LUCIO FLAVIO VISMARI

Garantia da segurança crítica em sistemas complexos: uma abordagem orientada a riscos para o gerenciamento de recursos de comunicação em Sistemas de Transporte Inteligentes Cooperativos (C-ITS)

São Paulo
2023

LUCIO FLAVIO VISMARI

ENGENHEIRO ELETRICISTA PELA ESCOLA POLITÉCNICA DA UNIVERSIDADE DE SÃO PAULO, 2002.
MESTRE EM ENGENHARIA ELÉTRICA PELA ESCOLA POLITÉCNICA DA UNIVERSIDADE DE SÃO PAULO, 2007.

Garantia da segurança crítica em sistemas complexos: uma abordagem orientada a riscos para o gerenciamento de recursos de comunicação em Sistemas de Transporte Inteligentes Cooperativos (C-ITS)

Versão Corrigida

Tese apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do Título de Doutor em Ciências.

Área de Concentração:
Engenharia de Computação (3141)

Orientador: Prof. Dr. João Batista Camargo Júnior

São Paulo
2023

Este exemplar foi revisado e alterado em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 16 de novembro de 2023

Documento assinado digitalmente
LUCIO FLAVIO VISMARI
Data: 16/11/2023 11:57:05-0300
Verifique em https://validar.iti.gov.br

Assinatura do autor: _____

Assinatura do orientador: ____

Documento assinado digitalmente
JOAO BATISTA CAMARGO JUNIOR
Data: 16/11/2023 13:36:28-0300
Verifique em https://validar.iti.gov.br

Catalogação-na-publicação

VISMARI, LUCIO FLAVIO
Garantia da segurança crítica em sistemas complexos: uma abordagem orientada a riscos para o gerenciamento de recursos de comunicação em Sistemas de Transporte Inteligentes Cooperativos (C-ITS) / L. F. VISMARI – versão corrig. -- São Paulo, 2023.
183 p.

Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Computação e Sistemas Digitais.

1.Computação Aplicada 2.Engenharia de Sistemas de Computação 3.Sistemas Colaborativos 4.Segurança de Tráfego 5.Sistemas Inteligentes de Transporte I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Computação e Sistemas Digitais II.t.

VISMARI, L. F. Garantia da segurança crítica em sistemas complexos: uma abordagem orientada a riscos para o gerenciamento de recursos de comunicação em Sistemas de Transporte Inteligentes Cooperativos (C-ITS). 2023. Tese (Doutorado) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2023.

Aprovado em: 27/10/2023

Banca Examinadora

Prof. Dr.	João Batista Camargo Júnior
Instituição:	EP - USP
Prof. Dr.	Gilberto Francisco Martha de Souza
Instituição:	EP - USP
Profa. Dra.	Selma Shin Shimizu Melnikoff
Instituição:	EP - USP
Prof. Dr.	Ricardo Caneloi dos Santos
Instituição:	UFABC
Prof. Dr.	Marcelo José Ruv Lemes
Instituição:	EMBRAER

Aos meus filhos Ana, Helena e Leonardo.

AGRADECIMENTOS

O período de desenvolvimento do projeto de pesquisa que resultou nesta tese foi, sem dúvida, o mais intenso da minha vida. Nestes anos, vivenciei o impacto que mudanças imprevisíveis no sistema 'vida' podem causar na sua capacidade de lidar com situações potencialmente danosas. Entendi que, como sistema complexo *de facto*, a melhor forma de lidar com a vida não é tentando controlar seus elementos de forma centralizada, mas por meio de mecanismos que orientem a evolução deste sistema para que se autoorganize da melhor forma possível. Ter controle sobre a vida é falácia. Mas, podemos ser o mecanismo que orquestra os resultados.

Dada a quantidade de interações que ocorreram neste período para que fosse possível chegar até aqui me obrigaria a agradecer a uma lista muito extensa de pessoas. Por isso, a única certeza que tenho é que, se tentasse ser completo nesta tarefa, esqueceria de muita gente. Portanto, de forma ampla, agradeço muito a todos que fizeram parte de minha vida nestes últimos anos, seja em nível profissional ou pessoal. Saibam que cada um de vocês contribuiu com o resultado aqui apresentado!

Além disso, registro aqui meus agradecimentos e considerações:

Aos meus filhos, Ana, Helena e Leonardo, minha esposa, Simone, e meus pais, Dorival e Liderce. Obrigado pela paciência e apoio, e perdão por não estar presente como deveria. Enfim, etapa que finaliza. Vida que se retoma.

Ao meu orientador, Prof. João Batista, obrigado pela compreensão, apoio, insistência e confiança, mesmo nos muitos momentos de explicável, mas não justificável, procrastinação acadêmica desta jornada.

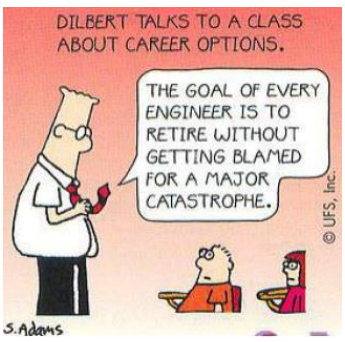
Aos meus amigos diretamente relacionados a estes resultados, Alexandre Nascimento e Caroline Molina, obrigado pela paciência, apoio, motivação e suporte.

Ao meu amigo Ricardo Gimenes, meu obrigado pelos ouvidos, motivação, tempo e energia.

Aos membros da banca examinadora, muito obrigado pelos apontamentos e contribuições.

E, por fim, agradeço a você, leitor. Seu interesse por esta obra possibilita sua disseminação, motivando outros a continuar esta jornada. Sempre há muito ainda a ser feito ...

EPÍGRAFE



“Dilbert fala a uma classe a respeito de opções de carreira.
A meta de todo engenheiro é se aposentar ***sem ser culpado por uma catástrofe****.*”
(ADAMS, 1993, tradução nossa, negrito nosso)

RESUMO

VISMARI, L. F. Garantia da segurança crítica em sistemas complexos: uma abordagem orientada a riscos para o gerenciamento de recursos de comunicação em Sistemas de Transporte Inteligentes Cooperativos (C-ITS). 2023. Tese (Doutorado) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2023.

Sistemas de engenharia são projetados para atender necessidades específicas, benéficas à sociedade. Contudo, quando não atendem ao comportamento desejado, alguns destes sistemas podem produzir perdas inaceitáveis, como mortes, danos à saúde ao meio-ambiente e à propriedade. Desta forma, para que os benefícios esperados superem os riscos de perdas, estes sistemas críticos devem ser seguros ao longo de todo seu ciclo de vida. Um sistema pode ser considerado seguro se ele não causar danos enquanto opera em condições normais, bem como apresentar um nível aceitável de risco de segurança quando operar em situações anormais previstas. Desta forma, as situações inseguras que possam vir a ocorrer ao longo da vida útil do sistema devem ser sistematicamente identificadas e incorporadas por sua especificação. Assim, quando ocorrerem, o sistema poderá lidar com a situação de forma a manter os riscos de segurança em níveis aceitáveis. Por outro lado, caso alguma situação insegura não tenha sido prevista durante o projeto do sistema, sua manifestação terá potencial de causar danos devido ao sistema não estar preparado para tratá-la, reduzindo a capacidade de garantir a segurança deste sistema. Esse limitação é acentuada pela evolução no paradigma dos sistemas críticos em segurança. Tecnologias de comunicação e informação (ICT), sobretudo comunicação colaborativa e Inteligência Artificial, estão tornando estes sistemas inerentemente complexos. Nestes sistemas complexos de engenharia, as limitações intrínsecas na capacidade de se prever situações que possam ocorrer durante sua operação – seja por modificação imprevisíveis no sistema ou no seu contexto de operação, seja por situações latentes não previstas – produzem especificações incompletas, o que compromete a garantia da segurança destes sistemas. Portanto, o objetivo deste trabalho é obter uma abordagem que permita lidar, durante a operação, com situações inseguras imprevisíveis ou/e imprevistas no projeto, possibilitando garantir níveis aceitáveis de risco de segurança durante a operação e contribuindo no processo de garantia de segurança de sistemas complexos de engenharia. Nesta abordagem, o processo de especificação de requisitos de segurança é incorporado à operação do sistema, buscando por situações onde o sistema atende a especificação vigente, mas o nível de risco de segurança observado na aplicação é inaceitável. Nestas condições, a configuração observada no sistema é reclassificada como uma situação insegura. Então, ela é incorporada à especificação do sistema para que, caso ocorra, não exponha os envolvidos à uma condição insegura. Como resultado, foi proposta uma abordagem que permitiu, por meio do gerenciamento de seus recursos de comunicação, garantir níveis aceitáveis de risco de segurança em uma aplicação crítica no contexto de C-ITS. Para isso, ao identificar uma situação insegura (configuração de parâmetros de comunicação do sistema que expunha a aplicação a níveis inaceitáveis de risco de segurança), a abordagem busca por uma configuração de recursos de comunicação que pemite recuperar um nível mínimo aceitável de risco de segurança. Portanto, além de uma abordagem, espera-se que este trabalho possa contribuir com a Engenharia de Segurança de Sistemas Complexos de Engenharia, mostrando que a mudança de mentalidade pode ser caminho para lidar uma mudança de paradigma de sistema de engenharia.

Palavras-Chave: segurança crítica (safety), sistemas complexos, sistemas de transporte inteligentes.

ABSTRACT

VISMARI, L. F. Safety critical assurance in complex systems: a risk-oriented approach to communication resources management in Cooperative Intelligent Transport Systems (C-ITS). 2023. Thesis (Doctorate) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2023.

Enginered systems are designed to meet specific needs, beneficial to society. However, when they do not meet the specified behavior, some of these systems can produce unacceptable losses, such as deaths, injuries and environment and property damages. Thus, in order for the expected benefits outweigh the risks, critical systems must be safe throughout their entire life cycle. A system is safe whether it does not cause damage while operating under normal conditions, and presents an acceptable level of safety risk when operating under anticipated abnormal situations. In this way, unsafe situations that may occur over the system lifetime must be systematically identified and incorporated into its specification. Thus, when they occur, the system will be able to deal with the situation in order to keep safety risks at acceptable levels. On the other hand, if any unsafe situation was not foreseen during the system design, its manifestation will have the potential to cause damage due to the system not being prepared to deal with it, reducing the capacity to guarantee system safety. This limitation is stressed by the evolution in the paradigm of safety-critical systems. Information and Communication Technologies (ICT), especially collaborative communication and Artificial Intelligence, are making these systems inherently complex. In these complex engineered systems, the intrinsic limitations in the ability to predict situations that may occur during their operation – whether due to unpredictable changes in the system or in its operating context, or due to unforeseen latent situations – produce incomplete specifications, which compromises the ensuring the systems safety. Therefore, the objective of this work is to obtain an approach that allows dealing, during the operation, with unpredictable or/and unforeseen unsafe situations in the project, making it possible to guarantee acceptable levels of safety risks during the operation and contributing to the process of systems safety assurance of complex engineered systems. In this approach, the safety requirements specification process is incorporated into the system operation, looking for situations where the system meets the current specification, but the level of safety risk observed in the application is unacceptable. Under these conditions, the configuration observed in the system is reclassified as an unsafe situation. Then, it is incorporated into the system specification so that, if it occurs, it does not expose those involved to an unsafe condition. As a result, an approach was proposed that allowed, through the management of its communication resources, to guarantee acceptable levels of safety risks in a critical application in the context of C-ITS. For this, when identifying an unsafe situation (configuration of system communication parameters that exposed the application to unacceptable levels of safety risk), the approach sought a communication resources configuration that allowed recovering a minimum acceptable level of safety risk. Therefore, in addition to an approach, it is expected that this thesis can contribute to the Safety Engineering of Complex Engineered Systems, showing that a change in mindset can be the way to deal with a paradigm shift in systems engineering.

Keywords: safety, complex systems, cooperative inteligent transportation systems.

LISTA DE ILUSTRAÇÕES

Figura 1 – Interação Elementos-Sistema-Ambiente 32
Figura 2 – Diagrama de entidade-relacionamento entre os conceitos relacionados a sistema 33
Figura 3 – Sistema de interesse, estrutura e nível de abstração 34
Figura 4 – Exemplo um ciclo de vida em 'V' 36
Figura 5 – Relação entre sistema e aplicação crítica 38
Figura 6 – Evolução dos sistemas de sinalização e controle no domínio de aplicação metroviário 39
Figura 7 – Ciclo evolutivo dos sistemas e aplicações 39
Figura 8 – Relação entre CPS e domínio de aplicação crítica em segurança 40
Figura 9 – Exemplos de sistemas críticos, seus ciclos de vida e eventos catastróficos 41
Figura 10 – Abordagens de garantia de segurança na SSE 42
Figura 11 – Relação causal entre falha-erro-disfunção (*fault-error-failure*) 45
Figura 12 – Universo de estados do sistema e propagação de falhas 46
Figura 13 – Ciclo de vida do sistema crítico em segurança (especificação e validação) 49
Figura 14 – Características dos sistemas complexos 54
Figura 15 – Modelo conceitual de referência para a ATN 58
Figura 16 – Visão física do ARC-IT: classes de objetos físicos e interconexões 60
Figura 17 – Relação entre grupos de conceitos em sistemas autônomos 65
Figura 18 – Uso de compensação funcional na garantia da segurança crítica 69
Figura 19 – Arquitetura de sistema crítico (a.) e a comunicação (C_i) como elemento do sistema (b) 72
Figura 20 – Fluxo de dados entre componentes do sistema (incluindo Ci) e Abordagem Proposta 73
Figura 21 – Modelo de referência para CAV 74
Figura 22 – Arquitetura de implementação da Abordagem Proposta 76
Figura 23 – Relacionamento entre elementos do cenário-base (CAV, CSP e CCO) 79
Figura 24 – Ambiente de tráfego do cenário-base 79
Figura 25 – Controle Automatizado de Tráfego de Cruzamento (CaTraCa) 80
Figura 26 – Gerenciamento da ocupação temporal da região de cruzamento (AV1 e AV2) 81
Figura 27 – Fluxo temporal de dados trocados entre elementos do cenário-base 82

LISTA DE ILUSTRAÇÕES

Figura 28 – Exemplo de envelope operacional de comunicação em aplicações críticas 85
Figura 29 – Envelopes operacionais (Máxima Latência Fim-a-Fim [ms] x Taxa de Transmissão mínima [Hz]) para dois cenários de direção avançada utilizando comunicação V2X 85
Figura 30 – Arquitetura do ambiente computacional para análise de segurança de RTS. 89
Figura 31 – Modelo detalhado do cenário-base do estudo de caso 90
Figura 32 – Componentes principais do Módulo FTS e componente de automação. 91
Figura 33 – Exemplo de implementação do cenário-base no SUMO (unidimensional) 92
Figura 34 – Exemplo de regiões de interseção entre trajetórias 93
Figura 35 – Exemplo de regiões de interseção e colisão entre vias 1 e 2 94
Figura 36 – Estrutura de dados gerado a cada evento de cruzamento de AVs em um ciclo de simulação 95
Figura 37 – Estrutura de dados do registro gerado por um ciclo de simulação 95
Figura 38 – Ampliação esperada dos Envelopes operacionais para dois cenários de direção avançada utilizando comunicação V2X 97
Figura 39 – Relacionamento entre elementos do cenário-base e a Abordagem Proposta 98
Figura 40 – Arquitetura da Abordagem Proposta para o Estudo de Caso 99
Figura 41 – Detalhe do processo de geração de estatísticas relacionadas à “distância mínima entre veículos” 101
Figura 42 – Implementação da Abordagem Proposta no cenário-base 103
Figura 43 – Fluxo temporal de dados entre cenário-base e Abordagem Proposta . 104
Figura 44 – Processo de simulação do cenário-base 105
Figura 45 – Algoritmo de automação do processo de simulação 106
Figura 46 – Tipos e exemplos de técnicas de ML (baseado em MATHWORKS (2023)) 107
Figura 47 – Algoritmo de emulação da Abordagem Proposta 109
Figura 48 – Algoritmo de AtualizaConfiguracao() 110
Figura 49 – Resultados da CAMPANHA II (μ[dmin] x Latência | Taxa de Transmissão) 122
Figura 50 – Resultados da CAMPANHA II (μ[dmin] x Taxa de Transmissão | Latência) 123

LISTA DE ILUSTRAÇÕES

Figura 51 – Resultados da CAMPANHA II (σ[dmin] x Taxa de Transmissão | Latência).....123
Figura 52 – Resultados da CAMPANHA II (Duração x Taxa de Transmissão).....124
Figura 53 – Resultados da CAMPANHA II (Duração x Latência).....125
Figura 54 – Resultados da CAMPANHA III (%Colisão x Latência | Taxa de Transmissão).....126
Figura 55 – Resultados da CAMPANHA III (%Colisão x Latência | Taxa de Transmissão) – destaque.....127
Figura 56 – Regras de decisão induzidas nos Modelos 1 e 2.....129
Figura 57 – Teste de inserção do Modelo #1, A (1.000ms; 1/5Hz).....131
Figura 58 – Teste de inserção do Modelo #1, B (2.000ms; 1/5Hz).....131
Figura 59 – Teste de inserção do Modelo #1, C (2.500ms; 1/5Hz).....132
Figura 60 – Ampliação observada para o envelope operacional do cenário-base..139

LISTA DE TABELAS

Tabela 1 – Definições de conceitos diretamente relacionados a sistemas críticos ...37
Tabela 2 – Atributos de dependabilidade de um sistema (AVIZIENIS *et al.*, 2004)...44
Tabela 3 – Definição de dependabilidade (IEC ref. 192 01 02) 47
Tabela 4 – Cenário de Desempenho da Comunicação V2X a serem avaliados (ETSI, 2022) 86
Tabela 5 – Resultados das campanhas de simulação 1 (2s, 200m) e 2 (3s, 200m)115
Tabela 6 – Resultados das campanhas de simulação 1 (2s, 200m) e 3 (2s, 100m)116
Tabela 7 – Resultados da simulação (CAMPANHA I.) 119
Tabela 8 – Resultados da CAMPANHA I – Cenários A, B e C.....120
Tabela 9 – Resultados da simulação (CAMPANHA II.) 121
Tabela 10 – Resultados da CAMPANHA II – Cenários A, B e C.....122
Tabela 11 – Resultados da simulação (CAMPANHA III.) 126
Tabela 12 – Definição dos limiares de classes para aplicação de AI/ML 127
Tabela 13 – Resultados da CAMPANHA III – classificação (**S**eguro; **I**nseguro) 128
Tabela 14 – Avaliação dos Modelos #1 e #2 (resultados) 130
Tabela 15 – Resultados dos testes de inserção, Modelo #1 132

LISTA DE SIGLAS

3GPP	*3rd Generation Partnership Project*
AI/ML	*Artificial Intelligence / Machine Learning* (Inteligência Artifcial / Aprendizado de Máquina)
ATN	*Aeronautical Telecommunications Network* (Rede de Telecomunicações Aeronáuticas)
AV	*Autonomous Vehicle* (Veículo Autônomo)
BER	*Bit Error Rate* (Taxa de Erro de Bit)
BSM	*Basic Safety Message* (Mensagem Básica de Segurança)
CAM	*Cooperative Awarenness Message* (Mensagem de Consciência Cooperativa)
CATraCa	Controle Automatizado de Trafego de Cruzamento
CAV	*Connected Autonomous Vehicle* (Veículo Autônomo Conectado)
CCO	Centro de Controle Operacional
C-ITS	*Cooperative ITS* (ITS Cooperativo)
CNS/ATM	*Communication, Navigation, Surveillance / Air Traffic Management* (Comunicação, Navegação, Vigilância / Gerenciamento de Tráfego Aéreo)
CPS	*Cyber-Physical System* (Sistema Ciberfísico)
CoES	*Complex Engineered Systems* (Sistemas Complexos de Engenharia)
CoSE	*Engenharia de Sistemas Complexos* *(Complex Systems Engineering)*
COTS	*Commercial-of-the-Shelf*
CSP	*Communication Service Provider* (Provedor de Serviços de Comunicação)

LISTA DE SIGLAS

DP(x)	Desvio-Padrão da variável ‘x’
DSRC	*Dedicated Short-Range Communications* (Comunicações Dedicadas de Curto Alcance)
E/E/PES	*Electric/Electronic/Programable Electronic Systems*
ES	*End System* (Sistema Final)
ETSI	*European Telecommunications Standards Institute* NTHSA (Instituto de Normas Européias de Telecomunicações)
EU	*End User* (Usuário Final)
EUA	Estados Unidos da América
FTS	*Fast Time Simulation* (Simulação em Tempo Acelerado)
HAV	*Highly Automated Vehicles* (Veículos altamente Automatizados)
NTHSA	*National Highway Traffic Safety Administration* (Administração Nacional de Segurança Rodoviária, EUA)
ICT	*Information and Communication Technologies* (Tecnologias de Informação e Comunicação)
IEC	*International Electrotechnical Commission*
IEEE	*Institute of Electrical and Electronic Engineers* (Instituto dos Engenheiros Eletricistas e Eletrônicos)
ITS	*Intelligent Transportation Systems* (Sistemas de Transporte Inteligentes)
ISO	*International Organization for Standardization*
LTE	*Long Term Evolution* (padrão para comunicação de banda larga sem fio)
ML	*Machine Learning* (Aprendizado de Máquina)
NASEM	*National Academy of Sciences, Engineering and Medicine (USA)*
Omnet++	*Objective Modular Network Testbed in C++*
OACI	Organização da Aviação Civil Internacional (International Civil Aviation Organization - ICAO)

LISTA DE SIGLAS

OoI	*Object of Interest* (Objeto de Interesse)
OSI	*Open System Interconnection reference model*
PCS	Processo Crítico em Segurança
PoC	*Proof of Concept* (Prova de Conceito)
QoS	*Quality of Service* (Qualidade de Serviço)
RC	Região de Colisão
RCt	Região de Controle
RCz	Região de Cruzamento
RFC	*Request for Comments*
RSU	*RoadSide Unit* (Unidade à margem da via)
ReTS	*Real Time Simulation* (Simulação em Tempo Acelerado)
RTS	Road Transportation Systems (Sistemas de Transporte Rodoviário)
SAE	Sociedade dos Engenheiros Automotivos
SLA	*Service Level Agreement* (Acordo de Nível de Serviço)
SSE	*System Safety Engineering* (Engenharia de Segurança de Sistemas)
Sumo	*Simulation of Urban MObility*
TraCI	*Traffic Control Interface*
TRB	*Transportation Research Board* (Conselho Nacional dos Transportes, NASEM)
TS	*Technical Specification* (Especificação Técnica)
V2I	*Vehicle to Infrastructure (communication)*
V2N	*Vehicle to Network (communication)*
V2P	*Vehicle to Pedestrian (communication)*

LISTA DE SIGLAS

V2V	*Vehicle to Vehicle (communication)*
V2X	*Vehicle to Everything (communication)*
VEINS	*Vehicles in Network Simulation*
VRU	*Vulnerable Roadway Users* (Usuários da Via Vulneráveis)

SUMÁRIO

1. INTRODUÇÃO 19

1.1. Objetivos 24

1.2. Justificativa 26

1.3. Organização do trabalho 29

2. CONCEITOS E FUNDAMENTOS 31

2.1. Sistemas críticos em segurança 31

2.2. Sistemas complexos de engenharia (CoES) 50

2.3. Sistemas de Transporte Inteligente Cooperativos (C-ITS) 57

3. UMA ABORDAGEM ORIENTADA A RISCOS PARA O GERENCIAMENTO DE RECURSOS EM C-ITS 62

3.1. CoES e os sistemas autônomos cooperativos 62

3.2. Garantia de segurança crítica em sistemas autônomos cooperativos 65

3.3. Compensação funcional e a garantia de segurança crítica em CoES 69

3.4. A Abordagem Proposta no contexto dos C-ITS 73

4. ESTUDO DE CASO NO CONTEXTO C-ITS 77

4.1. O cenário-base de tráfego rodoviário no contexto C-ITS 78

4.1.1. Definição do cenário-base de tráfego rodoviário 78

4.1.2. Implementação do cenário-base no ambiente computacional virtual 88

4.2. A Abordagem Proposta para o cenário-base 96

4.2.1. Definição da Abordagem Proposta orientada pelo cenário-base 96

4.2.2. Implementação e avaliação da Abordagem Proposta 104

5. RESULTADOS OBTIDOS E DISCUSSÃO 112

5.1. Resultados obtidos 112

5.1.1. Definição das Configurações de Simulação (CS) 113

5.1.2. Comportamento da segurança crítica *vs.* comunicação V2X 118

5.1.3. Desenvolvimento do Modelo Executivo 125

5.1.4. Verificação da efetividade do Modelo Executivo no cenário-base 130

5.2. Discussão dos resultados 134

6. CONCLUSÕES 141

6.1. Conclusões 141

6.2. Contribuições 144

6.3. Trabalhos futuros 147

6.4. Considerações finais 149

REFERÊNCIAS 150

SUMÁRIO

APÊNDICE I. MÉTODO DE CÁLCULO DA DISTÂNCIA MÍNIMA ENTRE OBJETOS EM CENÁRIOS DE MOVIMENTO GUIADO.....157

ANEXO I. INFERÊNCIA DE MODELOS EXECUTIVOS UTILIZANDO APRENDIZADO DE MÁQUINA EM INTELIGÊNCIA ARTIFICIAL.....169

1. INTRODUÇÃO

"*Scientists study the world as it is, engineers create the world that never has been.*"
(Theodore von Kármán, 1881-1963)[1]

Sistemas de engenharia são projetados, desenvolvidos e operados para atender necessidades específicas, prestar serviços e executar tarefas necessárias à vida cotidiana. Quando um sistema não é capaz de realizar os requisitos desejados e projetados, a missão para a qual foi concebido não pode ser devidamente cumprida. O funcionamento anormal de muitos sistemas do dia-a-dia tem consequências menores, não prejudicando significativamente as pessoas envolvidas ou o ambiente, causando, no máximo, irritação ou decepção.

