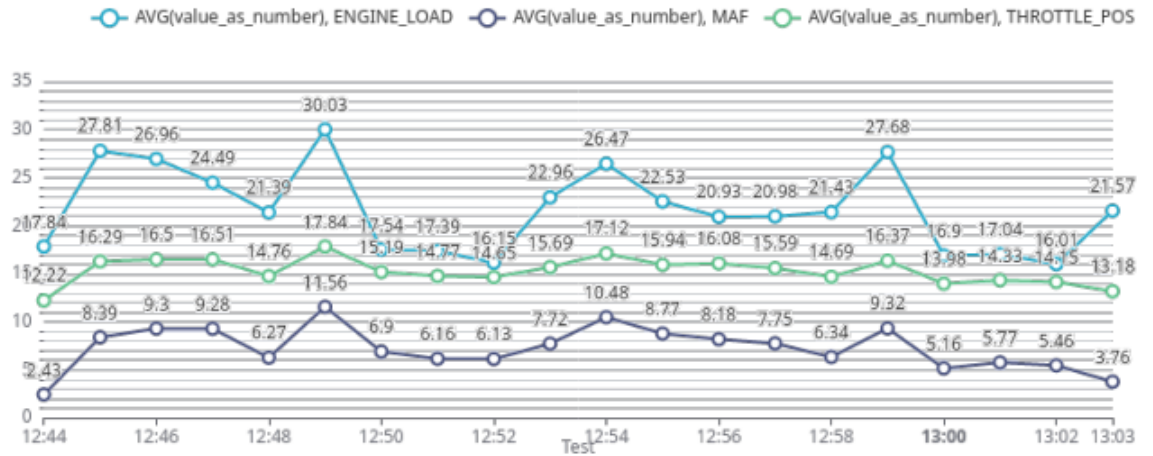
APÊNDICE A – Relatórios

Figura 19 – Relatório 1 gerado através do Apache Superset

Relatório ☆ Draft EDIT DASHBOARD ...



Discussao1 (/explore/?dashboard_page_id=i9zjW4lBr&slice_id=3)



Bars (/explore/?dashboard_page_id=i9zjW4lBr&slice_id=4)

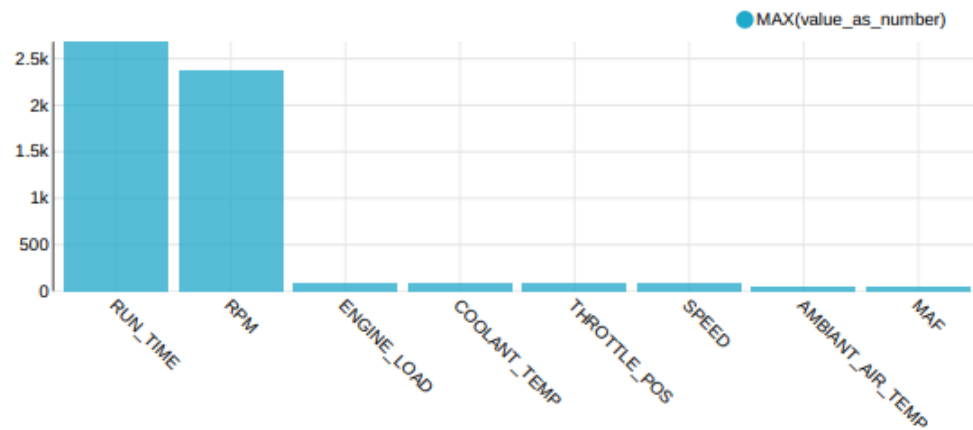
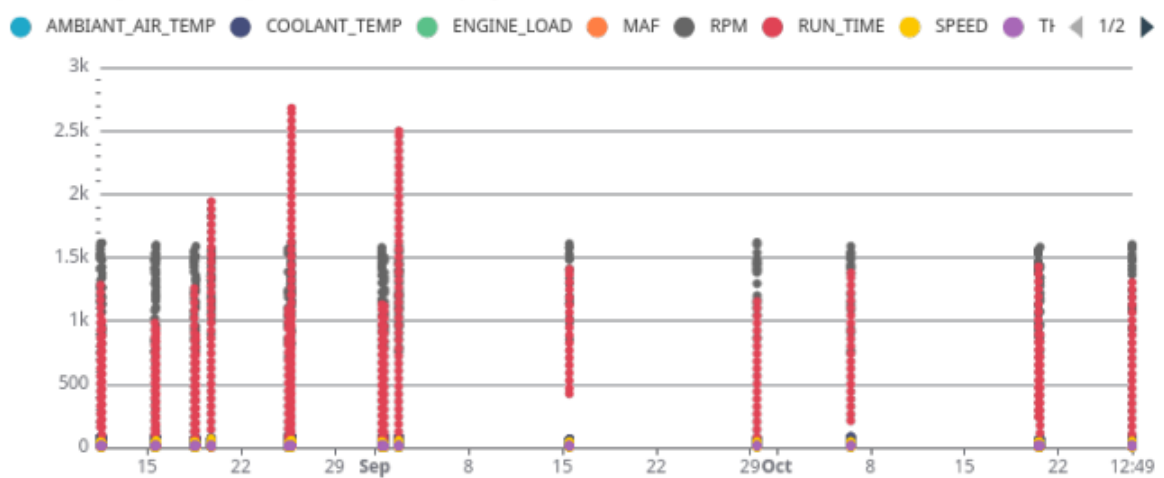


Figura 20 – Relatório 2 gerado através do Apache Superset

Sensores por dia (/explore/?dashboard_page_id=TrvTF1E99i&slice_id=6)



Num acontecimentos (/explore/?dashboard_page_id=TrvTF1E99i&slice_id=5)

key ↕	AVG(value_as_number) ↕
RPM	1.27k
RUN_TIME	809.55
COOLANT_TEMP	72.93
AMBIANT_AIR_TEMP	33.3
SPEED	24.09
ENGINE_LOAD	22.87
THROTTLE_POS	15.47
MAF	7.35