No entanto, o funcionamento anormal de alguns sistemas pode levar a consequências socialmente inaceitáveis, com perdas catastróficas. Sistemas que podem causar danos à vida humana e ao ambiente no qual está inserido são definidos como sistemas críticos em segurança (SOMMERVILLE, 2011). Sistemas de supervisão e controle de reatores em usinas nucleares – onde seu funcionamento anormal pode levar ao vazamento de elementos radioativos, causando mortes, ferimentos e danos ambientais – e sistemas de frenagem em veículos – cujo funcionamento anormal pode levar a colisões e a atropelamentos – são bons exemplos de sistemas críticos em segurança.

Um estado operacional ou condição na qual um sistema pode expor as pessoas e o meio ambiente a um acidente – evento não intencional que produz perdas inaceitáveis – é definido como um estado ou condição insegura (ou perigosa). Estados inseguros violam a propriedade (objetivos) de segurança crítica de um sistema e, portanto, os riscos relacionados à segurança devem ser minimizados a níveis considerados aceitáveis por todos os envolvidos e interessados. Os sistemas nos quais os riscos de perdas superam os benefícios para as partes interessadas tendem a ser retirados de operação – ou mesmo não serem desenvolvidos, pois são considerados inseguros dentro do contexto no qual foram concebidos.

[1] https://www.nsf.gov/news/special_reports/medalofscience50/vonkarman.jsp

Assim, a manutenção de níveis de risco aceitáveis relacionados à operação de sistemas críticos em segurança demanda abordagens robustas e bem estabelecidas de garantia de segurança (ALMEIDA JR, 2003), aplicadas ao longo de todo o ciclo de vida dos sistemas onde problemas (condições adversas para o funcionamento desejado) possam ser introduzidos e se manifestarem de forma imprevista ao longo de seu ciclo de vida. A garantia de segurança de um sistema é multidimensional, abrangendo diversas dimensões no ciclo de vida do sistema, entre elas a cultura de segurança dos desenvolvedores e dos operadores; os procedimentos de desenvolvimento e operação; as características intrínsecas do produto, os processos de desenvolvimento, implementação, operação e manutenção; entre outros (CAMARGO JUNIOR, 2002).

As questões relacionadas à segurança crítica de sistemas – sobretudo as abordagens de garantia de segurança – são tratadas pela Engenharia de Segurança de Sistemas (SSE – *System Safety Engineering*), cujo objetivo principal é prevenir acidentes (HARDY, 2010), identificando e eliminando condições inseguras (perigosas) e reduzindo, a níveis aceitáveis, os riscos de segurança associados ao sistema. Atualmente, a engenharia de segurança de sistemas possui um arcabouço consolidado de princípios, processos e métodos, aplicados ao longo de todo o ciclo de vida dos sistemas.

Entre suas atribuições, é responsabilidade da SSE garantir – e demonstrar (justificar) às partes interessadas, sobretudo autoridades de segurança ou/e certificadoras responsáveis por autorizar a operação do sistema – que um sistema é seguro. Um sistema pode ser considerado seguro se, para uma determinada aplicação e em um determinado ambiente operacional, ele não irá causar adversidades (perdas, acidentes, catástrofes) aos envolvidos ou a si mesmo enquanto estiver operando em condições normais; possuir um risco aceitável quando operando em condições anormais causadas por falhas; bem como prover proteção no caso de adversidades previstas (ERICSON II, 2011)).

Domínios de aplicação críticas em segurança – como energia nuclear, petroquímica e transportes – possuem abordagens de garantia de segurança robustas e bem estabelecidas, muitas delas regulamentadas. Estas abordagens, cujo cumprimento é mandatório – quando regulamentadas, por meio do atendimento compulsório a bases normativas – em diversos domínios de aplicação em muitos países, regem o ciclo de vida destes sistemas. Mais especificamente, regem a forma como devem ser especificados, projetados, verificados,

validados, produzidos e operados, mantidos e até mesmo descartados. Portanto, por princípio, estas abordagens deveriam ser suficientes para garantir baixos níveis de risco de segurança, tornando muito raras as ocorrências de incidentes e de acidentes.

Contudo, apesar da extensa e (quase) sempre mandatória base normativa de abordagens bem estabelecidas e da existência de uma disciplina da engenharia dedicada à garantia da segurança de sistemas, acidentes e condições inseguras (que, quando identificadas, geram processo de *Recalls* dos sistemas e produtos afetados) continuam ocorrendo a taxas que podem vir a comprometer a percepção de risco das partes envolvidas e interessadas, sobretudo seus usuários, comprometendo a permanência da operação destes sistemas. Por exemplo, tem-se observado um aumento significativo no número de *Recalls* de produtos em diversos setores produtivos, em especial no setor automotivo. Problemas de segurança relacionados a softwares e novas tecnologias embarcadas, sobretudo Tecnologias de Informação e Comunicação (ICT – *Information and Communication Technologies*), são causadores de uma parcela significativa destes *recalls*.

Uma investigação preliminar (VISMARI; CAMARGO JUNIOR, 2012) discutiu a hipótese de que as abordagens de garantia de segurança preconizadas pelas bases normativas (regulatórias) atuais seriam inadequadas para lidar com características de complexidade sistêmica[2] introduzidas pelas novas tecnologias e pelas demandas de novos serviços e funcionalidades. Consequentemente, estas abordagens viriam a produzir, em escala crescente, sistemas com níveis de segurança reais (ou seja, observados durante a operação) abaixo dos níveis esperados pelo desenvolvedor e certificador, bem como aceitos pela autoridade de segurança[3].

---

[2] Não houve uma definição formal de 'sistema complexo' em (VISMARI; CAMARGO JUNIOR, 2012). Considerou-se 'sistemas complexos' como:

> "*... increasingly composed of interconnected parts that as whole present one or more properties (behavior among the possible properties) not obvious from the properties of the individual parts (this property is called "**emergence**"). System complexity arises from the interaction between two or more components of a system, and such interactions lead to a system behavior that is difficult to determine by analyzing their components in isolation (by functional hierarchical decomposition). The cause-effect relationships of problems are not evident, and coupled systems are prone to failure propagation.* (…)."

[3] Devido às características intrínsecas do seu processo de elaboração (KNIGHT, 2014), normas e regulamentos sempre estão defasadas em relação ao estado-da-arte das tecnologias disponíveis. Como forma de contornar esta limitação inerente, as autoridades de regulação e certificação têm mecanismos para lidar com questões não tratadas pelas normas vigentes. Por exemplo, no setor aeronáutico, a autoridade brasileira (ANAC) possui a "Ficha de Controle de Assuntos Relevantes" (FCAR); a autoridade norte-americana (FAA), o *Issue Paper*; a autoridade europeia (EASA), o *Certification Review Item* (CRI).

Particularmente, os transportes são um dos domínios críticos em segurança significativamente impactados pelos avanços nas Tecnologias de Informação e Comunicação (ICT). Dentro desse contexto, há o desafio no uso de técnicas de Inteligência Artificial (AI – *Artificial Intelligence*), as quais possibilitam que máquinas (sistemas computacionais) percebam o ambiente – bem como tomem decisões – sem a necessidade de intervenção humana. Estas funcionalidades têm sido utilizadas na implementação de processos automatizados nos veículos, contemplando desde mecanismos de assistência automatizada até sua condução completamente autônoma (SAE, 2014).

Além do impacto no paradigma dos elementos veiculares, avanços nas ICT têm promovido a evolução nos paradigmas dos sistemas de transporte. Fazendo uso integrado das ICTs – sobretudo redes de comunicação de dados, o paradigma de transportes tem evoluído para um conceito de alto nível de automação. Neste novo paradigma, os elementos do sistema são nós (ES – *End Systems*) de uma rede de comunicação de dados. Serviços e funcionalidades (incluindo aquelas críticas em segurança) são implementados e providos entre seus sistemas finais (ES) na camada de aplicação de rede. Os serviços de comunicação de dados são fornecidos por Provedores de Serviços de Comunicação (CSP – *Communication Service Providers*), comumente uma infraestrutura de comunicação de propósito geral para transporte de dados administrada por uma terceira parte não envolvida com a aplicação.

Espera-se que estes novos paradigmas de transportes tragam benefícios para a sociedade, com a redução de acidentes e fatalidades, aumento da eficiência e da capacidade dos sistemas de transporte. Contudo, observa-se que estes paradigmas apresentam uma elevada complexidade sistêmica, especialmente nos cenários futuros de transporte onde frotas de veículos operam de forma autônoma[4] e cooperativa[5]. Consequentemente, a capacidade de prever e lidar com condições inseguras será prejudicada, inviabilizando a garantia da segurança crítica. Sobretudo, devido às normas de segurança não serem capazes de garantir a segurança

[4] A tomada de decisão na condução de veículos em sistemas de tráfego é sempre suportada por mecanismos baseados em inteligência, seja humana (veículos tradicionais) ou de máquina (veículos autônomos). Assim, pode-se considerar que ambos os tipos de veículos são agentes autônomos no sistema de transporte. Contudo, as novas tecnologias – não comprovadas pelo uso - e a AI – normalmente, de comportamento não determinístico – limitam a compreensão do comportamento dos veículos autônomos e, consequentemente, a capacidade de prever e identificar condições inseguras no sistema. Assim, em relação aos veículos conduzidos por humanos, veículos autônomos possuem maiores incertezas quanto às suas decisões tomadas, tornando-os escopo de pesquisa e desenvolvimento.

[5] A capacidade de cooperação entre elementos autônomos amplia a interação e, consequentemente, a possibilidade de surgirem comportamentos emergentes inesperados. Sobretudo quando o horizonte de consciência situacional é ampliado pelas ICT.

do sistema, pois suas abordagens não permitem eliminar do projeto as condições inseguras não previstas (ERICSON II, 2011).

Nas abordagens de SSE, condições inseguras são definidas como condições anormais ocasionadas por situações nas quais os elementos do sistema – individualmente ou em conjunto – expõem os envolvidos ou a si mesmos a perigos. Assim, estas situações inseguras devem ser gerenciadas de forma a manter os riscos de segurança do sistema em níveis aceitáveis. Neste processo, as situações inseguras devem ser sistematicamente identificadas, obtendo-se todas as situações previstas e plausíveis que possam levar o sistema a uma condição insegura. Por fim, o sistema é especificado, projetado e implementado para que as situações inseguras previstas sejam eliminadas ou minimizadas a níveis aceitáveis de risco de segurança.

Vale frisar que um sistema é seguro para a aplicação e para o ambiente considerados no projeto. Consequentemente, as especificações de um sistema são válidas – para fins de garantia de segurança – apenas enquanto as premissas adotadas no projeto não forem alteradas, sobretudo aplicação e ambiente operacional. Portanto, para um sistema justificado como seguro, espera-se que não ocorram condições inseguras enquanto o sistema estiver operando em atendimento às suas especificações.

Porém, caso as especificações do sistema sejam deficientes, sobretudo incompletas, um sistema pode ser exposto a situações inseguras mesmo enquanto opera aderente às suas especificações[6]. Isso pode ocorrer quando a situação específica que levou o sistema à condição insegura durante a operação:

i. Não foi prevista durante o projeto, mesmo que as premissas adotadas no projeto estejam válidas no momento que ocorreu o evento inseguro;
ii. Não era previsível durante o projeto, pois as premissas adotadas sofreram modificações ao longo da operação e, consequentemente, a situação que levou à condição insegura não era plausível de ocorrer.

Portanto, em ambos os casos, a especificação está incompleta devido às abordagens de SSE utilizadas durante o projeto do sistema não serem capazes de prever as situações que levam a condições inseguras durante a operação. Consequentemente, esta limitação pode ser contornada

[6] Portanto, falhas inseguras são situações nas quais os elementos do sistema expõem os envolvidos ou a si mesmos a condições inseguras devido ao não atendimento das suas especificações.

ao se gerenciar a especificação do sistema (seus recursos e condições operacionais) durante sua operação. Assim, baseado nas relações entre as características dos sistemas (sobretudo para os novos paradigmas) e os riscos de segurança crítica de suas aplicações, as especificações de segurança – representadas pela configuração do sistema que a implementam – seriam readequadas sempre que uma condição insegura (uma exposição a um perigo) fosse observada e o sistema estivesse operando aderente às especificações vigentes (que, até aquele momento, eram consideradas seguras).

Com esta abordagem, seria possível ampliar a capacidade do sistema em prever e lidar com condições inseguras, independentemente de suas causas, mantendo os níveis de risco aceitáveis durante sua operação. Consequentemente, esta abordagem contribuiria com o processo de garantia de segurança crítica no contexto de sistemas de transporte cooperativos e, de forma mais ampla, de Sistemas Complexos de Engenharia (CoES – *Complex Engineered Systems*).

As abordagens de garantia de segurança de sistemas críticos tendem a estar vinculadas ao domínio de aplicação daquele sistema. Portanto, esta pesquisa busca delimitar seu escopo, quando necessário à discussão, aos transportes rodoviários no contexto dos Sistemas de Transporte Inteligentes Cooperativos (C-ITS – *Cooperative Intelligent Transport Systems*). Contudo, os achados e conclusões obtidas neste escopo são extrapolados, quando possível, para outros contextos de CoES críticos em segurança. Portanto, onde for aplicável, esta generalização é realizada.

1.1. Objetivos

Características de complexidade sistêmica introduzidas, sobretudo, pelas novas Tecnologias de Comunicação e Informação (ICT) estão alterando o paradigma de sistemas críticos e limitando a aplicação das abordagens atuais de Engenharia de Segurança de Sistemas (SSE – *System Safety Engineering*). Desta forma, novas abordagens em SSE são necessárias para garantia de segurança destes sistemas complexos artificiais produzidos pelo homem, denominados como Sistemas Complexos de Engenharia (CoES).

Considerando as limitações atuais da SSE em prever e lidar com condições inseguras em CoES, o objetivo geral desta pesquisa é obter uma abordagem – a ser incorporada e utilizada durante a operação do CoES – que permita lidar com situações não previstas ou/e previsíveis

durante o projeto e que possam levar a condições inseguras durante a operação do sistema, possibilitando a manutenção de níveis aceitáveis de risco de segurança durante a operação e contribuindo no processo de garantia de segurança do sistema de CoES.

Em decorrência deste objetivo geral, esta pesquisa possui 3 objetivos específicos:

i. Identificar o conjunto de características intrínsecas aos CoES relacionadas ao novo paradigma de sistemas de transporte cooperativos e compreender a forma como estas características influenciam os – e são influenciados pelos – riscos de segurança crítica de suas aplicações. Compreender estas características e relações habilita à obtenção da abordagem de garantia de segurança e pode orientar novas práticas e processos mais apropriados para lidar, no ciclo de vida de engenharia, com CoES.

ii. Obter uma abordagem conceitual que gerencie a configuração do sistema em função de condições inseguras (exposição a perigo) observadas durante sua operação e que, ao ser incorporada aos sistemas de transporte rodoviário no contexto C-ITS, possibilite a manutenção de níveis aceitáveis de risco de segurança em uma aplicação de tráfego cooperativo.

iii. Instanciar e avaliar se uma abordagem conceitual baseada na relação entre características funcionais e riscos de segurança de um CoES e aplicada durante a operação do sistema pode ser utilizada para otimizar recursos do sistema em momentos onde seja observado a menor demanda da aplicação quanto à exposição aos riscos.

Por fim, esta pesquisa visa contribuir com inovação[7] no campo da Engenharia de Segurança de Sistemas (SSE), produzindo meios que promovam a evolução dos processos de garantia de segurança de sistemas críticos e complexos de engenharia.

[7] **Inovação**: "*Introdução de novidade ou aperfeiçoamento no ambiente produtivo e social que resulte em novos produtos, serviços ou processos ou que compreenda a agregação de novas funcionalidades ou características a produto, serviço ou processo já existente que possa resultar em melhorias e em efetivo ganho de qualidade ou desempenho*". [Lei nº 13.243, de 11 de janeiro de 2016]

1.2. Justificativa

A questão da complexidade em engenharia não é um tema recente. Por exemplo, o Capítulo 1 do livro[8] "*System Engineering: An Introduction to the Design of Large-scale Systems*" (GOODE; MACHOL, 1957) é intitulado "*Complexity – The Problem*". Nesse capítulo, o autor caracteriza e apresenta os Sistemas de Larga Escala (*Large-scale systems*)[9] como complexos, pois: mudanças em uma variável afeta muitas outras, raramente de forma linear; o efeito no sistema não é sempre bem compreendido; existem múltiplos laços de realimentação, incluindo laços de realimentação dentro de outros laços (sobretudo de informação); as entradas são descritas estatisticamente; os aspectos competitivos afetam o sistema; entre outros.

Além disso, GOODE (1957) aponta que a especialização crescente havia trazido problemas para a capacidade de intercomunicação, e que faltava coordenação entre as partes para lidar com os problemas. Portanto, para lidar com a complexidade, o autor apontou a necessidade de uma abordagem sistêmica / generalista / holística, a qual demandava por '*engineering scientists*' (um sinônimo para engenheiros de sistemas). Desta forma, surge de forma explícita o conceito da Engenharia de Sistemas, que vem evoluindo ao longo dos anos.

Atualmente, abordagens da Engenharia de Sistemas Tradicional (TSE – *Traditional Systems Engineering*) – e, analogamente, da Engenharia de Segurança de Sistemas (SSE)[10] – adotam um paradigma estruturado e reducionistas, sobretudo baseadas em decomposição/integração funcional hierárquica, inclusive na verificação, validação e comissionamento destes sistemas. Este paradigma permite garantir apenas as propriedades concebidas e especificadas no início do seu ciclo de vida. Desta forma, a partir do início da operação, tanto o sistema quanto o seu ambiente devem permanecer com suas características inalteradas ao longo de sua vida útil. Caso algo seja alterado, não se pode mais garantir as propriedades validadas no comissionamento do sistema, incluindo a segurança crítica.

---

[8] Considerado como o primeiro livro publicado no tema 'engenharia de sistemas'.

[9] Para o GOODE (1957), "*large-scale systems are characterized by a large number of parts, in number of different types of parts, in number of functions performed, inputs and in absolute cost ... boundaries are embedded in large-gray areas ...*".

[10] A Engenharia de Segurança de Sistemas (SSE) está intimamente relacionada à engenharia de sistemas, tendo os mesmos princípios e abordagens e as mesmas condições de contorno para sua aplicação. Portanto, as mesmas limitações (gargalos) identificadas à aplicação das abordagens da engenharia de sistemas tradicional em sistemas complexos podem ser observadas para a aplicação das abordagens da SSE tradicional em sistemas complexos críticos em segurança.

Conforme discutido anteriormente, estas premissas não são aplicáveis aos novos paradigmas de sistemas complexos de engenharia. Alguns autores buscaram identificar quais deveriam ser as diferenças entre as abordagens da Engenharia de Sistemas Tradicional (TSE) e da Engenharia de Sistemas Complexos (CoSE – *Complex Systems Engineering*).

NORMAN (2004) identificou como requisitos à utilização da TSE: sistemas com fronteiras e condições estáveis e bem definidas (de sistema e de ambiente); que o sistema interage com o ambiente, mas não o altera (e vice-versa); e, sobretudo, que é possível aplicar o conceito de reducionismo no desenvolvimento e análise dos sistemas tradicionais. Assim, existem condições de contorno para a aplicação da TSE, onde: o resultado desejado deve ser conhecido a priori e deve ser claro e inequívoco; a gestão do projeto do sistema deve ser realizada por um único gerente, e as mudanças devem ser introduzidas e gerenciadas de forma centralizada; e deve haver recursos intercambiáveis que possam ser aplicados e realocados conforme a necessidade.

Por outro lado, em sistemas complexos de engenharia (CoES), NORMAN (2004) identificou a inexistência de uma divisão de bases conceituais comuns, e seus elementos não são construídos para o propósito da aplicação-fim. Os sistemas dividem um ambiente de aquisição que os colocam de forma autônoma, bem como não têm gerenciamento e controle comuns e não possuem fundos comum para solução de problemas (competição por recursos). Os sistemas possuem múltiplos clientes e, sobretudo, seus elementos evoluem – geralmente, de forma descoordenadas – em diferentes taxas em função de diferentes pressões e necessidades impostas pelo ambiente.

Com isso, o autor concluiu que a CoSE muda o foco de "… *here is the solution designed from the requirements, now go to implement it* …" para "… *here are the selective pressures acting on the elements present (likely built using TSE), now resolve or reduce them* …". Assim, o mecanismo primário de mudança em sistemas de engenharia muda de "requisitos definindo efeitos" – abordagem "Top-Down" ou "Structured Design" da TSE – para a "evolução", onde o processo de engenharia deve guiar a evolução (abordagem "Bottom-Up") por meio de pressões seletivas locais (recompensa/punição) para se obter soluções globais nos CoES.

SOMMERVILLE *et al* (2012) apresenta uma relação entre as abordagens de engenharia e o tipo de complexidade (epistêmica ou inerente) observado nos sistemas. Discute que as

abordagens reducionistas de dividir-e-conquistar, base da TSE, são relativamente eficientes para gerenciar a complexidade epistêmica, pois se desenvolvem as partes e suas interfaces e não se considera suas interações e dinamismos. No caso da complexidade inerente, as abordagens atuais da TSE levam a atrasos nas entregas, elevação de custos e ao não atendimento das necessidades.

De acordo com LEVESON (2011b), a segurança crítica é uma propriedade emergente (de sistema), e não uma propriedade de componente. Assim, um dos desafios da CoSE é identificar e analisar os possíveis comportamentos emergentes inseguros que um sistema complexo pode vir a manifestar ao longo de seu ciclo de vida. Contudo, isso implica também em considerar a evolução do sistema complexo. Somada à característica de fronteiras difusas, tem-se uma limitação das abordagens atuais de análise de segurança, que demandam uma definição e descrição claras do sistema, incluindo fronteiras e comportamentos.

BAR-YAM (1997) aponta a modelagem como uma abordagem de análise de sistemas complexos. Neste caso, deve-se modelar cada parte e suas interações e, depois, obter o comportamento do todo emergindo do sistema. Modelagem e simulação computacional têm potencial para aplicação em engenharia de sistemas complexos, pois é uma técnica 'Bottom-up' indicada para identificar comportamentos esperado em sistemas de engenharia clássicos, observáveis por simulação um-para-um (uma mesma entrada gera sempre a mesma saída). Além disso, a teoria de Hayek dos fenômenos complexos (WIBLE, 2000) reforça o uso de simulação para prever padrões ('*pattern prediction*') em fenômenos complexos, dado que somente é possível prever valores precisos em fenômenos não complexos.

De forma pragmática, pode-se inferir que grande parte dos sistemas complexos de engenharia serão 'projetados' – ao menos, no seu instante inicial – e suas propriedades serão validadas para as características projetadas naquele momento. Mas, evoluirão ao longo de seu ciclo de vida, dado que possuem características complexas. Desta forma, em Sistemas Complexos de Engenharia e críticos em segurança, deve-se garantir a segurança crítica quando submetidos a mudanças (evolução), conceito definido por LAPRIE (2008) como '*Resiliência*'. Assim como a dependabilidade (NATELLA; COTRONEO; MADEIRA, 2016) é garantida por meio do gerenciamento de falhas, a resiliência de sistemas complexos será garantida por meio do gerenciamento de mudanças.

LEVESON *et al.* (2006) define resiliência como "*the ability of systems to prevent or adapt to changing conditions in order to maintain (control over) a system property.*" Assume-se como resiliente um sistema que pode se adaptar a mudanças internas (sistema) e externas (ambiente), modificando as ações de seus elementos internos durante sua operação continuada. Portanto, os componentes internos do sistema podem se reconfigurar e adaptar a novos ambientes. Desta forma, ser resiliente é mais do que ser robusto. Ser robusto implica no sistema continuar a atender aos requisitos mesmo na presença de falhas internas (o que é tratado pela dependabilidade). Ser resiliente demanda ser robusto e continuar a atender seus requisitos na presença de ameaças externas (do ambiente).

Portanto, com base na discussão apresentada, observa-se que a garantia da segurança crítica em CoES poderia ser obtida por meio do gerenciamento de mudanças internas (sistema) e externas (ambiente) ao longo da operação continuada destes sistemas. Este gerenciamento deve ser capaz de modificar, em tempo de execução, as ações (comportamentos) de seus elementos internos como forma de manter a aplicação em níveis de risco de segurança aceitáveis. Ou seja, uma abordagem de gerenciamento de recursos do sistema baseado nas relações entre características intrínsecas do sistema (complexo) e o nível de risco da aplicação tem potencial de aplicação nos processos de garantia de segurança crítica dos Sistemas Complexos de Engenharia.

1.3. Organização do trabalho

Este trabalho está estruturado com base em 7 seções, um Apêndice e um Anexo.

A seção 2 apresenta os conceitos fundamentais necessários ao desenvolvimento deste trabalho de investigação e cumprimento dos objetivos de pesquisa. Desta forma, são apresentados conceitos relacionados aos Sistemas Críticos em Segurança (*safety*), aos Sistemas Complexos de Engenharia (CoES) e aos Sistemas Inteligentes de Transporte Cooperativo (C-ITS), escopo de aplicação deste trabalho.

A seção 3 apresenta a abordagem conceitual proposta pelo autor para a garantia de segurança de CoES baseada no gerenciamento de recursos de comunicação e orientado aos riscos observados na aplicação crítica.

A seção 4 apresenta, em detalhes, o estudo de caso elaborado para apoiar a concepção,

desenvolvimento, validação e demonstração da abordagem proposta. Tanto os modelos conceituais do cenário-base de tráfego e da abordagem proposta quanto suas implementações computacionais são apresentadas nesta seção.

A seção 5 apresenta, analisa e discute os resultados das campanhas de simulação dos modelos implementados na seção anterior.

A seção 6 apresenta as conclusões e contribuições obtidas neste trabalho de investigação científica, bem como propõe linhas de investigação futuras para continuação deste trabalho.

O Apêndice I detalha a proposta de um método desenvolvido para auxiliar na estimativa eficiente do risco de colisão entre objetos móveis em movimento guiado, especialmente no contexto de Sistemas de Transporte Inteligentes. Este método é utilizado na implementação do estudo de caso, apresentado na seção 4.

O Anexo I apresenta um estudo realizado para identificar quais técnicas de aprendizado supervisionado em inteligência artificial produziriam o modelo executivo mais eficiente a ser utilizado no contexto do estudo de caso. Os resultados obtidos são utilizados na implementação do estudo de caso (seção 4).

2. CONCEITOS E FUNDAMENTOS

"*A man will turn half a library to make one book*" [11]
Samuel Johnson (1709-1784)

Esta seção apresenta e discute os conceitos fundamentais ao desenvolvimento deste trabalho de investigação, como aqueles relacionados aos Sistemas Críticos em Segurança (safety), aos Sistemas Complexos de Engenharia (CoES) e aos Sistemas Inteligentes de Transporte Cooperativo (C-ITS), escopo de aplicação deste trabalho.

2.1. Sistemas críticos em segurança

A apresentação e discussão a respeito de sistemas críticos em segurança inicia pela definição de 'sistema'. Este termo é intensivamente utilizado tanto nesta tese quanto nos mais diversos aspectos da vida cotidiana. Como forma de reduzir a possibilidade de múltiplas interpretações e direcionar o desenvolvimento desta investigação, faz-se necessário formalizar os conceitos relacionados aos sistemas e aderentes ao escopo de aplicação deste trabalho, orientado à engenharia de sistemas complexos críticos em segurança. Portanto, os conceitos e definições adotados neste trabalho estão relacionados aos **sistemas de engenharia** – sistemas projetados ou adaptados para **interagir** com um **ambiente (contexto) operacional previsto** e atender a **propósitos pretendidos** (**missões**) pelos seus **envolvidos (**usuários, desenvolvedores, proprietários, operadores, entre outros**)** enquanto cumpre com as restrições aplicáveis (INCOSE, 2023).

Um **sistema** é formado por um conjunto de **elementos** – físicos e/ou conceituais, naturais e/ou artificiais, como pessoas, produtos, serviços, informação e/ou processos – que **interagem** entre si e com o ambiente operacional. Os resultados produzidos pelo sistema (comportamentos ou significados pretendidos) não seriam obtidos pelos elementos isoladamente, e são percebidos pelo ambiente operacional na **fronteira do sistema**. Um sistema pode ser considerado tanto como serviços quanto como produtos que ele provê, e se relaciona com outros sistemas do ambiente operacional por meio de **cooperação** ou/e **competição**.

[11] James Boswell (1791). *The Life of Samuel Johnson, LL.D*. Disponível em https://www.gutenberg.org/files/1564/1564-h/1564-h.htm.

A Figura 1 ilustra esta relação Elemento-Sistema-Ambiente, bem como os demais conceitos mencionados.

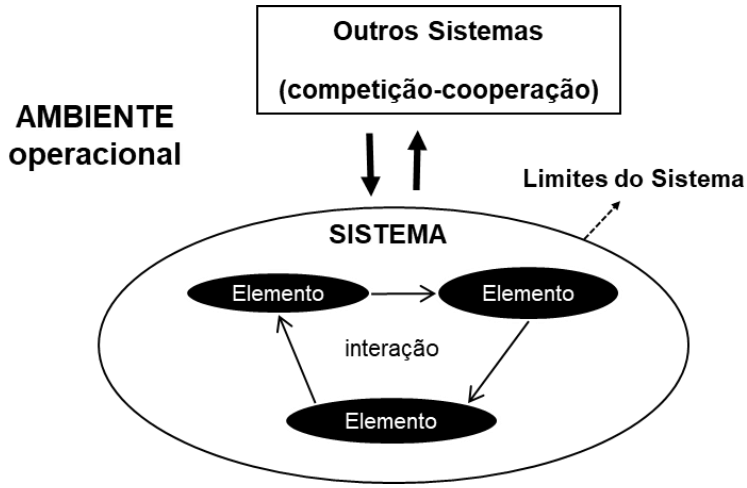


Figura 1 – Interação Elementos-Sistema-Ambiente

A **arquitetura** de um sistema representa sua **estrutura de elementos**, o modo como se relacionam entre si e com o ambiente operacional, bem como os princípios e regras que governam o projeto e evolução do sistema (IEEE, 2000). Em outras palavras, uma arquitetura descreve um sistema. A arquitetura de um sistema pode ser descrita por meio de diversas ***visões***, que representam um sistema (funcional, comportamental, estrutural, entre outras) sob a perspectiva de um conjunto de interesses (considerações importantes sobre o sistema, como desempenho, confiabilidade, tamanho, custo, entre outras) de seus interessados. A Figura 2 ilustra o diagrama de entidade entidade-relacionamento entre os conceitos apresentados e relacionados ao conceito de sistema.

A **fronteira do sistema** delimita quais elementos fazem parte do **sistema de interesse** (*System of Interest* – SoI), a qual define os elementos que são necessários o sistema existir de maneira autossuficiente dentro de seu ambiente operacional. Ou seja, a fronteira é definida com base na missão (propósitos pretendidos) para aquele sistema dentro de seu ambiente operacional previsto. Como exemplo, a Figura 3 ilustra o sistema de interesse S_1. Para atender aos propósitos previstos por seus interessados, S_1 é formado por uma **estrutura** de três **elementos** (S_{11}, S_{12} e S_{13}), que interagem entre si e se relacionam com os sistemas S_2, S_3 e S_4 no seu ambiente operacional previsto.

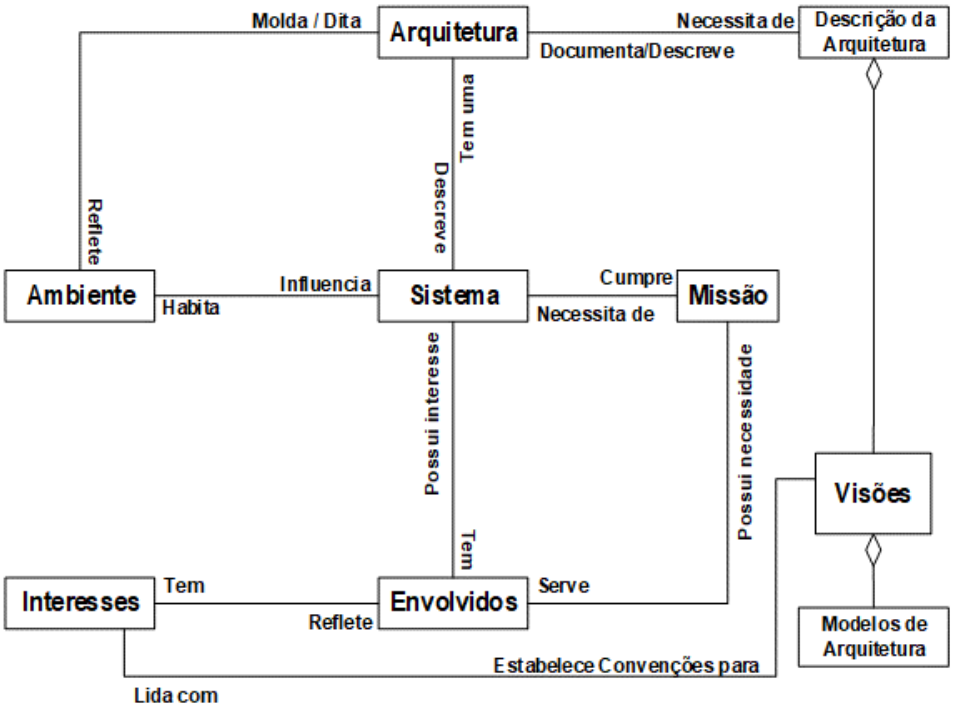


Figura 2 – Diagrama de entidade-relacionamento entre os conceitos relacionados a sistema
Fonte: adaptado de IEEE (2000)

Um elemento é uma parte discreta do sistema que pode ser desenvolvida e implementada (ou incorporada) para cumprir com requisitos especificados (INCOSE, 2023). Ou seja, um elemento cumpre propósitos previstos dentro de um contexto operacional que, neste caso, são os elementos internos à fronteira do sistema S_1. Portanto, um **elemento** do sistema S_1 pode ser observado como um **sistema de interesse** (SoI) em um nível mais detalhado de **abstração** de um sistema. Por exemplo, S_{12} é considerado como um elemento de S_1 dentro do nível 1 de abstração deste sistema. Contudo, em um nível mais detalhado de abstração (nível 2), pode ser abstraído como um SoI e **decomposto** como um **subsistema** de S_1. Este subsistema, agora um SoI, é formado por uma estrutura de cinco elementos que interagem no contexto delimitado pela fronteira de S_{12}. Essa decomposição dos elementos de um sistema é realizada até a menor parte necessária para desenvolvimento do sistema (nível de **componente)**. No exemplo, nível 3 da decomposição para um dos elementos do subsistema S_{12}.

A premissa de que o **reducionismo** é aplicável aos sistemas de engenharia é o que possibilita decompor um sistema hierarquicamente em partes, suas partes serem tradadas independentemente umas das outras e, ao final, estas partes podem ser (re)integradas para representar o sistema. Esta premissa um dos princípios centrais da Engenharia de Sistemas

(*Systems Engineering* – SE)[12], orientando suas ações ao longo do ciclo de vida[13] de sistemas de engenharia.

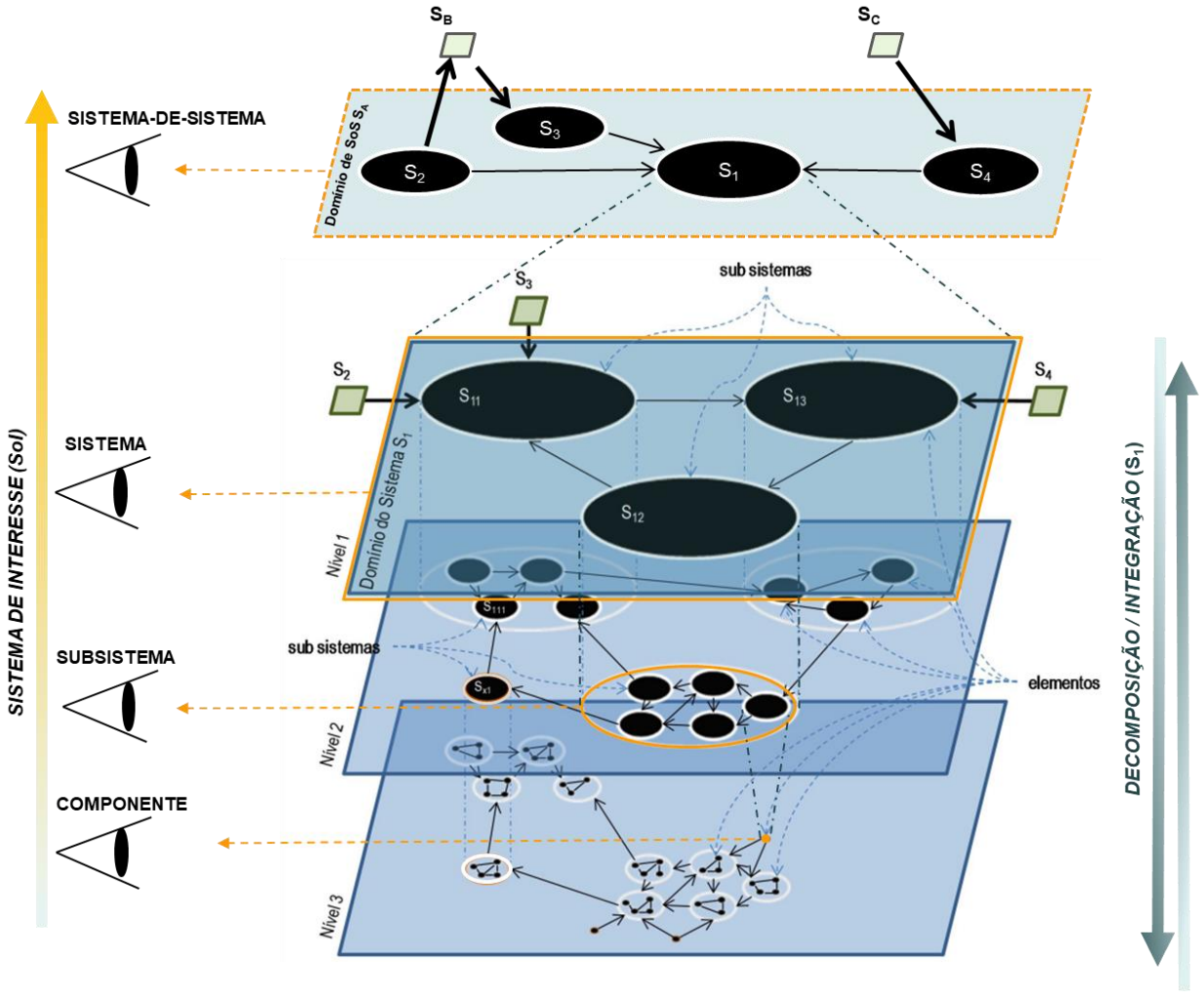


Figura 3 – Sistema de interesse, estrutura e nível de abstração

Fonte: adaptado de LEMES (2011)

Pode-se dividir o ciclo de vida completo de sistemas de engenharia em 5 fases principais: **Concepção** (identificação das necessidades dos envolvidos, exploração de conceitos e proposição de soluções viáveis), **Desenvolvimento** (especificação, projeto, implementação, verificação e validação do sistema conforme sua especificação), **Operação/Manutenção** (sistema é utilizado no ambiente operacional previsto e mantido nas condições funcionais

[12] A Engenharia de Sistemas é definida como uma abordagem transdisciplinar e **integrativa** para permitir a concepção, uso e desativação bem-sucedida de sistemas de engenharia, usando princípios e conceitos de sistemas e métodos científicos, tecnológicos e de gerenciamento (INCOSE, 2023).

[13] Ciclo de vida do sistema é o período que inicia quando o sistema é concebido e termina quando o sistema não está mais disponível para uso (ISO/IEC/IEEE, 2019)

especificadas, cumprindo a missão concebida), **Modificação/Atualização** (quando necessário lidar com questões não previstas durante o desenvolvimento) e **Desativação/Substituição** (sistema deixa de ser utilizado).

Figura 4 ilustra um modelo de ciclo de vida em 'V', tradicional na engenharia de sistemas, desde sua fase de concepção até sua desativação. Este modelo engloba uma abordagem '*Top-Down*' de concepção, planejamento e desenvolvimento, onde cada fase decompõe e refina a fase anterior, e uma abordagem '*Bottom-up*' de Integração, Verificação[14] e Validação[15], resultando em um sistema validado para sua fase de operação/manutenção. Vale frisar que um modelo de ciclo de vida é definido e aplicado ao sistema de interesse (SoI) em consideração.

A engenharia de sistemas é aplicável a sistemas de engenharia, projetados para atender necessidades específicas, prestar serviços e executar tarefas necessárias à vida cotidiana. Quando um sistema não é capaz de cumprir os requisitos desejados e projetados, a missão para a qual foi concebido não pode ser devidamente cumprida. O funcionamento anormal de muitos sistemas do dia-a-dia tem consequências menores, não trazendo perdas significativas às pessoas envolvidas ou o ambiente, causando, no máximo, aborrecimentos ou decepção.

No entanto, sistemas cujas disfunções podem levar a consequências inaceitáveis – como mortes, prejuízos financeiros, materiais e ambientais, a perda de uma missão ou oportunidade – são definidos como '**Sistemas Críticos**' (SOMMERVILLE, 2011; SOMMERVILLE *et al.*, 2012). São definidos três tipos de sistemas críticos, em função das consequências (perdas) que podem produzir em caso de disfunção[16]: (i.) Sistemas críticos quanto aos negócios (*business critical systems*); (ii.) Sistemas críticos quanto à missão (*mission critical systems*); e iii. Sistemas críticos quanto à segurança (*safety critical systems*). No primeiro tipo as consequências catastróficas são financeiras (como em sistemas de e-commerce ou de reservas de passagens). No segundo, estão relacionadas à perda catastrófica de missão (como um sistema de gerenciamento de bagagens em aeroportos ou de lançamento de satélites).

[14] Verificação: atividade que avalia se os requisitos de uma determinada fase do clico de vida foram atendidos de forma apropriada (i.e., "*Are you building it right?*").
[15] Validação: atividade que avalia se os requisitos particulares para um uso específico pretendido foram atendidos de forma apropriada (i.e., "*Are you building the right thing*?")
[16] Neste trabalho, o termo 'disfunção' é a tradução para o termo 'failure', definido na seção 2.1 como o "*fim da capacidade do sistema de entregar ao seu usuário uma função, uma operação ou um serviço de acordo com sua respectiva especificação*" (AVIZIENIS *et al.*, 2004).

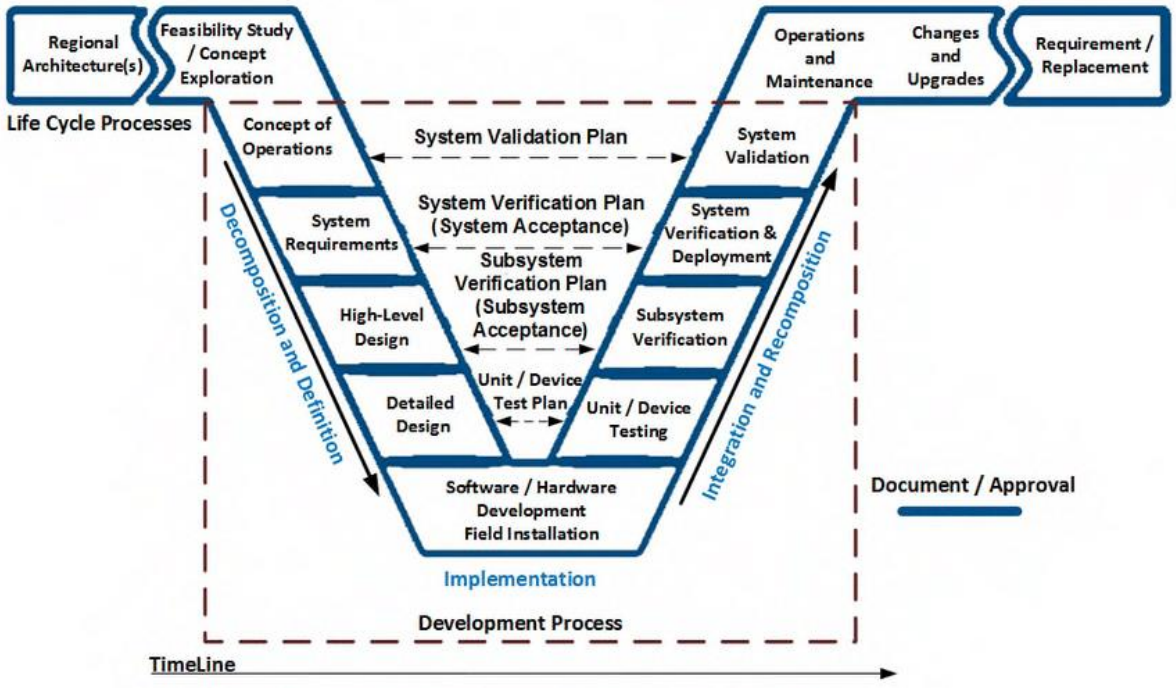


Figura 4 – Exemplo um ciclo de vida em 'V'
Fonte: (MAINDZE; SKAF; JENNIONS, 2019)

O terceiro tipo – os **sistemas críticos quanto à segurança** (*safety-critical systems*) corresponde aos sistemas cujo funcionamento anormal pode causar **perdas e danos à vida humana e ao ambiente** no qual estão inseridos. Sistemas de supervisão e controle de reatores em usinas nucleares são bons exemplos de sistemas críticos em segurança, pois seu funcionamento anormal pode levar ao vazamento de elementos radioativos, causando mortes, ferimentos e danos ambientais.

A definição de alguns dos principais aspectos relacionados à sistemas críticos em segurança estão apresentados na Tabela 1. Ao comparar e relacionar estas definições, cujos conceitos, sobretudo de risco, são gerais e vão além do domínio de sistemas críticos em segurança, é possível redefinir o conceito de **sistema crítico** como "*sistemas nos quais o desvio em relação ao comportamento esperado* (desejado) *pode levar a efeitos negativos nos objetivos*". Além disso, o conceito de **sistemas críticos em segurança** pode ser redefinido como: "*sistemas nos quais o desvio em relação ao comportamento esperado* (desejado) *pode levar a lesões e danos para a saúde das pessoas, ou dano à propriedade e ao ambiente*". Desta forma, estas redefinições permite uma abordagem mais ampla do conceito de disfunção e de

consequências, incluindo os casos onde as fronteiras do sistema são difusas ou os modos de disfunção (*failure modes*)[17] não são muito bem definidos.

Tabela 1 – Definições de conceitos diretamente relacionados a sistemas críticos

CONCEITO	DEFINIÇÃO NORMATIVA (DOD, 2012; ISO/IEC, 2014)
Segurança Crítica (*Safety*)	[1]. Livre de **riscos** que não são toleráveis; [2]. Livre de condições que podem causar morte, ferimentos, doenças ocupacionais, dano ou perda de equipamento propriedade, ou dano ambiental.
Risco (*Risk*)	Efeito[18] da incerteza[19] sobre os objetivos (frequentemente caracterizado por referência a eventos potenciais e consequências); Combinação da probabilidade de ocorrência de **dano** e a severidade do **dano**.
Dano (*Harm*)	Ferimento ou prejuízo para a saúde das pessoas ou prejuízos para propriedades ou para o ambiente.
Perigo (*Hazard*)	Potencial fonte de **dano** (perigo pode ser uma **fonte de risco**).
Fonte de risco (*Risk source*)	Elemento que tem um potencial intrínseco para dar origem a risco.

Dado que um sistema crítico é definido pelo efeito (consequência) que provoca no ambiente que opera (e está inserido), a distinção e relação entre a '***aplicação***' e o '***sistema***' crítico propostos por ALMEIDA JR (2003) é útil ao processo de desenvolvimento, avaliação e garantia de propriedades dos sistemas críticos. Um **sistema crítico** supervisiona e controla uma **aplicação crítica**, que pertence a um domínio de aplicação crítica. É na **aplicação crítica** onde as consequências inaceitáveis (eventos catastróficos, com mortes, ferimentos e prejuízos materiais e ambientais) podem ocorrer, pois nela os perigos se manifestam (pois existem condições necessárias para levar a lesões e ferimentos, havendo energia envolvida em seus processos de transformação). Estes efeitos negativos sobre os objetivos (como exemplo, transportar) podem ser causados pelo desvio do comportamento esperado / desejado (*failure*) do **sistema crítico**. A Figura 5 ilustra a relação entre o sistema e a aplicação crítica, conforme proposto por ALMEIDA JR (2003).

[17] Para manter a consistência dos termos adotados neste trabalho, utilizou-se 'modo de disfunção' ao invés de 'modo de falha' (*fault mode*).
[18] Efeito é um desvio do esperado – positivo e/ou negativo.
[19] Objetivos podem ter diferentes aspectos (tais como financeiros, saúde e segurança e metas ambientais) e podem ser aplicadas em diferentes níveis (tais como estratégico, organizacional, projeto, produtos e processos).

Figura 5 – Relação entre sistema e aplicação crítica
Fonte: adaptado de ALMEIDA JR (2003)

Ao confrontar a evolução de alguns **domínios de aplicação crítica** em transporte com a taxonomia proposta por ALMEIDA JR (2003), observa-se que a evolução vem ocorrendo sobre os **sistemas críticos** (onde as disfunções podem levar a consequências inaceitáveis). As **aplicações críticas** permanecem praticamente inalteradas, processando a energia envolvida em seus processos de transformação para o atendimento dos objetivos (nos exemplos de transportes, transportar bens e pessoas entre locais). A Figura 6 ilustra a evolução ocorrida nos sistemas de sinalização e controle em transporte metroviário ao longo das últimas décadas. A aplicação crítica de transporte continua, em essência, inalterada. Já seus sistemas críticos sofreram mudanças significativas, incorporando novas tecnologias de forma evolutiva.

Ao longo das últimas décadas, a evolução dos sistemas críticos vem ocorrendo baseada em Tecnologias de Informação e Comunicação (ICT – *Information and Communication Tecnologies*), sobretudo por meio de sistemas Elétricos, Eletrônicos e Programáveis (E/E/PES – *Electric/Electronic/Programmable Electronic Systems*). O surgimento de novas demandas nos **domínios de aplicação** (ex.: busca por maior capacidade, maior eficiência, menores custos operacionais e de implantação) motiva o desenvolvimento de novas tecnologias e conceitos que possam suprir as novas demandas.

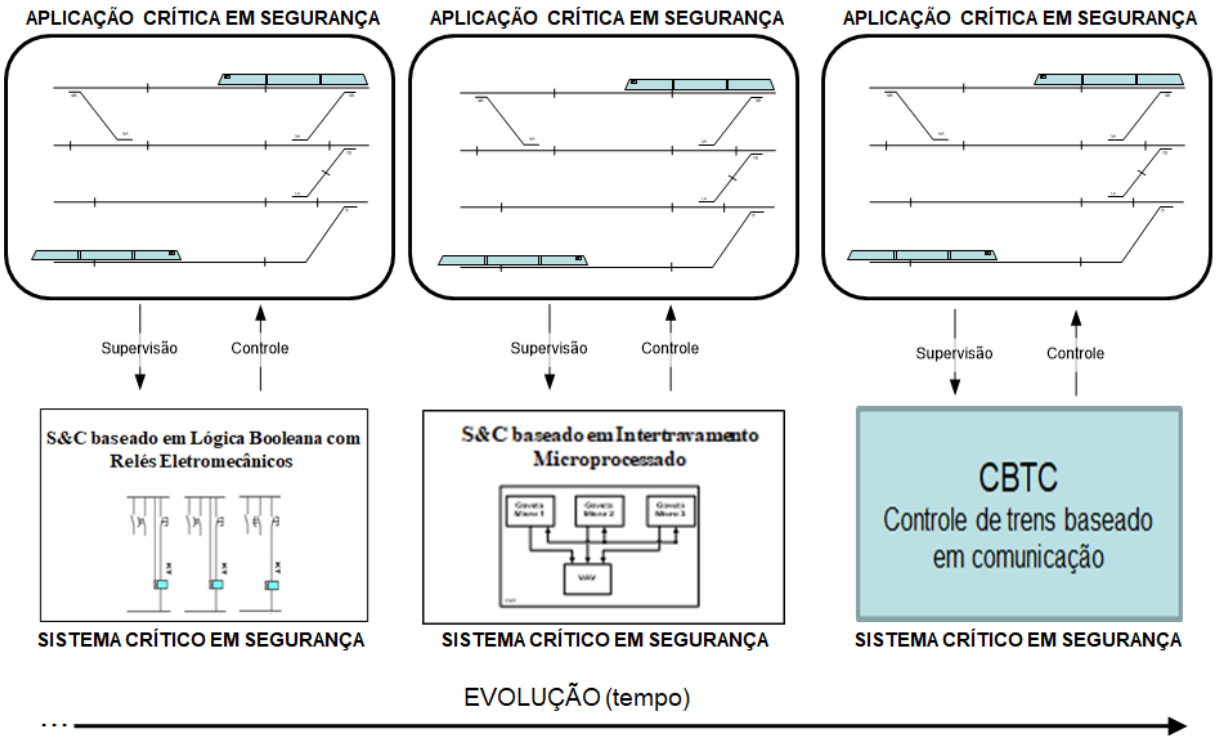


Figura 6 – Evolução dos sistemas de sinalização e controle no domínio de aplicação metroviário

Estas novas tecnologias e conceitos, aplicados nos **domínios de sistema**, trazem consigo novos (e ainda desconhecidos) modos de falha para o sistema. Consequentemente, estes novos modos de falha podem provocar consequências inaceitáveis no domínio da aplicação. Portanto, novas demandas trazem consigo novos desafios para a engenharia quanto à **garantia dos objetivos** dos sistemas e aplicações. A Figura 7 ilustra este ciclo evolutivo dos sistemas e aplicações.

Figura 7 – Ciclo evolutivo dos sistemas e aplicações

Este estudo observou que a evolução dos sistemas críticos, sobretudo quanto à segurança crítica (*safety*), baseada em Tecnologias de Informação e Comunicação (ICT), é representativamente aderente ao paradigma atual de "**Sistemas Ciberfísicos**" (CPS – *Cyber-Physical Systems*). CPS podem ser definidos por:

> *"Systems that offer* ***integrations of COMPUTATION, NETWORKING, and PHYSICAL PROCESSES*** *or, in other words, as the systems where physical and software components are deeply intertwined, each operating on different spatial and temporal scales, exhibiting multiple and distinct behavioral modalities, and interacting with each other in a myriad of ways that change with context"* (KHAITAN; MCCALLEY, 2015)

Conforme discutido anteriormente, domínios de aplicação crítica em segurança possuem um 'lado' físico (processo físico/químico/biológico que envolve transformação/transferência de energia) e um 'lado' 'ciber' (sistemas críticos em segurança baseados em ICT). O mesmo ocorre com o paradigma de Sistemas Ciberfísicos (CPS). A Figura 8 ilustra uma relação entre o conceito de CPS e de sistema e aplicação em domínio crítico em segurança. Pode-se concluir que os domínios de aplicação crítica em segurança são, ou virão a se tornar, sistemas ciberfísicos.

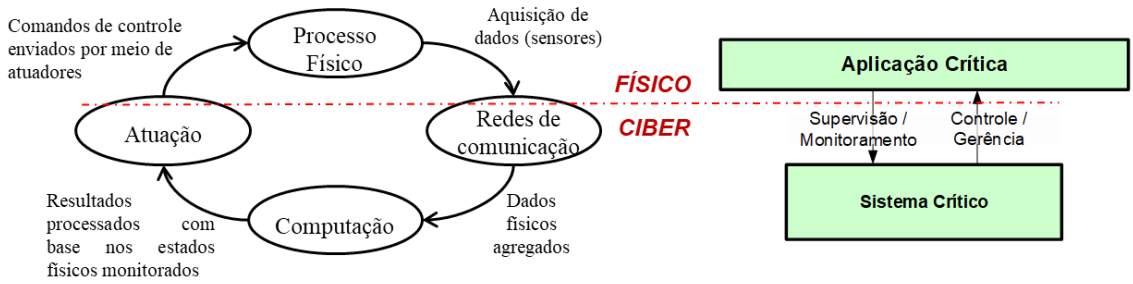


Figura 8 – Relação entre CPS e domínio de aplicação crítica em segurança

Dada às características dos sistemas críticos em segurança e seus possíveis impactos negativos (catastróficos) aos envolvidos quando apresentam comportamento indevido (não desejado/inesperado), duas questões podem ser formuladas quando da concepção destes sistemas (AVIZIENIS *et al.*, 2004):

1. Como desenvolver um sistema que forneça um serviço em que as pessoas possam justificadamente (com base em evidências) confiar?
2. Como evitar que suas disfunções (*failures*) sejam inaceitavelmente frequentes e/ou severas (*i.e.* o risco de eventos catastróficos seja aceitável)?

A Figura 9 ilustra três exemplos de sistemas de engenharia críticos em segurança no domínio dos transportes (uma aeronave, um trem e um carro autônomo), desde a fase de concepção até sua operação, bem como exemplos de eventos catastróficos que ocorrem com os mesmos. O objetivo da engenharia de sistemas críticos é lidar com as duas questões anteriormente formuladas, transformando a concepção (desejo) destes sistemas em realidade (sistema

finalizado e operacional) e, quando operacionais, o risco de perdas catastróficas seja muito baixo.

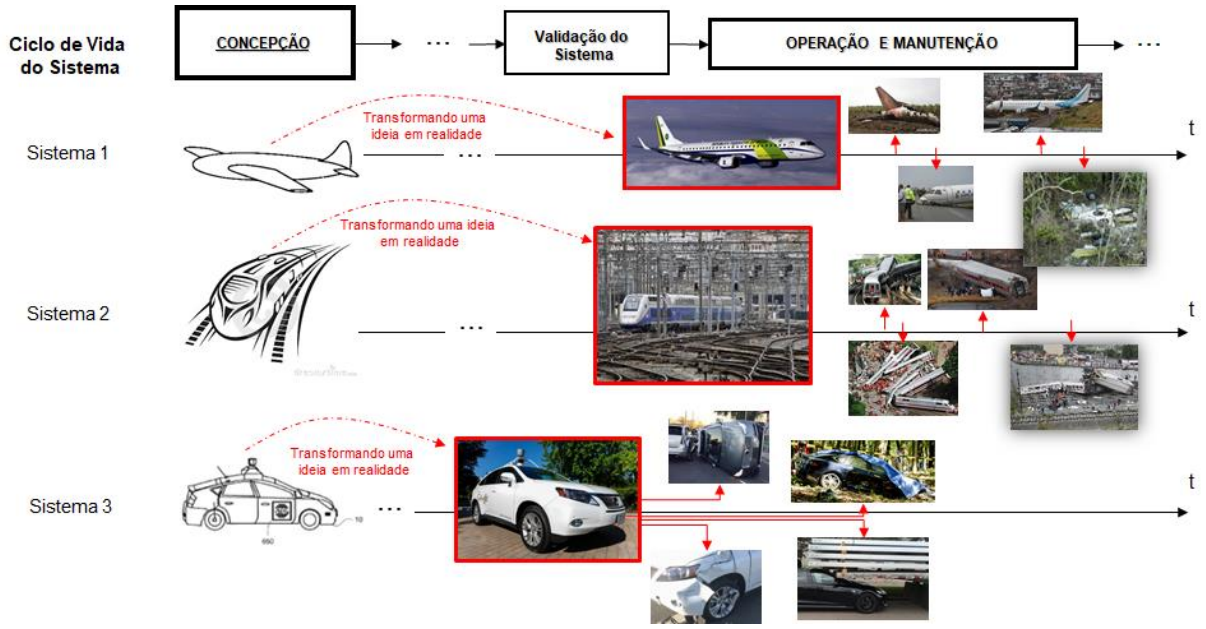


Figura 9 – Exemplos de sistemas críticos, seus ciclos de vida e eventos catastróficos

A disciplina da engenharia que lida com questões de segurança crítica é a Engenharia de Segurança de Sistemas (*System Safety Engineering* – SSE). Por definição, a SSE é definida como uma disciplina da engenharia que emprega conhecimento especializado e habilidades em aplicar princípios científicos e de engenharia, critérios e técnicas para identificar perigos e, então, eliminar os perigos ou reduzir os riscos associados quando um perigo não pode ser eliminado (DOD, 2012).

Na SSE, o conceito de segurança crítica do sistema (***system safety***) é definido como a aplicação de princípios de gerenciamento e de engenharia, critérios e técnicas para atingir riscos aceitáveis de acidente (***safety risks***), dentro das restrições de eficácia e adequação operacional, tempo e custo, em todas as fases do ciclo de vida do sistema (DOD, 2012). Ou seja, a SSE lida com perigos e os riscos associados, e com os cenários que podem levar à exposição aos perigos, por meio de duas abordagens distintas, contudo, complementares, o **Gerenciamento de Risco (*Risk Management*)** e o **Gerenciamento de Falhas (*Fault Management*)**:

- O **Gerenciamento de Risco** é uma abordagem '*Top-Down*' que lida com os dois termos relacionados aos riscos de dano: **probabilidade de ocorrência (exposição)** ao dano e **severidade (consequências)** do dano. Assim, a gerência de risco busca-se tanto

eliminar ou reduzir a exposição aos perigos (e **fontes de risco**) quanto reduzir a severidade (consequência) de um acidente caso ele ocorra.

- O **Gerenciamento de Falhas** é uma abordagem '*Bottom-up*' que lida com as causas que levam à uma disfunção e, consequentemente, a um evento de perda no domínio da aplicação. Assim, busca-se evitar a introdução de falhas no sistema durante o desenvolvimento, remover as falhas introduzidas antes que o sistema inicie sua operação ou tolerar a ocorrência das falhas que ainda permanecerem no sistema quando estiver operando.

A Figura 10 ilustra estas duas abordagens e sua relação com o conceito de Aplicação/Sistema crítico proposta por (ALMEIDA JR, 2003).

Figura 10 – Abordagens de garantia de segurança na SSE

O **Gerenciamento de Riscos** é uma abordagem que se inicia (e se concentra) na **camada de aplicação crítica**. O gerenciamento de riscos identifica e lida com **situações de risco** no domínio de aplicação crítica (ambiente operacional previsto) que podem ser causadas pelo sistema crítico no escopo de operação. Consequentemente, as situações de risco devem ser controladas – eliminadas ou reduzidas – tanto por ações ou/e mecanismos no sistema crítico quanto no ambiente. Quanto à sua abordagem, o gerenciamento pode ser realizado tanto de forma **genérica** (ABNT, 2018; ISO, 2009) sobre o conceito de risco – ou seja, o efeito da incerteza sobre os objetivos, onde os perigos (que podem levar a danos a vida humana) são apenas uma das possíveis fontes de risco – quanto de forma **específica** (ISO/IEC, 2014), centrada na segurança crítica, onde a única fonte de risco são os perigos.

Em ambas as abordagens de gerenciamento de riscos, o processo é composto pelas atividades de **definição de contexto** – onde se identifica e descreve o sistema, o ambiente e suas interfaces, seus usuários, qual o uso pretendido, quais as possíveis falhas e usos indevidos, entre outros; de **identificação de risco (ou perigos)** – onde as fontes de risco (perigos) são identificadas, seus riscos são estimados e avaliados (contra um nível tolerável/aceitável); e de **redução/tratamento de riscos** – onde os riscos são reduzidos até que o risco residual esteja em níveis considerados aceitáveis, tanto pela redução da exposição às fontes de risco (perigos) ou/e pela redução da severidade dos acidentes;

O **Gerenciamento de Falhas** é uma abordagem que inicia (e se concentra) na **camada de sistema crítico**. O gerenciamento de falhas identifica e lida com **situações anormais** (falhas) no sistema crítico, interno às suas fronteiras, que podem levar a situações de risco de danos no domínio de aplicação crítica (ambiente operacional previsto). Consequentemente, as situações anormais devem ser controladas – eliminadas ou reduzidas – por ações ou/e mecanismos no sistema crítico. O gerenciamento de falhas lida com as causas (**ameaças**) que levam a disfunções do sistema, cujos resultados indesejados (comportamentos anormais) percebidos na **fronteira do sistema** pela aplicação podem expor os envolvidos a perigos.

Por ser uma abordagem ‘botton-up’, existe uma relação de ‘muitos-para-um’ entre uma falha e uma situação de risco, onde diversos tipos e combinações de falhas no sistema podem levar à uma situação de risco na aplicação. Portanto, a abordagem do gerenciamento de falhas deve lidar com a **capacidade de o sistema desempenhar suas funções como esperado e quando necessário**, e não apenas lidar com falhas que levem a situações de risco.

Portanto, o gerenciamento de falhas está baseado no conceito de **dependabilidade**[20]. Neste caso, um sistema crítico deve apresentar **atributos** de dependabilidade (ser um ‘*dependable system*’) como forma de garantir um nível de serviço em que as pessoas possam justificadamente confiar, e onde os riscos de eventos catastróficos são aceitáveis. Dependabilidade em sistemas foi primeiramente definido, para sistemas computacionais, na década de 1980 (AVIZIENIS; LAPRIE, 1986). Depois, o conceito foi refinado para sistemas computacionais críticos, onde dependabilidade foi definida como (AVIZIENIS *et al.*, 2004):

[20] “dependabilidade” é um anglicismo para a palavra “dependability”, que significa “*the quality of being trustworthy and reliable.*” No contexto de engenharia, indica a qualidade do serviço fornecido por um sistema e a confiança depositada no serviço fornecido.

- A capacidade para entregar serviço que justificadamente possa ser confiável; ou
- *A capacidade para evitar serviços disfuncionais que sejam mais frequentes e mais severos do que o aceitável.*

Aqueles autores definiram uma taxonomia para o conceito de dependabilidade formado por 3 componentes: **Atributos** de dependabilidade, **Ameaças** à dependabilidade e **Meios** para garantir a dependabilidade. A definição dos **Atributos (*attributes*)** de dependabilidade são listados na Tabela 2 (AVIZIENIS *et al.*, 2004).

Tabela 2 – Atributos de dependabilidade de um sistema (AVIZIENIS *et al.*, 2004)

Atributo	Definição
Disponibilidade (*Availability*)	Prontidão para serviço correto *(readiness for correct service)*
Confiabilidade (*Reliability*)	Continuidade para serviço correto (continuity of correct service)
Segurança Crítica (*Safety*)	Ausência de consequencias catastróficas para os usuários e o meio ambiente. (*absence of catastrophic consequences on the user(s) and the environment*)
Integridade (*Integrity*)	Ausência de alterações impróprias do sistema. (*absence of improper system alterations*)
Manutenibilidade (*Maintainability*)	Capacidade de sofrer modificações e reparos. (*ability to undergo modifications and repairs*)
Confidencialidade *(Confidentiality*)	Ausência de divulgação não autorizada de informação. (*absence of unauthorized disclosure of information*)

As **Ameaças (*Threats*)** à dependabilidade de um sistema são as **Falhas** (*Faults*), **Erros** (Errors) e **Disfunções** (Failures) que o sistema possui e manifesta. Elas agem em uma cadeia de causa-consequência, onde a disfunção (*failure*) é o efeito percebido na fronteira do sistema (considerado como 'interface de serviço'), conforme ilustrado na Figura 11. Portanto, a definição/caracterização de uma disfunção depende do **nível de abstração** com a qual se define o sistema (e suas fronteiras), conforme discutido anteriormente.

As definições literais de Falha (*Fault*), Erro (*Error*) e Disfunção (*Failure*) apresentadas por AVIZIENIS et al (2004) são:

> **FALHA:** condição anormal que pode levar à redução ou à perda da capacidade de um componente de desempenhar sua função (ou seja, 'errar'). Uma falha é causada (ativada) por alguma vulnerabilidade do sistema (tanto no hardware quanto no software) e pode ser causada por problemas (decisões ou e/ações incorretas) tanto na especificação quanto na implementação, em componentes defeituosos, etc.; tornando assim o sistema vulnerável a perturbações.

ERRO: manifestação de falha. É um desvio entre a especificação do sistema (ou seja, o comportamento desejado) e seu desempenho real. Um erro pode levar o sistema a uma disfunção.

DISFUNÇÃO: fim da capacidade do sistema de entregar ao seu usuário uma função, uma operação ou um serviço de acordo com sua respectiva especificação (características desejadas). É a manifestação (consequência) de um Erro na interface de serviço do sistema.

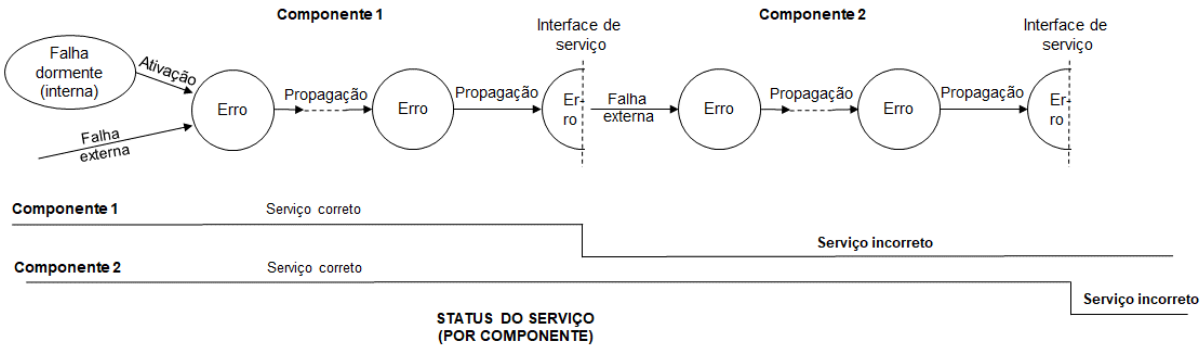


Figura 11 – Relação causal entre falha-erro-disfunção (*fault-error-failure*)
Fonte: (AVIZIENIS *et al.*, 2004)

Por fim, os **Meios** (*Means*) para garantir os atributos de dependabilidade de um sistema seguem a abordagem do **Gerenciamento de Falhas**, mencionado anteriormente. **Prevenção de Falhas** (*Fault Avoidance*) é aplicada como forma de evitar que as falhas sejam introduzidas durante o projeto; **Remoção de Falhas** (*Fault Removal*) é aplicado durante o projeto do sistema como forma de evitar que as falhas introduzidas permaneçam quando o sistema entrar em operação; **Tolerância a Falhas** (*Fault Tolerance*) é aplicada para que falhas sejam toleradas (sistema se comporte de forma desejada) caso ativadas durante a operação.

Figura 12 ilustra um diagrama de Venn representando o universo de estados do sistema, destacando os estados defeituosos (em falha, erro ou disfunção) e seu fluxo de manifestação, desde a ativação de uma falha dormente no sistema até uma disfunção que expõe o sistema a uma situação perigosa, podendo levar à um acidente. O gerenciamento de falhas age sobre este fluxo de manifestação dos defeitos sobre regiões específicas: fronteira externa dos estados defeituosos, evitando que falhas latentes (ou condições de ativação) permaneçam no sistema ou sejam ativadas durante sua operação; e até antes da fronteira de disfunção, tolerando as falhas que tenham sido ativadas durante a operação. Em sistemas críticos em segurança, a tolerância

a falhas pode agir até a fronteira de estados inseguros, evitando que o sistema exponha os envolvidos a condições inseguras.

Figura 12 – Universo de estados do sistema e propagação de falhas
Fonte: adaptado de VISMARI (2007)

Além desta definição acadêmica para dependabilidade, o IEC criou na década de 1990 a Comissão Técnica 56 (IEC/TC56)[21] para desenvolver padrões internacionais sobre dependabilidade. A IEC/TC56 estruturou essas elaborações de normas em três (3) níveis: gestão, processo e ferramentas ou normas de suporte. A família de normas IEC 60300 (*Dependability Management*) é o ponto focal desses padrões, e as versões mais antigas do IEC 60300-1 definiram dependabilidade como sinônimo de desempenho de disponibilidade, que é afetado pela confiabilidade, capacidade de manutenção e desempenho do suporte à manutenção. As versões mais recentes do padrão IEC (2014 e posteriores) incorporam outras características (atributos) na definição de dependabilidade, conforme apresentado na Tabela 3.

Conforme definido pela IEC/TC56, dependabilidade é uma característica de qualidade relacionada ao tempo ('capacidade de desempenhar quando necessário'). Contudo, isso não significa que a dependabilidade lida com as mudanças não previstas no sistema durante seu desenvolvimento. A capacidade de os sistemas fornecerem serviços, conforme e quando necessário, que possam ser justificadamente confiáveis, pode ser afetada por sistemas que enfrentam mudanças, sobretudo não previstas.

[21] http://tc56.iec.ch/

Tabela 3 – Definição de dependabilidade (IEC ref. 192 01 02)[22]

dependability *<of an item>:*	*ability to perform as and when required*
	Note 1 to entry: Dependability includes **availability** (192-01-23), **reliability** (192-01-24), **recoverability** (192-01-25), **maintainability** (192-01-27), and **maintenance support performance** (192-01-29), and, in some cases, other characteristics such as **durability** (192-01-21), **safety and security**.
	Note 2 to entry: Dependability is used as a collective term for the time-related **quality** characteristics of an item.

A persistência da dependabilidade quando o sistema enfrenta mudanças é definida por LAPRIE (2008) como **resiliência**. O autor definiu o conceito de **mudanças** em três dimensões: mudanças pela sua **natureza** (funcional, ambiental ou tecnológica, podendo esta última dizer respeito a um ou a ambos hardware e software), por **perspectiva** (prevista, previsível e imprevisível) e por **horizonte de tempo** (curto, médio e longo prazo). Essa definição de resiliência está bem alinhada ao contexto de sistemas ciberfísicos e críticos em segurança, e é uma evolução natural da definição de dependabilidade.

Resiliência é um termo usado há muito tempo em diversos domínios de conhecimento (por exemplo, ecologia, ciência de materiais, negócios, psicologia e assim por diante). A resiliência pode ser definida como a capacidade do sistema de absorver perturbações e adaptar-se para manter as mesmas funções, estruturas e comportamentos (FOSTER, 1993). Em geral, um evento de perturbação pode ser potencialmente prejudicial a um ou mais elementos ou processos do sistema. Sobre condições nocivas e resiliência de sistemas, uma definição relacionada à segurança para resiliência é apresentada por LEVESON *et al.* (2006):

> "... a resiliência é a capacidade dos sistemas de prevenir ou se adaptar às mudanças nas condições, a fim de manter (controlar) uma propriedade do sistema (...) a propriedade que nos preocupa é a segurança ou o risco".

Pode-se considerar que atributos e ameaças são, em essência, os mesmos para dependabilidade e resiliência. No entanto, a dependabilidade implementada em um sistema irá tolerar as ameaças (falhas) consideradas no contexto operacional previsto para aquele sistema

[22] http://www.electropedia.org/iev/iev.nsf/display?openform&ievref=192-01-22

durante seu desenvolvimento. Caso as **características do ambiente ou do sistema** mudem ao longo da operação, o sistema pode não estar preparado para tolerar novas ameaças. Para garantir disponibilidade, confiabilidade, segurança, confidencialidade, integridade e manutenibilidade em quaisquer circunstâncias, o sistema deve estar preparado para **tolerar mudanças**, tanto internas quanto externas ao sistema, ao longo do seu ciclo de vida.

Sistemas críticos em segurança possuem ciclos de vida normatizados e bem definidos. As normas lidam com o conceito de segurança crítica ao longo de todo ciclo de vida do sistema. No caso da norma IEC 61508 (IEC, 2010), uma norma básica de segurança crítica para sistemas elétricos, eletrônicos e eletrônicos programáveis (E/E/PE), lida com a segurança crítica desde a fase de concepção até a desativação do sistema. Além disso, considera a segurança crítica tanto no nível do sistema quanto no nível do contexto geral onde o sistema estará inserido. Os perigos e riscos são identificados neste contexto, bem como os **requisitos gerais de segurança** (*overall safety requirements*) são especificados. Em seguida, os requisitos gerais de segurança são mapeados para o nível de sistema, que são refinados, desenvolvidos, implementados e validados. Após a validação da segurança do sistema, realiza-se a validação geral de seus requisitos para seu contexto. A Figura 13 ilustra o ciclo de vida de desenvolvimento de sistema crítico em segurança e os 2 níveis de abstração adotados (geral e de sistema).

As normas de segurança crítica atuais, bem como a Engenharia de Segurança de Sistemas (SSE), especificam abordagens tradicionais de desenvolvimento de sistemas: a definição de missão, fronteiras, interfaces e ambiente operacional previsto no início do desenvolvimento; a aplicação de abordagens reducionistas, com decomposição/integração hierárquica; a verificação, validação e avaliação dos sistemas contra a especificação dos requisitos do sistema; entre outros. Além disso, as técnicas e métodos aplicáveis nos processos de avaliação, verificação e validação (incluindo certificação) de segurança crítica são baseadas em decomposição funcional hierárquica – como exemplo, as Árvores de Falha (FTA), a Análise de Modos de Falha e seus Efeitos (FMEA), entre outras; em testes, ensaios, inspeções, modelagem e simulação.

Após a aplicação dos processos normativos e das abordagens tradicionais de SSE, incluindo as abordagens de desenvolvimento, avaliação e validação de segurança, e a partir da validação geral **de segurança** (*overall safety validation*), o sistema comissionado é colocado em operação. Desta forma, o sistema pode ser considerado seguro durante sua operação apenas

para a missão, fronteiras, interfaces, ambiente – incluindo os modos de operação e de falha considerados plausíveis – que foram identificados, **especificados**, projetados e implementados no início do ciclo de vida, durante o desenvolvimento do sistema. Desta forma, as abordagens Tradicionais de Engenharia estão aptas a garantir a segurança crítica para sistemas cujas características **previstas e implementadas** – tanto do sistema quanto do seu ambiente – permaneçam inalteradas ao longo de toda sua vida útil.

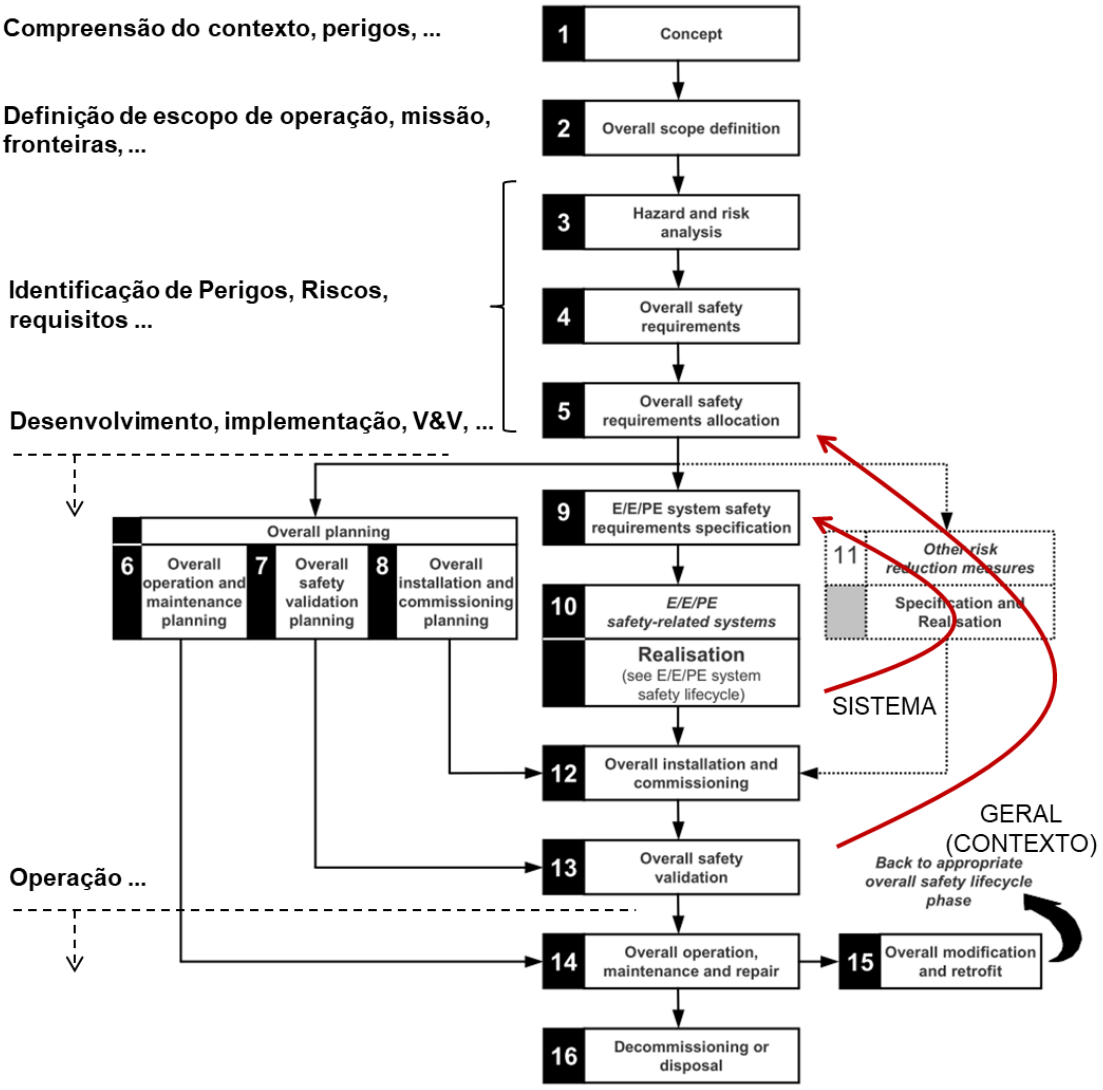


Figura 13 – Ciclo de vida do sistema crítico em segurança (especificação e validação)
Fonte: adaptado de IEC (2010)

Porém, a Engenharia de Sistemas – e, por consequência, a Engenharia de Segurança de Sistemas – não está apta a garantir a segurança de sistemas de engenharia que sofram mudanças não previstas ou/e imprevisíveis, tanto internas (sistema) quanto externas (ambiente), ao longo da

operação continuada destes sistemas. Nestes casos, sobretudo para Sistemas Complexos de Engenharia, novas abordagens são necessárias.

2.2. Sistemas complexos de engenharia (CoES)

Os termos '*complexidade*' e '*complexo*', da mesma forma que ocorre há tempos com o termo '*sistema*', têm sido amplamente utilizadas ao longo dos últimos anos nos mais diversos contextos e aplicações. Combinações como '*estrutura complexa*', '*redes complexas*', '*complexidade algorítmica*', '*gerenciamento complexo*', '*sistema complexo*', '*complexidade da economia*', '*tema complexo*', entre outros, são comuns no dia-a-dia tanto em círculos especializados de conhecimento – como engenharia e ciências – quanto ao público geral.

Observa-se na bibliografia que não há um consenso quanto à definição dos termos '*complexidade*' ou de '*sistemas complexos*'. Segundo os dicionários da língua portuguesa[23], as definições de 'complexo' e 'complexidade' são:

- ✓ ***complexo***:

  > *Diz-se de ou conjunto, tomado como um todo mais ou menos coerente, cujos componentes funcionam entre si em numerosas relações de interdependência ou de subordinação, de apreensão muitas vezes difícil pelo intelecto*.
  > *Que abrange ou contém muitos elementos ou aspectos diversos, com diferentes formas de inter-relação, às vezes de difícil apreensão ou compreensão*.
  > *Complicado, difícil [antôn.: descomplicado, fácil]*.
  > *A que falta clareza; CONFUSO; OBSCURO [Antôn.: claro, compreensível]*
  > *Passível de ser encarado ou apreciado sob diversos ângulos*.

- ✓ ***complexidade***: qualidade, condição ou estado do que é complexo.

Para o termo '*complexo*', observa-se que ele é definido pela combinação de um numeroso **conjunto de elementos**, tanto em quantidade quanto em tipos, e de **relacionamentos** entre estes elementos, levando a uma **estrutura** e a um **comportamento**, muitas vezes, **difícil de compreender e/ou prever**.

Assim, dado que um '*sistema*', por definição, é um **conjunto de elementos** (funcionais) que interagem com o objetivo de cumprir uma missão ou produzir um resultado, conclui-se que

[23] Foram consultados os dicionários Priberiam, Houaiss e Aulete. As definições foram combinadas na forma apresentada neste texto.

as definições de ‘complexo’ (e complexidade) são vinculadas ao conceito de ‘*sistema complexo*’, conforme observado por XIONG (2011).

Com base nestas definições de dicionário, entende-se que um sistema complexo é um sistema cujas estruturas e comportamentos são difíceis de compreender e/ou prever. E esta dificuldade estaria mais relacionada às limitações do intelecto humano do que a características intrínsecas do sistema complexo. Portanto, duas questões podem ser elaboradas:

i. A ‘*complexidade*’ dependeria (apenas) da capacidade intelectual do observador?
ii. O que torna um sistema ‘*complexo*’ (i.e. “*difícil de compreender e/ou prever*”)?

NORMAN (2010) apresenta uma distinção entre os conceitos de ‘*complexo*’ e ‘*complicado*’, definindo o escopo de aplicação para cada termo. Ele cita que a palavra “*complexidade*” descreve um *estado do mundo* (‘*state of the world*’). Já a palavra ‘*complicada*’ descreve um *estado da mente* (‘*state of the mind*’), que é o “*psychological state of a person in attempting to understand, use, or interact with something in the world*”. TESLER e SAFFER (apud NORMAN (2010)) definem que cada aplicação possui uma quantidade inerente *complexidade irredutível* (‘*irreducible complexity*’), independente do observador, e que a questão principal é quem deve lidar com ela: o usuário ou o desenvolvedor (programador ou engenheiro).

SOMMERVILLE (2012) lida com as questões acima elaboradas e divide a complexidade em dois grupos: *complexidade epistêmica* e *complexidade inerente*. A ***complexidade epistêmica*** (relativo ao intelecto, ao conhecimento; cognitivo) cresce (apenas) com o tamanho do sistema e, como é difícil conhecer o sistema como um todo em detalhes (devido ao seu tamanho), é difícil prever/explicar seu comportamento. Já na ***complexidade inerente***, a dificuldade de prever/explicar o comportamento não depende apenas do tamanho do sistema, mas de outras **características**. Contudo, em ambos os casos, a **complexidade** está relacionada às **limitações em se prever os resultados (comportamento)** produzido por um sistema.

Buscou-se na bibliografia pelas características – além do tamanho do sistema – que estão relacionadas à complexidade inerente. Como resultado, o único consenso encontrado é que não há consenso sobre uma definição universal ou conjunto de características que definem um sistema (inerentemente) complexo. Percebe-se que as características consideradas em uma

definição de sistema complexo são enviesadas pela área de conhecimento que a define (por exemplo, Matemática, Biologia, Economia, Sociologia, Engenharia, entre outras). As definições mais representativas são:

> "*Sistemas complexos são sistemas que compreendem muitas partes interativas com a capacidade de gerar uma nova qualidade de comportamento coletivo macroscópico cujas manifestações são a formação espontânea de estruturas temporais, espaciais ou funcionais distintas... de componentes interagindo simultânea e não linearmente uns com os outros e com seu ambiente em vários níveis – e na rica diversidade de comportamento de que são capazes.*" (MURPHY; ELLIS; O'CONNOR, 2009)

> "*Sistema complexo é um sistema com múltiplos componentes interativos, dos quais o comportamento geral não pode ser inferido (facilmente explicado) simplesmente a partir do comportamento (especificando o papel/regra) dos componentes, mas emerge da interação de seus componentes e da interação entre ele e seu ambiente. Este conceito contrasta com a máquina tradicional ou constructos Newtonianos, que assumem que todas as partes de um sistema podem ser conhecidas, que o planejamento detalhado produz resultados previsíveis e que a informação flui ao longo de um caminho predeterminado*. " (XIONG, 2011)

> "*Um sistema complexo consiste em um grande número de agentes conectados/interagindo que, como um todo, exibem um comportamento coordenado sem nenhum controle centralizado. Ou seja, um sistema complexo exibe propriedades emergentes que obviamente não decorrem das propriedades dos agentes individuais*" (BOCCARA, 2010)

> "*Sistemas complexos são sistemas que não possuem uma autoridade centralizada e não são projetados a partir de uma especificação conhecida, mas envolvem diferentes partes interessadas criando sistemas que são funcionais para outros propósitos e só são reunidos no sistema complexo porque os "agentes" individuais do sistema veem tal cooperação como sendo benéfica para eles*" (SHEARD; MOSTASHARI, 2009)

> "*Sistemas complexos são formados por elementos independentemente desenvolvidos e controlados, cujo comportamento é regido por negociação (competição ou cooperação) entre os elementos e sua adaptação no ambiente (não há controle centralizado)*". (CLEARY, 2005)

Ao avaliar as definições identificadas na literatura, pode-se listar o seguinte conjunto de características comuns e intrínsecas aos sistemas **inerentemente complexos:**

a. **Não possuem autoridade ou controle centralizado** (não há "entidade onisciente");
b. **Cada elemento** (agente/subsistema) no ambiente **decide** a forma de **interagir** (**cooperativa** ou **competitivamente)** com os demais elementos do ambiente como forma de cumprir sua missão. Esses elementos executam suas **regras**, em nível local, de forma a ter benefícios pessoais.

c. Formação espontânea (**auto-organização**) de estruturas **temporais**, **espaciais** e **funcionais** distintas – formação de **grupos de hierarquia** e de diferentes **escalas**, que influenciam a evolução e o comportamento do sistema.
d. Sistema não é projetado a partir de uma especificação.
e. **Fronteiras difusas**. Elementos influenciam (*upward causation*) e são influenciados (*downward causation*) pelo ambiente.
f. Interações **não lineares** e entre **múltiplos níveis de abstração** ([*feedbacks*]^{n}).
g. Sensibilidade às condições iniciais.
h. Sensibilidade à mudança.
i. **Comportamento coordenado (EMERGENTE) sem qualquer controle centralizado**
j. Elementos podem **evoluir** (de forma independente) para se adaptar ao ambiente.

A Figura 14 ilustra as principais características que compõe um sistema complexo. Partindo de um grande número de elementos (sistemas) simples que **interagem** dinamicamente (de forma não linear e possivelmente efêmera) em múltiplos níveis de abstração e escalas (níveis hierárquicos), **evoluindo** para a formação estruturas auto-organizadas e auto-reguladas (homeostáticas) que representam sistemas abertos e dissipativos que produzem **comportamentos emergentes** (que não podem ser inferidos do comportamento de seus elementos individuais). Estes sistemas complexos apresentam comportamentos comuns ao longo de **escalas de tempo e espaço**, ao longo de sistemas e de disciplinas.

Mesmo sem identificar um consenso sobre uma definição universal ou conjunto de características que caracterizem um sistema (inerentemente) complexo, o **comportamento (propriedade) emergente** é citado pela quase totalidade dos autores. Por definição, espera-se um comportamento 'emergente' ("*que emerge, que resulta ou procede*"*;* "*que sai de, se origina em, é consequência de*") para qualquer tipo de 'sistema' ('*combinação de partes reunidas para concorrerem para um resultado*', '*conjunto de elementos, concretos ou abstratos, interligados e que funciona como um todo estruturalmente constituído*'). No caso específico dos sistemas complexos, BOCCARA (2010), alinhado com diversos autores, menciona que o surgimento de propriedade [comportamento]" emergente é a característica essencial dos sistemas complexos. Ele define que:

> **"***Propriedades emergentes são efeitos em larga escala de um sistema de agentes que interagem localmente que são muitas vezes surpreendentes e difíceis de prever, mesmo no caso de interações simples.***".** (BOCCARA, 2010)

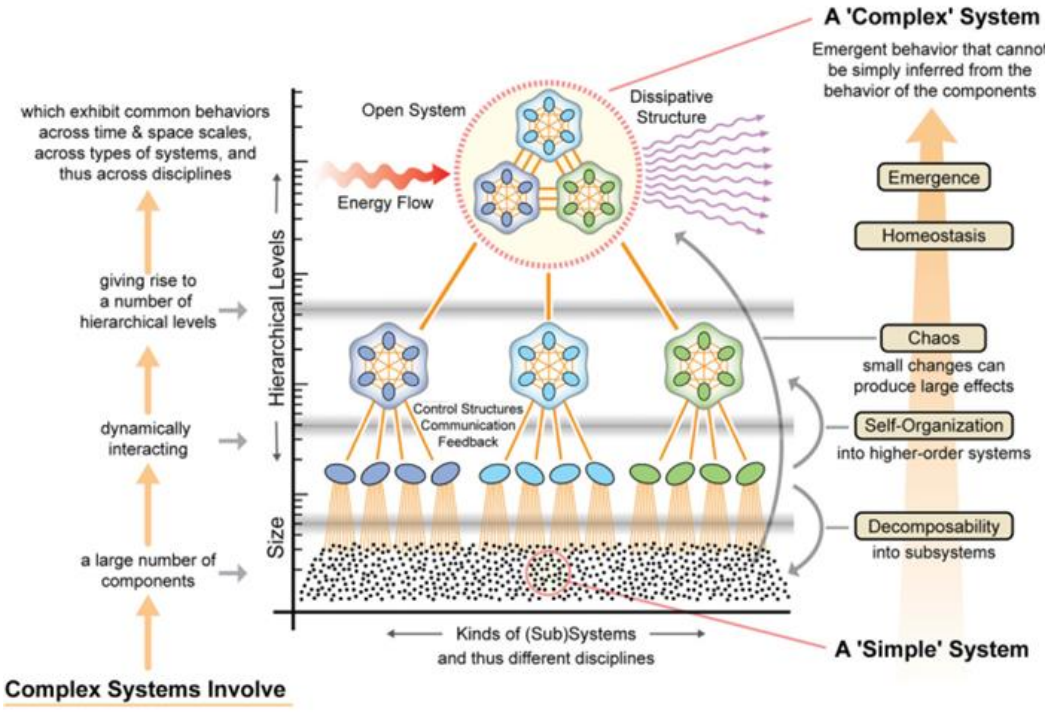


Figura 14 – Características dos sistemas complexos
Fonte: (SHEARD; MOSTASHARI, 2009)

Ou seja, o conceito de propriedade emergente em sistemas complexos se diferencia da emergência observada em sistemas 'não complexos' pela dificuldade em ser prevista (incerteza), tanto no tempo quanto no espaço. BEDAU (1997) define – e HOLLNAGEL, WOODS e LEVESON (2006) apresentam – três (03) tipos de emergência: **Nominal**, **Fraca** (*Weak*) e **Forte** (*Strong*). Para cada uma delas, temos (HOLLNAGEL; WOODS; LEVESON, 2006):

> **Emergência nominal:** "quando as propriedades de nível macro, embora sem sentido no nível micro, podem ser derivadas da montagem de propriedades de nível micro"
>
> **Emergência fraca**: "quando uma explicação microscópica, 'em princípio', do comportamento macroscópico ainda é possível, mas o comportamento detalhado e abrangente não pode ser previsto sem a realização de uma simulação um-para-um, porque não há explicação condensada da dinâmica causal do sistema"
>
> **Emergência forte**: "*quando as propriedades de nível macro não podem ser explicadas, e menos ainda previstas, mesmo em princípio, por qualquer causalidade de nível micro. A existência de emergência forte é uma questão controversa, pois é inconsistente com o dogma científico comum da causalidade ascendente (upward causation) e introduz uma presunção de que a causalidade descendente (downward ccausation) holística liga as leis subjacentes da física para impor princípios de organização sobre os componentes…. Ele afirma que um sistema acima de um nível específico de complexidade … não pode ser totalmente controlado por causalidade*

> *ascendente, devido à existência de limites superiores fundamentais no conteúdo da informação e na taxa de processamento da informação*".

A emergência **Nominal** é o tipo de emergência apresentado por sistemas ordinário ou 'não complexos'. Neles, é possível modelar sua dinâmica e a relação causal entre o comportamento dos elementos e suas consequências (conhecido como '*upward causation*', ou a relação causa-efeito dos elementos para a fronteira do sistema). Sistemas que apresentam este tipo de emergência podem ser projetados e desenvolvidos utilizando as abordagens clássicas (tradicionais) de engenharia, como o reducionismo (abordagens "*Top-Down*").

No caso da emergência 'fraca', ainda é possível caracterizar os efeitos do comportamento microscópico (em baixos níveis de abstração do sistema) sobre o sistema (altos níveis de abstração). Mas, diferentemente da emergência nominal, não é possível modelar a dinâmica causal do sistema e, consequentemente, prever o comportamento emergente de forma literal. Apenas é possível aplicar uma abordagem *bottom-up*, avaliando o comportamento do sistema por meio de simulações "um-para-um".

Sistemas que apresentam a emergência 'forte' são um caso à parte, atípico em sistemas de engenharia. Nestes sistemas, os níveis mais altos de abstração do sistema impõem as características dos níveis inferiores, onde o ambiente e a estrutura dos sistemas impõem os princípios de organização sobre seus componentes (conhecido como '*downward causation*'). JOHNSON (2006) apresenta uma distinção entre os conceitos de emergência Forte (*Strong*) e Fraca (*Weak*), onde:

> "*... distinções entre emergência fraca e forte também podem ser caracterizadas em termos de relações causais entre os diferentes níveis de um sistema complexo ... emergência fraca pode ser analisada usando técnicas reducionistas onde comportamentos complexos em um nível de sistema são causados por propriedades de componentes subjacentes ... emergência forte relaciona-se com uma forma de causalidade descendente' (downward causality) onde comportamentos em níveis mais baixos em um sistema são limitados por características de nível mais alto.*." (JOHNSON, 2006)

Neste trabalho, a investigação não se restringiu apenas ao domínio da engenharia. Assim, foi possível estender a investigação a um escopo amplo de domínios, tanto de sistemas naturais (maior parte da literatura em sistemas complexos) quanto artificiais (feitos pelo homem), obtendo um arcabouço bibliográfico. Dele, foram extraídos os conceitos e características fundamentais de sistemas complexos aplicáveis ao domínio da engenharia (sistemas de engenharia). Além do **comportamento (propriedade) emergente**, as demais características de sistemas complexos identificadas e listadas anteriormente são aplicáveis a qualquer sistema no

escopo da engenharia. Ou seja, em Sistemas Complexo de Engenharia, observam-se as seguintes características inerentes:

- o **controle e evolução não coordenada (sem gerência ou autoridade centralizada)** de partes constituintes dos sistemas;
- suas **interações dinâmicas, não lineares e por meio de competição e cooperação** seguindo regras locais;
- a formação (quase) espontânea de estruturas temporais, espaciais e funcionais como forma de cumprir uma missão, **não projetado a partir de uma especificação**;
- a existência de fronteiras difusas, não muito bem definidas entre sistemas.

Como discutido anteriormente, as abordagens de Engenharia de Segurança de Sistemas Tradicionais (TSE) são limitadas em garantir a segurança crítica de sistemas de engenharia que sofram **mudanças** não previstas ou/e imprevisíveis, tanto internas (sistema) quanto externas (ambiente), ao longo da operação destes sistemas. Isso é decorrente das condições de contorno necessárias para a aplicação das TSE, como haver fronteiras e condições estáveis e bem definidas; considerar que o sistema interage com o ambiente, mas não o altera (e vice-versa); e, principalmente, ser possível aplicar o reducionismo e decomposição hierárquica.

Sistemas complexos de engenharia (CoES) não possuem as condições mínimas necessárias para aplicação de TSE. CoES são sistemas intrinsecamente evolutivos, sofrendo mudanças ao longo de seu ciclo de vida – que não são projetadas a partir de uma especificação ou de um processo centralizado – e formando estruturas para cumprimento de missão. Segundo NORMAN (2004), as diferenças entre os mecanismos primários de mudança em engenharia de sistemas tradicionais (TSE) e de engenharia de sistemas complexos (CSE) são:

- TSE: requisitos geram mudanças, implementados por abordagem "*Top-Down*" ou "*Structured Design*";
- CSE: evolução geram mudanças, orientadas por abordagem "*Bottom-up*", onde pressões seletivas locais (recompensa/punição) levam a soluções globais.

Desta forma, uma abordagem que promova modificações (evolução) do sistema orientada pelo nível de risco de segurança observado na aplicação tem potencial de aplicação nos processos de garantia de segurança crítica dos Sistemas Complexos de Engenharia.

2.3. Sistemas de Transporte Inteligente Cooperativos (C-ITS)

Avanços nas Tecnologias de Informação e Comunicação (ICT – *Information and Communication Technologies*), sobretudo na Computação Pervasiva e na Inteligência Artificial / Aprendizado de Máquina (AI/ML – *Artificial Intelligence / Machine Learning*), têm afetado todos os domínios da vida cotidiana do ser humano, inclusive aqueles domínios onde podem ocorrer mortes e danos à saúde, ao meio ambiente e ao patrimônio. Particularmente, os transportes são um dos domínios críticos em segurança cujos paradigmas de sistemas estão sendo profundamente impactados por esta onda tecnológica.

Nos Transportes Aéreos, o paradigma de sistemas de controle de tráfego vem evoluindo para um conceito de **alto nível de automação e plenamente baseado em ICT**, nos quais **funcionalidades,** processos e serviços **vêm sendo implementados** seguindo um modelo de referência de interconexão baseado no modelo ISO/OSI. Neste modelo, os elementos do sistema são definidos como nós de uma rede de dados e denominados como Sistemas Finais (*End Systems* - ES), que interagem entre si por meio de protocolos e serviços de interfaces comuns (padronizadas) para o domínio de aplicação.

Como exemplo, a Figura 15 apresenta o modelo conceitual de referência desta rede de dados, definida e denominada pela Organização da Aviação Civil Internacional (OACI)[24] como Rede de Telecomunicações Aeronáuticas (ATN). Observa-se que a ATN é uma rede de comunicação entre ES (aeronaves, centros de controle, entre outros) que contenham as mesmas camadas de implementação do modelo e os mesmos processos de aplicação, representado pelas funcionalidades ATN utilizadas pelos ES e definidas na **camada 7** (*aplicação layer*) do modelo ISO/OSI. Assim, esta evolução de paradigma nos sistemas de tráfego aéreo – denominada como CNS/ATM (*Communication, Navigation, Surveillance / Air Traffic Management*), em alusão às principais funcionalidades do controle de tráfego (Comunicar, Navegar e Vigiar como forma de Gerenciar o Tráfego Aéreo) – possibilita o aumento da capacidade de transporte aéreo e de sua segurança operacional do espaço aéreo (FLAVIO VISMARI; CAMARGO JUNIOR, 2011; VISMARI, 2007).

[24] A OACI é uma agência especializada da Organização das Nações Unidas (ONU). Com a participação dos seus Estados membros, coordena os princípios e técnicas da navegação aérea internacional e promove o planejamento e o desenvolvimento do transporte aéreo internacional para garantir um crescimento seguro e ordenado (https://www.icao.int/).

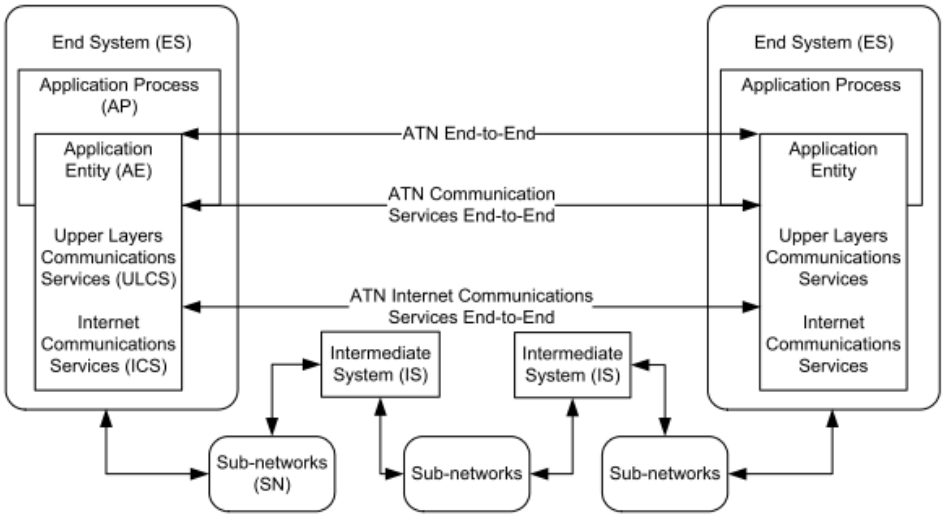


Figura 15 – Modelo conceitual de referência para a ATN

Fonte: (ICAO, 1999)

Nos Transportes Terrestres rodoviários[25], observa-se uma mudança de paradigma de sistemas análoga à ocorrida nos Transportes Aéreos com o CNS/ATM – também um conceito de **alto nível de automação** – denominado como **Sistemas de Transporte Inteligente (ITS)**. Contudo, diferentemente dos sistemas de transporte aéreo, os transportes terrestres têm caráter regional e, consequentemente, não existe uma agência ou autoridade em nível mundial que coordene ou promova o planejamento dos ITS. Desta forma, cada região ou Estado possui suas próprias iniciativas e programas para o desenvolvimento dos Sistemas de Transporte Inteligente. Atualmente, destacam-se as iniciativas do Departamento de Transportes (DoT) dos Estados Unidos da América (USA)[26], da União Européia (EU)[27] e da região da Ásia-Pacífico[28].

Ao avaliar as definições apresentadas por estas iniciativas, observa-se que os conceitos de ITS são similares entre si. Fazendo uso integrado das ICTs – sobretudo redes de comunicação de dados – na infraestrutura de transportes e nos veículos, os ITS têm como missão fornecer **serviços** relacionados a diferentes modais de transporte e de gerenciamento de tráfego,

[25] Mesmo que sua definição se aplique a diferentes modais de transporte, o ITS tem sido desenvolvido para os contextos de transportes terrestres que possam se movimentar com mais de um grau de liberdade. Inclusive, a própria Diretiva 2010/40/EU da União Europeia (COUNCIL EUROPEAN UNION, 2010) direciona sua aplicação para "*... road transport and for interfaces with other modes of transport*". Assim, neste trabalho, utiliza-se o termo 'rodoviário' como forma de **excluir** os modais de transporte terrestre guiados, como o transporte sobre trilhos (metro-ferroviário), em tubulações ('dutoviário'), entre outros.

[26] https://www.its.dot.gov/about/its_jpo.htm

[27] https://transport.ec.europa.eu/transport-themes/intelligent-transport-systems_en

[28] http://itsasia-pacific.com/

permitindo que usuários tenham maior consciência situacional e façam uso mais seguro, coordenado e inteligente das redes de transporte (COUNCIL EUROPEAN UNION, 2010).

Arquiteturas de referência têm sido elaboradas para orientar o planejamento, definição e integração dos sistemas de transporte no conceito ITS. Como exemplos, pode-se citar as iniciativas do DoT-USA – a ARC-IT (*Architecture Reference for Cooperative and Intelligent Transportation*)[29] e da EU – a FRAME (*European Intelligent Transport Systems (ITS) Framework Architecture*)[30]. Em ambos, uma abordagem de engenharia de sistemas é aplicada para definir objetos, relacionamentos, funções e serviços, entre outros.

Na ARC-IT (USA), estas definições são obtidas por visões de alto nível: **Empresarial** (relacionamento entre organizações e as regras que são seguidas), **Funcional** (descrevem elementos funcionais – processos – e suas interações lógicas – fluxos de dados – que satisfazem os requisitos do sistema), **Física** (descreve os objetos físicos – sistemas e dispositivos – e as interfaces de alto nível entre objetos) e **Comunicação** (descreve os **protocolos em camadas** utilizados para comunicação de dados entre objetos físicos).

Como exemplo, a Figura 16 ilustra a arquitetura de visão física do ARC-IT, com suas classes de objetos físicos (veículos, elementos de campo, pessoas e centros) e interconexões (comunicação entre objetos). Nesta arquitetura, observa-se que todas as classes de objetos físicos do ITS se interconectam umas com as outras e consigo mesmas. **Esta capacidade de duas ou mais classes de objetos ITS interagirem e cooperarem entre si e habilitar o provimento de serviços ITS aprimorados e de melhor qualidade é denominada "ITS Cooperativa" (C-ITS)**.

No caso da arquitetura de referência norte-americana (ARC-IT), o conceito de C-ITS é intrínseco à própria arquitetura, onde cada pacote de serviço (por exemplo, Operações de Veículos Automatizados) definem os objetos, interfaces, processos e fluxos de dados necessário para seu provimento. Consequentemente, a Comunicação contém a definição das soluções (conjunto de normas e protocolos) necessárias para suportar o fluxo de dados do serviço.

[29] https://www.arc-it.net/
[30] https://frame-online.eu/

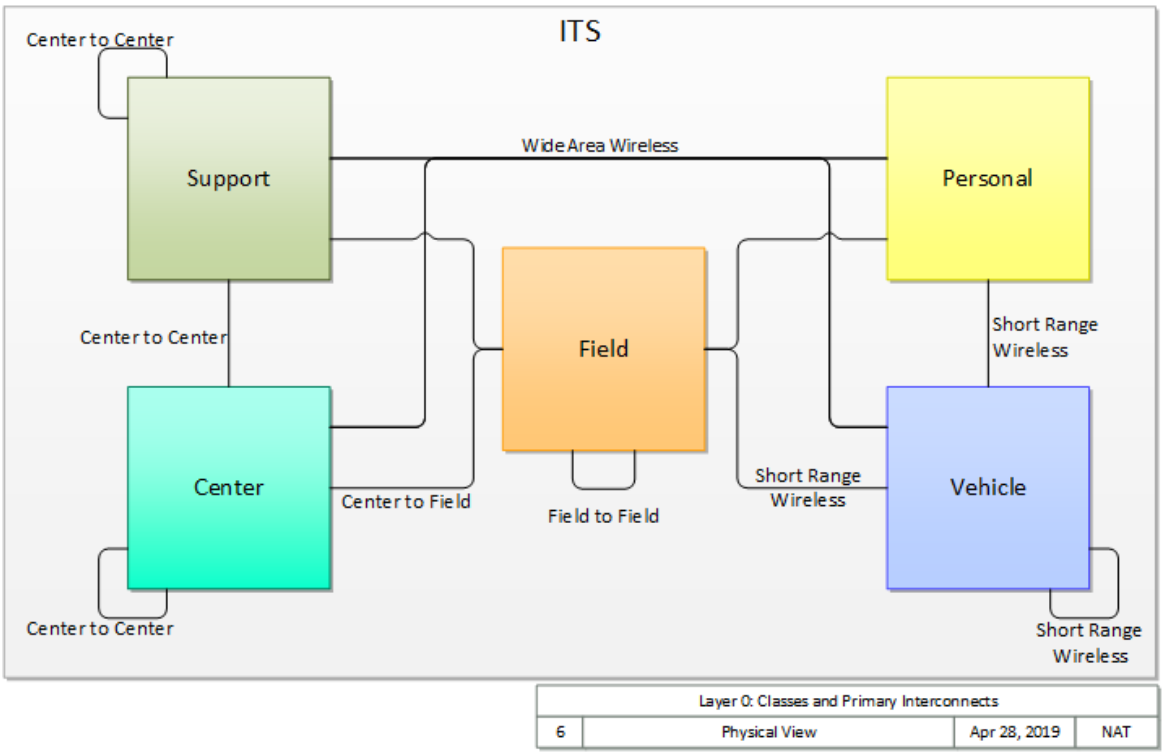


Figura 16 – Visão física do ARC-IT: classes de objetos físicos e interconexões
Fonte: (US-DOT, 2022)

Por outro lado, a União Europeia[31] define formalmente o C-ITS como uma categoria de serviço ITS baseada em redes de dados abertas e heterogêneas que possibilita aos veículos interagirem entre si e com a infraestrutura viária, possibilitando uma grande variedade de serviços de cooperação e informação. Tipicamente, observa-se a possibilidade de os **veículos** interagirem entre si (**comunicação V2V**); com **pessoas** – ciclistas pedestres, dispositivos pessoais, entre outros (**comunicação V2P**); e com **elementos de campo** – **equipamentos à margem de via (RSU),** semáforos, sinalizações, entre outros (**comunicação V2I** ou **V2N**).

Atualmente, dois grupos de tecnologias de acesso são utilizadas para a comunicação entre veículos e os demais elementos do sistema de transporte – denominado como comunicação V2X (*Vehicle-to-Anything*):

- **Comunicação Dedicada de Curto Alcance** (DSRC – *Dedicated Short Range Communication*). No DSRC, a comunicação é realizada de forma direta entre elementos do ITS, formando topologias espontâneas sem necessidade de infraestrutura de rede. A DSRC – denominada nos EUA como WAVE (*Wireless Access in Vehicular*

[31] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=PI_COM:C(2019)1789

Environments), e como ITS-G5 na EU – é baseada no protocolo IEEE 802.11p (ARENA; PAU; SEVERINO, 2020). DSRC é empregado em V2V, V2I e V2P.

- **Comunicação baseadas em redes Celulares** (*Cellular V2X* ou C-V2X). Em C-V2X, a comunicação entre elementos é realizada por **meio de uma infraestrutura de comunicação da rede celular**. Tecnologias baseadas em C-V2X são orientadas pelas normas desenvolvidas pelo 3GPP, e incluem protocolos da segunda (GSM/2G), terceira (UMTS/3G), quarta (LTE) e quinta (5G-NR) gerações de redes celulares. C-V2X é uma comunicação de longo alcance, empregado em V2N.

Algumas instituições de normatização vêm produzindo normas e especificações técnicas para definir os **requisitos de comunicação para serviços V2X**. Atualmente, as principais instituições são o IEEE e o ETSI, onde:

- O IEEE (*Institute of Electrical and Electronic Engineers*) vem desenvolvendo as normas IEEE 802.11p (ARENA; PAU; SEVERINO, 2020) e 802.11bd , definindo os protocolos WAVE e ITS-G5 baseados em tecnologias de comunicação WIFI.

- O ETSI (*European Telecommunications Standards Institute*) vem produzindo especificações técnicas (TS) para tecnologias de comunicação V2X celulares (C-V2X), como a TS 22.185 (ETSI, 2017) – com foco em aspectos básicos de segurança crítica usando tecnologias 4G-LTE – e a TS 22.186 (ETSI, 2022) – com foco em cenários de caso de uso V2X aprimorados baseados em tecnologia 5G.

Individualmente, DSRC e C-V2X não conseguem satisfazer a todos os requisitos funcionais e operacionais necessários ao atendimento das aplicações C-ITS (SEPULCRE; GOZALVEZ, 2018). Desta forma, são necessárias redes de dados abertas e heterogêneas que possibilitem interoperabilidade e conectividade entre todos os elementos do sistema e, consequentemente, a troca de **mensagens C-ITS** necessárias às aplicações do C-ITS. Sobretudo, Mensagens Básicas de Segurança (BSM) (SAE, 2022) e Mensagens de Consciência Cooperativa (CAM) (ETSI, 2019) – cuja equivalência entre os conjuntos de dados BSM e CAM pode ser observada em KIM *et al.* (2018) – fundamentais à obtenção de consciência situacional pelos envolvidos, permitindo uma operação coordenada e segura do sistema de transporte.

3. UMA ABORDAGEM ORIENTADA A RISCOS PARA O GERENCIAMENTO DE RECURSOS EM C-ITS

"Complex systems engineering is not a new or renewed attention to details; it is an attention to the overall coherence"
(NORMAN; KURAS, 2004)

Conforme discutido anteriormente, **Sistemas Complexos de Engenharia (CoES)** e críticos em segurança apresentam limitações à aplicação das abordagens atuais de **Engenharia de Segurança de Sistemas (SSE)**. Portanto, o objetivo desta pesquisa é contribuir com uma abordagem conceitual de garantia de segurança crítica no contexto de Engenharia de Sistemas Complexos.

Nesta seção, propõem-se uma abordagem **conceitual** de garantia de segurança crítica para CoES. Esta abordagem utiliza características intrínsecas dos sistemas complexos, as quais são identificadas e justificadas. Quando utilizada durante a operação de um **CoES** – nesta pesquisa, os sistemas de transporte no contexto de Sistemas de Transporte Inteligentes Cooperativos (C-ITS), esta abordagem promove a obtenção de níveis aceitáveis de risco de segurança ao contexto de aplicação do sistema, gerenciando as configurações dos recursos do sistema em função dos riscos de segurança observados durante a operação.

3.1. CoES e os sistemas autônomos cooperativos

Avanços nas Tecnologias de Informação e Comunicação (ICT), sobretudo na computação móvel e pervasiva e na Inteligência Artificial / Aprendizado de Máquina (AI/ML), têm evoluído o paradigma dos sistemas de transporte para um conceito de alto nível de automação, tanto no escopo dos veículos quanto do sistema de transporte. A expectativa é de que este paradigma, definido como Sistemas de Transporte Inteligentes (ITS) – incluindo a categoria de serviço ITS baseada em redes de dados abertas e heterogêneas que possibilitam aos veículos interagirem entre si e com a infraestrutura viária de forma colaborativa (C-ITS) – promova a segurança a eficiência e o desempenho ambiental dos transportes rodoviários.

Neste contexto, os veículos autônomos (AV – *Autonomous Vehicles*) podem ser considerados o mais promissor novo elemento pertencente aos paradigmas de Sistemas de Transporte Inteligentes. Espera-se que os veículos autônomos – e mesmo os veículos altamente

automatizados (HAV – *Highly Automated Vehicles*) (SAE, 2014)– representem vantagens para a sociedade, sobretudo a redução de riscos de segurança e o aumento da eficiência e produtividade dos sistemas de transporte. Relativo aos sistemas de transporte rodoviário, e dado que grande parte dos acidentes – aproximadamente 90% dos acidentes são atribuídos ao erro humano segundo SINGH (2015), espera-se que o emprego de veículos autônomos nos sistemas de transporte promova uma redução significativa na quantidade e/ou severidade dos acidentes de trânsito.

Apesar de promissores, os **veículos autônomos** – assim como qualquer novo sistema, conceito ou tecnologia de aplicação crítica – serão incorporados à vida humana diária apenas se seus benefícios tornarem aceitáveis os riscos de segurança associados. Desta forma, deve ser possível garantir que os futuros sistemas de transporte autônomo sejam seguros ao longo de todo o ciclo de vida, sobretudo durante a operação. Por isso, e mesmo antes que possam ser colocados em operação, é mandatório compreender a relação entre estes novos conceitos e tecnologias e os riscos de segurança crítica das aplicações autônomas.

Tanto a Comunicação Cooperativa (*Cooperative Communication* – CC) quanto a Inteligência Artifical (AI – *Artificial Intelligence*) / Aprendizado de Máquina (ML – *Machine Intelligence*) são dois conceitos habilitadores (pilares) dos sistemas autônomos, sobretudo os sistemas de transporte. A autonomia demandada por estes sistemas exige interação com o meio e raciocínio/aprendizado para tomada de decisão, sobretudo em situações imprevistas.

No caso dos Sistemas de Transporte Inteligentes Colaborativos (C-ITS), os veículos autônomos interagem entre si e com os demais elementos fisicamente distribuídos do ambiente por meio de comunicação V2X (*Vehicle to Everything*), provida por uma **infraestrutura de comunicação de dados**. Quanto a AI/ML, estas tecnologias permitem que os elementos autônomos dos sistemas tenham a capacidade de **aprender** com o meio e tomar decisões em situações não previstas durante o desenvolvimento, sobretudo situações imprevisíveis (BALLINGALL; SARVI; SWEATMAN, 2020; PICARDI *et al.*, 2020).

Desta forma, ao fazer uso da **comunicação cooperativa** e da **AI/ML**, os elementos autônomos do sistema podem agir de forma coordenada e cooperativa – em nível local (elemento) – cumprindo os objetivos do sistema no nível da aplicação mesmo em meio a situações imprevistas. Além disso, pode-se observar que este novo paradigma de sistemas de

transporte possui outras características intrínsecas de **complexidade inerente**, discutidas na 2.2, pois:

- Sistemas baseados em AI/ML (nos quais há aprendizado) podem operar de forma **não determinística**, produzindo resultados (comportamentos) diferentes quando submetidos a um mesmo conjunto de entradas. Isso pode ser resultado da **sensibilidade à mudança** (e **às condições iniciais**, comparando-se os processos de treinamento e operação do sistema). Nestes sistemas, pequenas alterações no conjunto de dados de entrada em relação àqueles utilizados para seu treinamento (por exemplo, mudança de poucos pixels em uma imagem) pode produzir resultados diferentes do esperado (por exemplo, falso positivo ou falso negativo no reconhecimento de uma imagem (LU *et al.*, 2017; RAJABLI *et al.*, 2020).
- **Falta de autoridade ou controle centralizado**, típica dos sistemas de tráfego rodoviários (em comparação com outros sistemas de transporte, como o aéreo ou ferroviário), pode ser maximizada devido a capacidade de autonomia dos agentes do sistema de tráfego. Cada elemento **toma decisões, em nível local** e de forma **cooperativa ou competitiva**, em função de suas regras e/ou outros valores (que podem evoluir com o tempo), em benefício próprio, para cumprir sua missão.
- A Comunicação Cooperativa entre os elementos de transporte (e com outros sistemas além do transporte) tende a intensificar **interações não lineares** e entre **múltiplos níveis de abstração**. Além disso, tendem a promover **fronteiras difusas**, pois seus elementos influenciam e são influenciados pelo ambiente, que pode ser qualquer elemento além da fronteira do próprio elemento de transporte (veículo ou outro).
- A **formação espontânea de estruturas** (sobretudo temporais e espaciais) já é característica inerente de sistemas de tráfego rodoviário tradicionais, baseados em operadores humanos. Contudo, este fenômeno tende a ser impactado pelo aumento local (sobre os elementos de transporte) da consciência situacional (em tempo e espaço) do ambiente trazido pela Comunicação Cooperativa.
- O aprendizado de máquina (*ML*) permite que os sistemas (e seus elementos) possam **evoluir**, de forma independente aos demais, para se adaptarem ao ambiente. E sua evolução pode tanto influenciar quanto ser influenciada pelo ambiente, em diversos níveis, devido à Comunicação Cooperativa.

Desta forma, **sistemas autônomos cooperativos** – sobretudo no domínio dos transportes – podem ser classificados como **Sistemas Complexos de Engenharia** (*Complex Engineered Systems – CoES*). Assim, as características de **complexidade inerente**, juntamente com a **Comunicação Cooperativa** e a **AI/ML**, tornam-se os pilares deste paradigma de sistema. A relação entre os conceitos relacionados aos Sistemas Autônomos – Comunicação Cooperativa, Complexidade Sistêmica, AI/ML e Segurança Crítica – é ilustrada na Figura 17.

Figura 17 – Relação entre grupos de conceitos em sistemas autônomos

3.2. Garantia de segurança crítica em sistemas autônomos cooperativos

Viabilizar os futuros sistemas autônomos cooperativos, sobretudo os sistemas de transporte, exige compreender e lidar com a forma como a Comunicação Cooperativa, a AI/ML e a Complexidade Sistêmica **influenciam os riscos de segurança crítica de suas aplicações**. Assim, questões fundamentais relacionadas a como garantir a operação segura dos **sistemas autônomos cooperativos**, a forma como a comunicação cooperativa poderia contribuir com a segurança crítica destes sistemas, entre outras, devem ser respondidas no contexto da Engenharia de Sistemas Complexos (CoES).

Estas questões foram exploradas em dois projetos de pesquisa vinculados a um Convênio de Cooperação Técnico-Científico com a *Ericsson Research*[32]. Atuando na temática da segurança crítica e resiliência em sistemas de transporte autônomos, um dos principais resultados tangíveis destes projetos foi o desenvolvimento e implementação de um **ambiente**

[32] Entre outubro de 2015 e novembro de 2019, dois projetos de pesquisa consecutivos e evolutivos – "*Segurança crítica em sistemas ciberfísicos automotivos*" (2015-2017) e "*Segurança e resiliência em sistemas veiculares autônomos* (2017-2019) – lidaram com questões de segurança crítica e resiliência em ITS, desde a conceituação de sistema ciberfísico no domínio automotivo até a utilização de veículos autônomos conectados (CAV), suportados por tecnologias de comunicação cooperativa e de AI/ML. **O presente autor atuou diretamente ao longo de todo o período dos projetos como um dos principais pesquisadores – bem como especialista em engenharia de segurança de sistemas críticos em sistemas de transporte, contribuindo com o desenvolvimento intelectual das pesquisas e na autoria e coautoria de produção técnico-científica**.

computacional virtual utilizado na análise de segurança crítica dos futuros sistemas de tráfego autônomos (VISMARI *et al.*, 2018b). Este ambiente computacional permite modelar **cenários de tráfego rodoviário** no contexto dos Sistemas de Transporte Inteligente (ITS), bem como simular e analisar o impacto de conceitos, tecnologias e procedimentos sobre os níveis de segurança crítica do sistema de transporte.

Neste ambiente computacional, os cenários são modelados por meio dos elementos básicos de um sistema de tráfego (vias, ambiente, veículos e condutores), bem como por funcionalidades embarcadas nestes elementos – como sensores nos veículos, algoritmos de condução autônoma (baseados, ou não, em AI/ML) nos condutores – e pelas características da comunicação entre elementos. Ao simular estes cenários de tráfego, é possível identificar a dinâmica da relação entre os níveis de segurança crítica dos sistemas de transporte e o comportamento dos elementos modelados, representado por suas características individuais e pela interação entre eles – incluindo abordagens de comunicação cooperativa.

O ambiente computacional foi aplicado na avaliação de diversas abordagens e conceitos relacionados a cenários de tráfego baseados em ITS. Nestes cenários, a segurança crítica do tráfego foi avaliada quando da introdução de uma nova abordagem baseada em lógica *Fuzzy* para controle de frenagem de veículos autônomos e/ou altamente automatizados (NAUFAL *et al.*, 2018; VISMARI *et al.*, 2018a); de uma nova arquitetura de sistemas veiculares orientada à garantia da segurança (MOLINA *et al.*, 2017); de uma nova abordagem de controle baseada em sensores avançados para garantia da movimentação segura (MOLINA *et al.*, 2018); entre outros.

Nestas avaliações, a segurança crítica do tráfego rodoviário foi estimada utilizando métricas relacionadas ao **risco de colisão** entre veículos, principalmente a taxa de colisão (**CR – *Collision Rate***) entre veículos. Durante o processo de avaliação, buscou-se identificar o comportamento da CR (aumento, manutenção ou redução) em função das características configuradas no cenário de tráfego e desempenhadas pelo sistema em análise.

Portanto, ao utilizar o ambiente computacional de modelagem e simulação, considere o processo de avaliação dos cenários de tráfego representado na forma:

CRi = CR(Mi), onde:

Mi = **{CS; E}i:** modelo do cenário de tráfego em avaliação, submetido à configuração ‘i’ de seus elementos.

CS: modelo representativo do cenário/contexto da aplicação (ex.: arquitetura do cenário, características do tráfego, entre outros).

E = {E1, E2, ..., En**}**: modelagem dos elementos do sistema do cenário de tráfego em avaliação, formado por n elementos.

Es = {*param.1*; *param.2*, ..., *param.k*}s: variáveis/parâmetros considerados na modelagem do elemento ‘Es' do sistema de tráfego, onde s = 1, 2, ...n;

CR(.): função de transferência entre parâmetros do sistema e a métrica de segurança no nível de aplicação (neste caso, Taxa de Colisões - CR).

CRi: Taxa de Colisão [#colisões/tempo] observada para a configuração ‘i’.

Ao longo das avaliações dos **cenários de tráfego no contexto ITS**, foi possível observar que, para um mesmo sistema (M), configurações “i” significativamente diferentes entre si produziram valores semelhantes de CR. Como exemplo, considerando as configurações ‘k’ e ‘j’, observou-se que **CR(Mk) ≈ CR(Mj) = *CR*,** $k \neq j$

Ao analisar os pares {k, j} de configurações de cenário, observou-se que alguns parâmetros em nível de sistema atuavam de forma antagônica entre si com relação à métrica CR (em nível de aplicação). A título de ilustração, considere um par de parâmetros {param.x, param.y}s de um modelo em análise submetido às configurações {k, j}. Considerando:

Es,k = {*param.1*, ..., *param.x*, *param.y*, ...}s,k = {*V1_k*, ..., *Vx_k*, *Vy_k*, ...}s,k

Es,j = {*param.1*, ..., *param.x*, *param.y*, ...}s,j = {*V1_j*, ..., *Vx_j*, *Vy_j*, ...}s,j

Caso se observe CR(Mk) ≈ CR(Mj) tanto para o cenário Vx_k << Vx_j e Vy_k >> Vy_j quanto para o cenário Vx_k >> Vx_j e Vy_k << Vy_j, define-se que {param.x, param.y}s são parâmetros antagônicos entre si.

Comportamento análogo foi observado pelo autor em uma análise sistemática de risco de segurança do Sistemas de Controle de Tráfego Aéreo (ATC – *Air Traffic Control*) implementado segundo o paradigma CNS/ATM (FLAVIO VISMARI; CAMARGO JUNIOR,

2011; VISMARI, 2007). Ao modelar e avaliar este sistema, observou-se um comportamento antagônico entre parâmetros do sistema que influenciava o nível de risco de colisão entre aeronaves. Sobretudo, entre a taxa de varredura do sistema de vigilância e o nível de integridade de posição (acurácia) do sistema de navegação. Foram obtidos níveis semelhantes de risco tanto para situações de baixa integridade de posição e alta taxa de varredura de vigilância quanto na situação oposta (baixa integridade de posição e alta taxa de varredura).

Conforme apresentado anteriormente, tanto o CNS/ATM (no domínio dos transportes aéreos) quanto o ITS (no domínio dos transportes terrestres rodoviários) são análogos na forma de evolução dos paradigmas de sistemas baseados em ICT. Ambos os sistemas estão evoluindo orientados a modelos de referência de interconexão em camadas (como o ISO/OSI), onde os seus processos e serviços são implementados computacionalmente e de forma distribuída, e seus usuários/sistemas finais (EU/ES) são nós de uma rede de comunicação de dados que interagem entre si por meio de protocolos e serviços de interfaces comuns no domínio de aplicação.

Portanto, considerando o domínio dos **Sistemas Complexos de Engenharia** (definidos como sistemas com características de complexidade inerente) e, mais especificamente, os **sistemas de transporte autônomos cooperativos**, elaborou-se a hipótese de que **existe** um conjunto de **parâmetros antagônicos e não dependentes entre si** no nível dos elementos destes sistemas cujos valores podem ser manipulados (controlados) de forma a **obter níveis de risco de segurança** crítica aceitáveis no nível da aplicação.

Como forma de identificar um conjunto de parâmetros com as características indicadas, é necessário compreender o **comportamento individual** (i.e.: impacto positivo ou negativo) **de cada um dos parâmetros do sistema em relação à métrica de segurança da aplicação crítica**. Além disso, é necessário identificar as tuplas (duplas, por exemplo) de parâmetros, não dependentes entre si, cujo comportamento seja antagônico entre si com relação à métrica de segurança e, ao menos, um dos parâmetros seja manipulável (controlável). Por fim, deve-se inferir a relação (função) entre as tuplas de parâmetros do sistema e a métrica de segurança considerada. Desta forma, é possível estimar os valores a serem aplicados sobre os parâmetros controláveis da tupla como forma de compensar possíveis variações nos valores dos demais parâmetros, mantendo o nível de risco desejado para a aplicação.

A título de ilustração, a Figura 18 apresenta um exemplo no qual foi identificada uma dupla (**param.x**, **param.y**) de parâmetros antagônicos e não dependentes, adotada em uma **abordagem de compensação funcional**. A relação (param.x, param.y)➔CR está representada tanto por meio de uma tabela (Figura 18.(a)), classificando CR em 2 classes distintas – CR ≤ 0,5 (seguro) e CR > 0,5 (inseguro) – quanto por meio de uma superfície (Figura 18.(b)). Observa-se que, no instante t_1, os parâmetros (param.x, param.y) possuem os valores ($Vx(t_1)$, $Vy(t_1)$). Para o sistema em análise, estes valores produzem um nível de segurança considerada aceitável (CR ≈ 0,3).

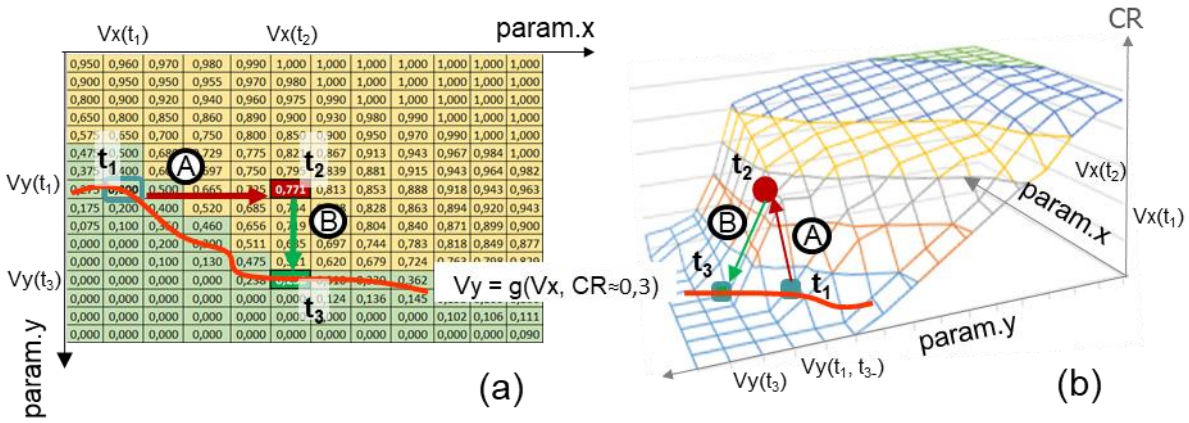


Figura 18 – Uso de compensação funcional na garantia da segurança crítica

No instante $t_2 > t_1$, observa-se uma elevação no valor de param.x ($Vx(t_2) > Vx(t_1)$), expondo a aplicação a um nível de risco não aceitável ($Vx(t_2)$, $Vy(t_2)$) ➔ CR >> 0,5 (transição 'A'). Então, o valor de param.y é ajustado no instante $t_3 > t_2$ (transição 'B'), compensando a elevação no valor do param.x de forma a manter CR em um nível considerado seguro. Desta forma, um novo valor é estimado para o param.y pela função Vy = g(Vx, CR < 0,3) e, quando o parâmetro controlado é reconfigurado com $Vy(t_3)$, a aplicação é considerada segura.

3.3. Compensação funcional e a garantia de segurança crítica em CoES

Com base nas observações apresentadas no contexto de sistemas de transporte autônomos cooperativos, infere-se ser possível gerenciar o comportamento **antagônico e não dependente entre si** entre características de elementos do sistema de forma a regular **o nível de risco de segurança** crítica da aplicação. Portanto, uma **abordagem de compensação funcional** utilizando estas características tem potencial de aplicação em **processos de garantia da segurança crítica** em CoES – em especial, sistemas de transporte autônomo cooperativos.

Portanto, uma abordagem baseada em compensação funcional deve ser capaz de realizar as seguintes atividades**:**

1. Inferir o comportamento individual dos parâmetros dos elementos do sistema em relação à métrica de segurança da aplicação crítica (CRn = CR(param.n));

2. Identificar as tuplas de parâmetros, não dependentes entre si, cujos comportamentos sejam antagônicos com relação à métrica de segurança e, ao menos, um dos parâmetros seja manipulável (controlável). Por exemplo, identifica-se a dupla {param.x, param.y}, pois (d(CR(param.x))/d(param.x) / d(CR(param.y))/(dparam.y)) < 0, onde o param.y é controlável.

3. Definir a relação (função) entre as tuplas de parâmetros do sistema identificadas e a métrica de segurança considerada. No exemplo, para a dupla {param.x, param.y}, são obtidos **{(param.x, param.y) ➔ CR}** e param.y = g(param.x, CR).

4. Compensar possíveis variações nos valores dos parâmetros, estimando os valores a serem aplicados sobre os parâmetros controláveis da tupla como forma de manter o nível de risco de segurança da aplicação em níveis considerados aceitáveis. Esta atividade pode ser aplicada, ao menos, em dois cenários distintos, relacionados ao monitoramento da métrica de segurança da aplicação crítica, nas quais:

    4.1 A Abordagem Proposta **não monitora** a métrica de segurança da aplicação crítica **em tempo de execução**.

    Neste caso, a relação **{(param.x, param.y) ➔ CR}** é obtida somente durante a fase de implementação da Abordagem Proposta. Assim, apenas os parâmetros da tupla identificada (param.x, param.y) são monitorados, sendo obtidos os valores (Vx(t), Vy(t)) em tempo de execução. Portanto, ao receber um valor ($Vx(t_i)$, $Vy(t_i)$) no qual **{**($Vx(t_i)$, $Vy(t_i)$) **➔ CR ≥ SL}**, onde SL é o valor limite de CR no qual sistema é considerado seguro (risco em nível aceitável), a Abordagem proposta estima $Vy(t_{i+1}) = g(Vx(t_i), CR < SL)$.

    Portanto, a Abordagem Proposta tem aplicação no processo de garantia de segurança do sistema em tempo de execução, apenas.

4.2 A Abordagem Proposta **monitora** a métrica de segurança da aplicação crítica **em tempo de execução**.

Neste caso, é possível aplicar $Vy(t_{i+1}) = g(Vx(t_i), CR(t_i))$, onde $CR(t_i)$ é a métrica de segurança obtida no instante t_i.

Portanto, a Abordagem Proposta, em tempo de execução, pode ser aplicada tanto no processo de garantia de segurança do sistema quanto em um processo de gerenciamento e otimização de recursos do sistema orientado aos riscos da aplicação.

Observa-se que capacidade de **monitoramento** do **comportamento** tanto dos elementos de sistema quanto do nível de risco (segurança crítica) da aplicação é característica fundamental à implementação da **abordagem baseada em compensação funcional**. Contudo, considerando as características de **Sistemas Complexos de Engenharia** justificadas anteriormente, os Sistemas Autônomo Cooperativos, por princípio, não possuem uma **autoridade ou controle centralizado**. Consequentemente, os elementos do sistema possuem capacidade limitada de obter consciência situacional do seu ambiente, podendo gerar restrições à aplicação desta **abordagem**.

Contudo, vale frisar que **sistemas críticos em segurança** baseados em ICT (Tecnologias de Informação e Comunicação) – entre eles, **os sistemas de transporte autônomo** – são inerentemente **sistemas ciberfísicos distribuídos**. Nestes sistemas, os elementos ciberfísicos apresentam um 'lado físico' – onde são realizados os processos que envolvem a transformação/transferência de energia no nível de aplicação crítica, onde danos podem ser produzidos – e o 'lado lógico' – onde se realiza **a interface com os demais elementos do ambiente**, o processamento de dados e o gerenciamento dos processos do lado físico.

Desta forma, entende-se que os dados trafegados por estas interfaces possam ser utilizados para inferir tanto o comportamento dos elementos do sistema críticos quanto do nível de segurança crítica da aplicação. Portanto, o monitoramento da comunicação entre os elementos ciberfísicos poderia ser uma solução para as restrições apresentadas à implementação da abordagem proposta.

Para fins de argumentação, considere a arquitetura de sistema distribuído baseado em ICT e crítico em segurança ilustrada na Figura 19.(a). Nesta arquitetura, tanto os elementos **ciberfísicos** (e_i, e_j e e_k) – cujos 'lados físicos' estão relacionados aos Processos Críticos em Segurança (PCS) da aplicação – quanto os elementos **lógicos** (e_h e e_m) – que lidam apenas com processamento de dados no domínio da informação – estão distribuídos e se comunicam entre si por meio de suas interfaces de comunicação de dados (setas contínuas). Consequentemente, e conforme ilustrado pela Figura 19**.(b)**, estas interfaces podem ser **agrupadas e abstraídas como sendo um dos elementos do sistema** (**C_i**). No exemplo, a arquitetura do elemento **C_i** é formada pelos componentes (subsistemas) **CSPa** e **CSPb** (duas redes de dados distintas e que possuem interface entre si).

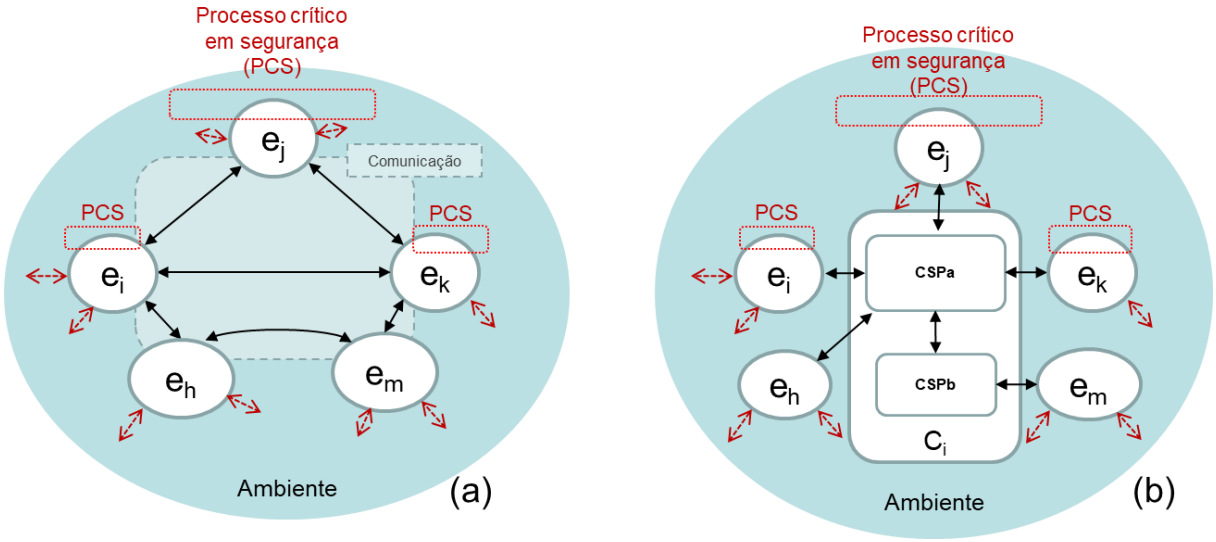


Figura 19 – Arquitetura de sistema crítico (a.) e a comunicação (C_i) como elemento do sistema (b)

Portanto, conclui-se que o monitoramento da comunicação entre os elementos do sistema deva ser realizado sobre o **elemento de comunicação (C_i)** do sistema em questão. Além de possibilitar o monitoramento dos dados trafegados entre os elementos ciberfísicos (e não ciberfísicos) – utilizados para compreender o comportamento da relação entre os parâmetros dos elementos do sistema e do nível de segurança crítica da aplicação – o monitoramento do elemento de comunicação (**C_i**) permitiria incluir e, se possível, considerar **os parâmetros de C_i no procedimento de compensação funcional**.

Desta forma, a Figura 20 apresenta o fluxo de dados esperado entre um **elemento que implemente a abordagem de compensação funcional proposta neste trabalho** (denominado como "**Abordagem Proposta**") e os elementos do sistema distribuído crítico em segurança

baseado em ICT. Observa-se que, por meio do elemento de comunicação do sistema (**C_i**), a Abordagem Proposta monitora parâmetros dos elementos do sistema, incluindo parâmetros de comunicação (indicado pelos fluxos ‘monitora(“elemento”)’). Também por meio de **C_i**, a Abordagem Proposta ajusta os parâmetros de elementos como forma de implementar a compensação funcional (indicado pelos fluxos ‘ajusta(“parâmetro”, “elemento”)’). O desenvolvimento do conceito da “Abordagem Proposta” é detalhado a seguir (sec.3.4).

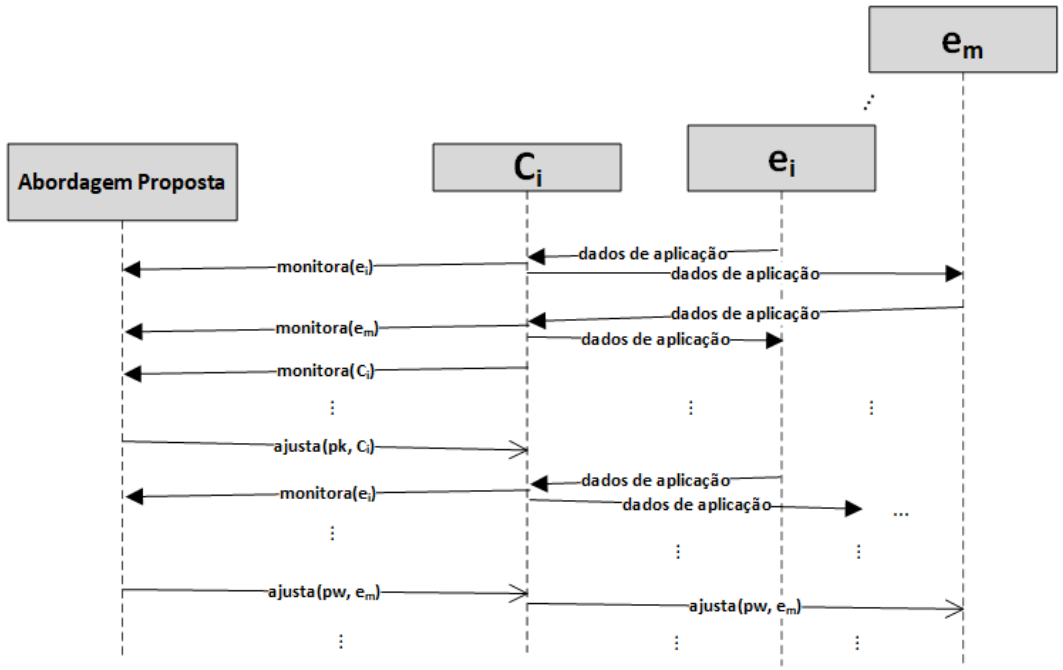


Figura 20 – Fluxo de dados entre componentes do sistema (incluindo Ci) e Abordagem Proposta

A efetividade da Abordagem Proposta é influenciada tanto pelas características do conjunto de dados trafegados pelo elemento de comunicação (**C_i**) que podem ser monitorados pela Abordagem Proposta quanto pelos parâmetros que podem ser controlados. Consequentemente, o **processo de implementação** (e, de forma mais ampla, o **ciclo de vida**) da Abordagem Proposta é orientado à contexto, dependendo das características do elemento de comunicação utilizado – como arquiteturas, protocolos e funcionalidades – no contexto de sistema crítico baseado em ICT.

3.4. A Abordagem Proposta no contexto dos C-ITS

O Veículo Autônomo Conectado (CAV – *Connected Autonomous Vehicle*) é o principal elemento ciberfísico e crítico em segurança dos **sistemas de transporte autônomos no contexto de C-ITS**. Um modelo ciberfísico de referência para CAV, adaptado de

NASCIMENTO *et al.* (2019)[33], é ilustrado na Figura 21. Neste modelo, o '**Veículo**'[34] é o 'lado físico' do CAV que interage com o '**Ambiente**'[35] e onde são realizados os processos críticos relacionados ao movimento do veículo, gerenciados pelo seu 'lado ciber' (percepção, controle e atuação). Além disso, este modelo adaptado permite representar as interfaces entre o CAV e os demais elementos do ambiente C-ITS (V2X).

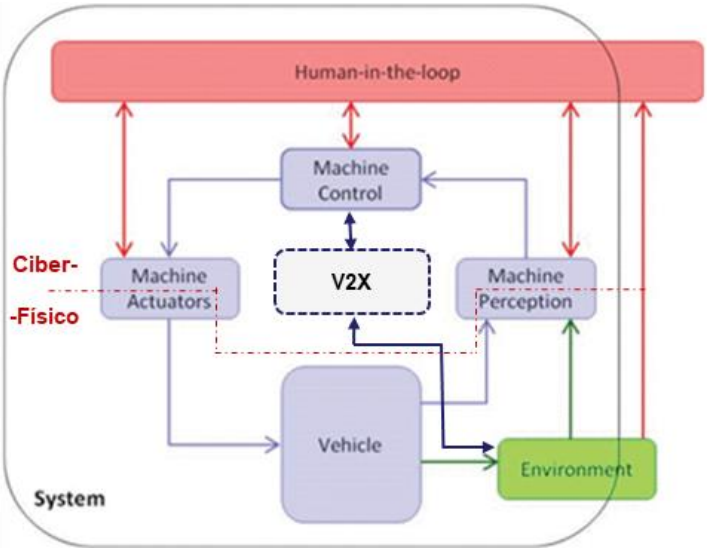


Figura 21 – Modelo de referência para CAV
Fonte: adaptado de NASCIMENTO *et al.* (2019)

Um **Sistema de Transporte Autônomo no contexto de C-ITS** é instanciado ao se utilizar o modelo de referência para CAV como modelo para os elementos ciberfísicos (e_i, e_j, e_k) da arquitetura apresentada na Figura 19. Além disso, o elemento CSPa representa a infraestrutura de rede de dados que provê a **comunicação V2X** (móvel) entre os nós desta rede (neste exemplo, e_i, e_j e e_k – CAV – e $\mathbf{e_h}$). Assim, o CSPa implementa a comunicação entre os CAV (**V2V**); entre veículos conectados e outros elementos móveis do sistema (**V2P**), como pedestres,

[33] O modelo proposto em pelos autores foi adaptado de CHRISTENSEN *et al.* (2015).O modelo original foi desenvolvido em um projeto de pesquisa patrocinado pelo NTHSA – *National Highway Traffic Safety Administration* (subordinado ao Departamento de Transportes dos EUA) – e conduzido pelo consórcio AVR (Automated Vehicle Research), formado por Ford, General Motors, Nissan, Mercedes-Benz, Toyota e Volkswagen/Audi. Seu objetivo principal era orientar o desenvolvimento da automação de veículos (níveis SAE L0 a L5, segundo a SAE (2014). Foi adaptado para que a literatura a respeito de segurança crítica e AI/ML em AV pudesse ser mapeada sobre cada elemento de um veículo autônomo.

[34] **Veículo** é o conjunto de elementos fundamentais – como chassi, motor, transmissão, suspensão, entre outros – que interage diretamente com o 'Ambiente' e que realiza os processos físicos de transformação e transferência de energia/grandezas relacionados ao cumprimento da sua missão (ou seja, movimento controlado).

[35] **Ambiente** representa todos os demais elementos de um sistema de tráfego – como vias, clima, sinalização viária, outros veículos, entre outros – com os quais o veículo (CAV) interage.

ciclistas e demais usuários da via vulneráveis (VRU – *Vulnerable Roadway Users*); entre CAV e infraestrutura de ITS[36] (**V2I**); e entre CAV e servidores de aplicação (elemento e_m) – **V2N** – disponíveis em redes de dados externas à V2X (CSPb).

Vale frisar que a diferença entre as comunicações V2I e V2N, em termos funcionais, é o tipo da interface de comunicação entre CAV e os demais usuários finais (EU – *End Users*), como os servidores de aplicação do sistema (ETSI, 2015). Na comunicação V2I, a interface é uma RSU[37] (*Road Side Unit*). Já na comunicação V2N, a interface pode ser qualquer ponto de acesso sem fio à rede que forneça serviços de aplicação.

Contudo, uma RSU pode ser implementada tanto por e**lementos dedicados ao ITS**, instalados na via, quanto por **elementos de comunicação de uso geral**, como um eNB[38] de uma rede móvel celular (ETSI, 2015). Consequentemente, os CAV podem acessar os serviços de aplicação externos à rede V2X (V2N) e a infraestrutura e serviços de ITS (V2I) por meio de uma **infraestrutura de rede móvel celular de propósito geral.** Porém, o **desempenho que pode ser obtido dos serviços de aplicação** por meio destes dois tipos de comunicação não **é o mesmo**, dado que quanto maior a distância de rede entre CAV e servidor de aplicação menor o nível de desempenho relacionado à comunicação.

Desta forma, a comunicação V2I poderia ser melhor empregada para serviços de aplicação críticos em desempenho, e implementados por meio de Computação de Borda (*Edge Computing*) – representado na Figura 19 por **e_h**. Já a comunicação V2N poderia ser utilizada em serviços de aplicação menos críticos em desempenho, e implementados por meio do conceito de Computação em Nuvem (*Cloud Computing*) – representado na Figura 19 por **e_m.**

Portanto, considerando estas características de comunicação de sistemas de transporte autônomo no contexto C-ITS, conclui-se que **a Abordagem Proposta (conceitual)** seria melhor implementada **(ferramenta)** como um serviço de aplicação em um servidor de computação de borda da infraestrutura de rede móvel (elemento e_h, Figura 19**.(b))** que provê os **serviços de comunicação V2X (componente CSPa,** Figura 19**.(b))**. A Figura 22 ilustra a

[36] Esta infraestrutura ITS pode ser representada por semáforos inteligentes, sensores e câmeras de tráfego, serviços de aplicação ITS (como gerenciamento de tráfego, emergência), entre outros.
[37] **RSU** é uma entidade que suporta serviços V2I e que pode se comunicar com um usuário final (EU).
[38] Um eNB, também denominado como *Evolved Node* B, é um equipamento que é conectado a uma rede móvel celular e que se comunica diretamente com os usuários móveis da rede.

arquitetura desta implementação, apresentando os fluxos de dados entre seus elementos, conforme ilustrado pela Figura 20.

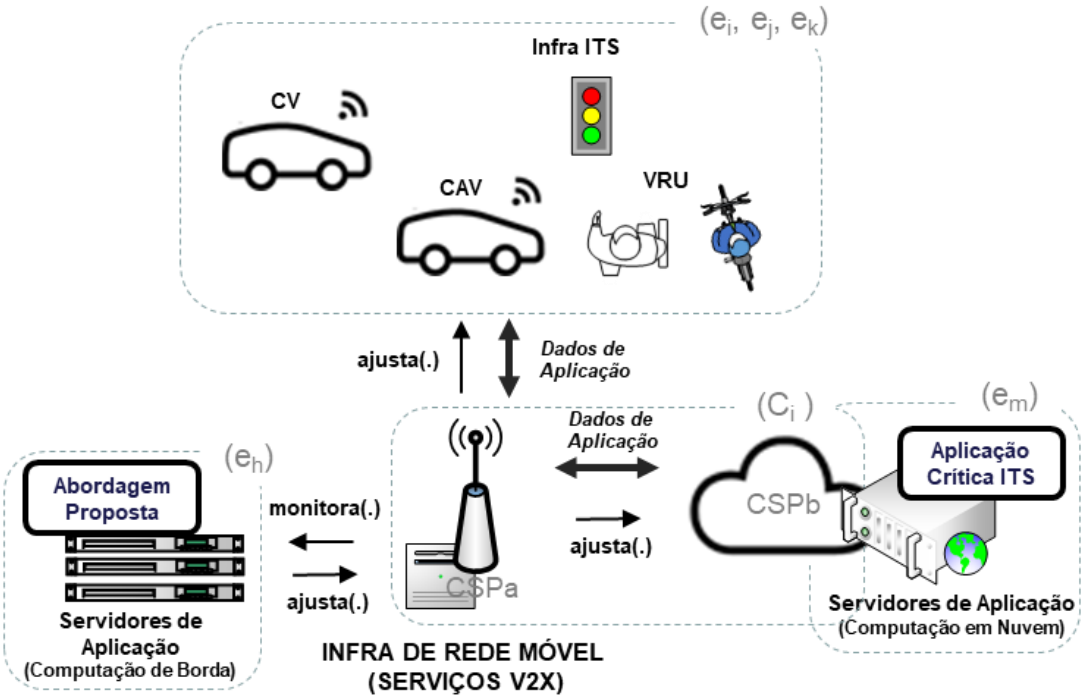


Figura 22 – Arquitetura de implementação da Abordagem Proposta

Com esta implementação, seria possível monitorar os dados relacionados à consciência situacional dos CAV por meio de mensagens BSM ou CAM que trafegam pela **infraestrutura de rede** dos serviços de comunicação V2X (*Dados de Aplicação*), permitindo inferir o nível de segurança crítica do processo de tráfego realizado pelos CAV. Também seria possível **monitorar e ajustar** parâmetros de rede do provedor de serviços de comunicação V2X (CSPa), bem como **monitorar e ajustar** parâmetros dos elementos que instanciam serviços de aplicação crítica ITS (**e_m**).

Nas próximas seções, um estudo de caso é apresentado em detalhes, onde um cenário-base de tráfego rodoviário em C-ITS é desenvolvido e implementado para apoiar o **desenvolvimento, análise e demonstração** do conceito da Abordagem Proposta nesta tese. Tanto a modelagem do **cenário-base** quanto da **abordagem proposta** (como ferramenta) são implementadas por meio do ambiente computacional anteriormente apresentado.

4. ESTUDO DE CASO NO CONTEXTO C-ITS

"Experience serves not only to confirm theory, but differs from it without disturbing it; it leads to new truths which theory only has not been able to reach."
(D'Alembert, Elémens de Philosophie[39]*.* tradução nossa*)*

Esta seção apresenta, em detalhes, o estudo de caso projetado e executado para apoiar o desenvolvimento, análise e demonstração da Abordagem Proposta neste trabalho.

Primeiro, elabora-se um cenário-base de tráfego rodoviário aderente ao contexto de Sistemas de Transportes Inteligentes Cooperativos (C-ITS). Este cenário-base tem as características demandadas pela pesquisa: um sistema distribuído e baseado em tecnologias de informação e comunicação – no caso, tecnologias ITS para comunicação V2X – em uma aplicação crítica em segurança – o ambiente de cruzamento de tráfego rodoviário sem semáforo. Por fim, este cenário é implementado, executado e avaliado por meio de modelagem e simulação computacional, utilizando o ambiente computacional virtual apresentado anteriormente.

Em seguida, uma instanciação do conceito da abordagem proposta é desenvolvido e implementado sobre o cenário-base. Assim, as características do cenário-base orientam as definições da arquitetura e do processo de desenvolvimento e operação da Abordagem Proposta. Esta arquitetura é implementada no modelo computacional do cenário-base e, por meio de simulação computacional, obtêm-se os dados para a avaliação da efetividade da abordagem proposta, como ferramenta, no contexto de C-ITS.

Portanto, este estudo de caso tem como objetivos: testar a hipótese de que é possível garantir propriedades emergentes no nível da aplicação – como segurança crítica de tráfego – por meio da compensação funcional entre parâmetros dos elementos no nível de sistema; e avaliar a efetividade da abordagem proposta na garantia da segurança (e na otimização de recursos) da aplicação crítica modelada no estudo de caso.

[39] Citado na *introdução de "Traite Analythique de la Resistance des Solides", de P.S Girard* (1798). Disponível em https://gallica.bnf.fr/ark:/12148/bpt6k1517747n.image

4.1. O cenário-base de tráfego rodoviário no contexto C-ITS

No cenário-base elaborado para este estudo de caso, um sistema de supervisão e controle suportado por ICT, supervisiona e controla uma aplicação crítica – um cruzamento de vias de tráfego rodoviário. A missão deste sistema crítico em segurança é prover o serviço de "Semáforo Virtual" aos veículos autônomos conectados (CAV) que trafegam no cruzamento de vias.

Neste sistema, a comunicação V2X entre os **elementos** – os veículos autônomos conectados (CAV) que trafegam nas vias do cruzamento de tráfego e os servidores de aplicação no Centro de Controle Operacional (CCO) que fornecem o serviço de semáforo virtual – é realizada por um **Provedor de Serviços de Comunicação (CSP) de propósito geral** baseado em tecnologias de comunicação sem fio, como 4G/LTE e 5G. Os parâmetros, configurações, métricas e resultados obtidos com este cenário-base são definidos, bem como os planos de teste a serem executados e que orientam a obtenção de resultados para análise.

O cenário-base é implementado e executado por meio do **ambiente computacional virtual** apresentado anteriormente. Neste ambiente computacional (VISMARI *et al.*, 2018b), é possível modelar sistemas de tráfego – considerando veículos, vias, condutores, comunicação e demais elementos – e simular situações definidas por seus parâmetros operacionais. Desta forma, o cenário-base é modelado computacionalmente e simulado segundos os parâmetros definidos nos Planos de Teste. Maiores detalhes a respeito deste ambiente são apresentados na seção 4.1.2.

4.1.1. Definição do cenário-base de tráfego rodoviário

Neste estudo de caso, um **serviço** de **Controle Automatizado de Tráfego de Cruzamento (CaTraCa)** gerencia a dinâmica dos Veículos Autônomos (CAV) trafegando por vias convergentes e cujo cruzamento não possui semáforos físico. Neste contexto de C-ITS, CAVs se comunicam com o provedor do serviço de semáforo virtual – neste caso, servidores computacionais no Centro de Controle Operacional (**CCO**) que executam os algoritmos do **CaTraCa** – por meio da **Infraestrutura de um Provedor de Serviços de Comunicação** (**CSP** – *Communication Service Provider*). Este CSP, de propósito geral, provê os serviços de comunicação V2X (*Vehicle to Everything*) entre CAV e CCO. A Figura 23 ilustra o

relacionamento entre CCO, CSP e os CAV, bem como a relação entre o Ambiente de Tráfego rodoviário e os CAV.

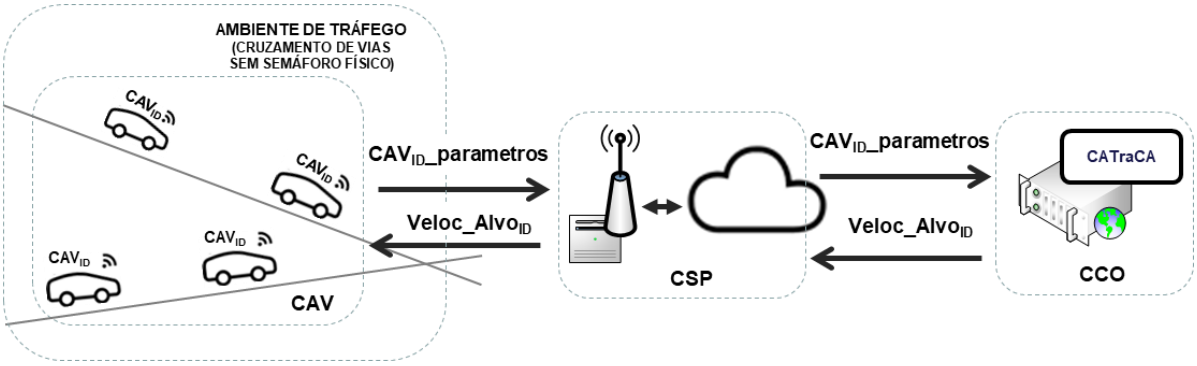


Figura 23 – Relacionamento entre elementos do cenário-base (CAV, CSP e CCO)

A Figura 24 ilustra o cenário-base, em especial o ambiente de tráfego rodoviário, adotado neste estudo de caso. Este ambiente de tráfego possui duas vias (**Via 1** e **Via 2**) simples (não permitem ultrapassagem) e de mão única, ortogonais entre si. Estas duas vias se cruzam na interseção entre vias (Região de Cruzamento – **RCz**). Esta interseção não possui sinais de trânsito ou semáforos físicos gerenciando o fluxo de tráfego no cruzamento. Assim, para garantir a segurança e fluidez do tráfego dos CAV que utilizam estas vias, o serviço de Controle Automatizado de Tráfego de Cruzamento (CaTraCa) é fornecido pelo **CCO** para os CAV dentro de uma **Região de Controle (RCt)**.

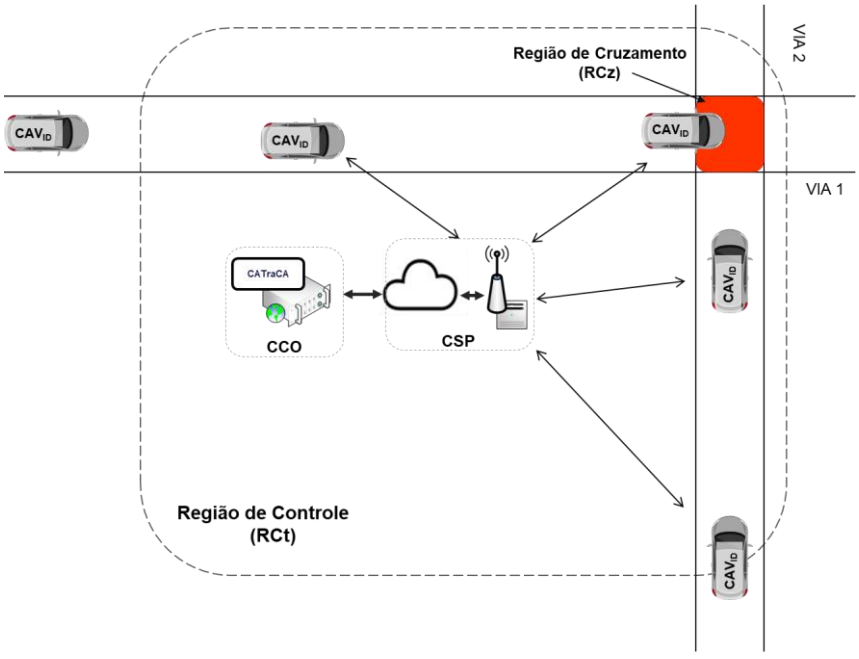


Figura 24 – Ambiente de tráfego do cenário-base

Portanto, o **CaTraCA** fornece um serviço de **semáforo virtual**, coordenando o fluxo de tráfego em ambas as vias e prevenindo acidentes (colisões entre CAV) na Região de Cruzamento (**RCz**). Esse conceito foi inspirado no trabalho de MICULESCU e KARAMAN (2016, 2020), que propuseram um algoritmo de coordenação de tráfego para o mesmo cenário de cruzamento sem semáforo adotado neste trabalho. Contudo, o algoritmo não permitia introduzir variações no desempenho dos elementos do sistema. Desta forma, foi necessário desenvolver um novo algoritmo, aplicável aos objetivos do presente trabalho.

O princípio de funcionamento do CaTraCA, desenvolvido ao longo deste trabalho, está baseado no gerenciamento da ocupação temporal da região de cruzamento (Rcz) por cada CAV. Continuamente, o CCO recebe mensagens contendo **parâmetros** (**CAV_{ID}_parameters**) de **todos os CAVs** univocamente identificados (ID) dentro da Região de Controle (**RCt**). As mensagens **CAV_{ID}_parameters** contêm os parâmetros de consciência situacional definidos pelo protocolo BSM (*Basic Safety Messages*) – *timestamp*, posição, velocidade, direção, comprimento do veículo e aceleração (SAE, 2022)**.**

Ao receber os dados, o **CaTraCa** – embarcado no CCO – estima o intervalo de tempo – instante inicial (Ti) e instante final (Tf) – no qual cada CAV ocupará a **RCzs**. Quando o CaTraCa identifica um **conflito futuro** – onde dois CAV ocuparão a **RCz** no mesmo intervalo de tempo, o CaTraCa determina as velocidades com as quais estes CAV deveriam trafegar para que este conflito não ocorra no futuro. Estas novas '**velocidades-alvo**' são enviadas aos CAV por meio da comunicação V2X provida pelo CSP. Ao receber uma mensagem-alvo ($Veloc\_Alvo_{ID}$), o $CAVI_D$ endereçado deve aplicá-la, aumentando ou reduzindo sua velocidade atual. A Figura 25 ilustra o diagrama de relacionamento entrada-saída do CaTraCA.

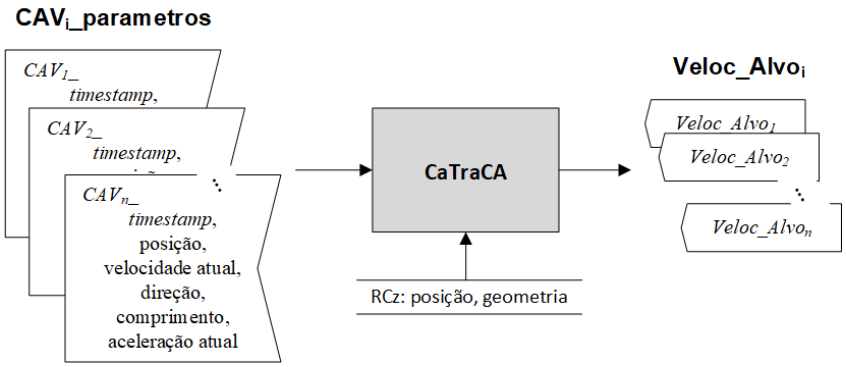


Figura 25 – Controle Automatizado de Tráfego de Cruzamento (CaTraCa)

A Figura 26 exemplifica o conceito do CATraCA instanciado neste estudo de caso. Cada via (Via 1 e Via 2) é representada por uma linha temporal, onde o AV_1 (CAV identificado com o ID 1)[40] está trafegando na via 1 e o AV_2 (CAV identificado com o ID 2) está trafegando na via 2. Os instantes futuros de entrada (Ti) e de saída (Tf) de cada AV na RCz são continuamente estimados. Neste exemplo, o CaTraCa identificou que a RCz será ocupada pelo AV_1 entre os instantes Ti_1 e Tf_1, e o AV_2 ocupará a RCz entre os instantes Ti_2 e Tf_2.

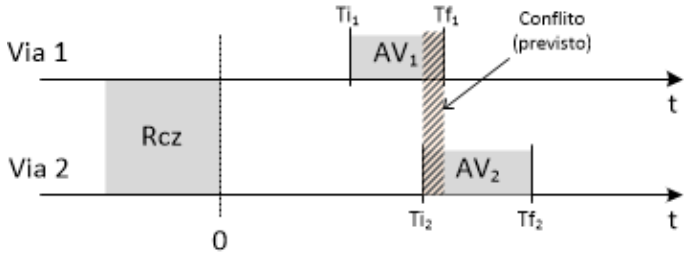


Figura 26 – Gerenciamento da ocupação temporal da região de cruzamento (AV1 e AV2)

Dado que há interseção entre estes intervalos de tempo – ou seja, AV_2 ingressará na RCz antes do AV_1 ter desocupado completamente a RCz, considera-se que haverá um conflito futuro (colisão) entre AV_1 e AV_2 na RCz. Neste caso, o CaTraCa determina quais velocidades deveriam estar trafegando o AV_1 e/ou AV_2 para que não haja o conflito. Essas velocidades-alvo são informadas aos AV_1 e/ou AV_2, que são responsáveis por aplica-las, aumentando ou reduzindo suas velocidades.

A Figura 27 ilustra o diagrama de tempo e o fluxo das informações trocadas entre os elementos do sistema (AV, CSP e CCO) deste cenário-base. Neste estudo de caso, adota-se um protocolo de comunicação V2X por radiodifusão (*broadcast*), periódico e não orientado à conexão (*connectionless*), onde as mensagens dos AVs e CCO são transmitidas periodicamente e a uma taxa de transmissão de mensagens (*msg_update_rate*) configurável no CSP para o provimento dos serviços V2X.

Garantir a qualidade de serviço (QoS) fornecido pelo(s) Provedor(es) de Serviços de Comunicação (CSP) é fundamental para o desempenho das **aplicações críticas em segurança suportadas por sistemas de comunicação**. No estado-da-arte normativo de sistemas críticos em segurança suportados por comunicação, a comunicação não é considerada um elemento

[40] **IMPORTANTE**: a partir deste ponto do texto, e apenas com relação ao estudo de caso apresentado neste trabalho, **os termos 'AV' (*Autonomous Vehicles*) e 'CAV' (Connected *Autonomous Vehicles*) devem ser compreendidos como sinônimos**.

crítico em segurança. Desta forma, a segurança crítica da aplicação deve ser garantida pelos ‘usuários finais’ (EU – *End Users*). Portanto, os EUs são responsáveis tanto por **detectarem** problemas na comunicação – como perda de **integridade** dos dados, perda da **disponibilidade** da comunicação (permanente ou transitória), alta latência, entre outros – quanto por **reagir** de forma segura.

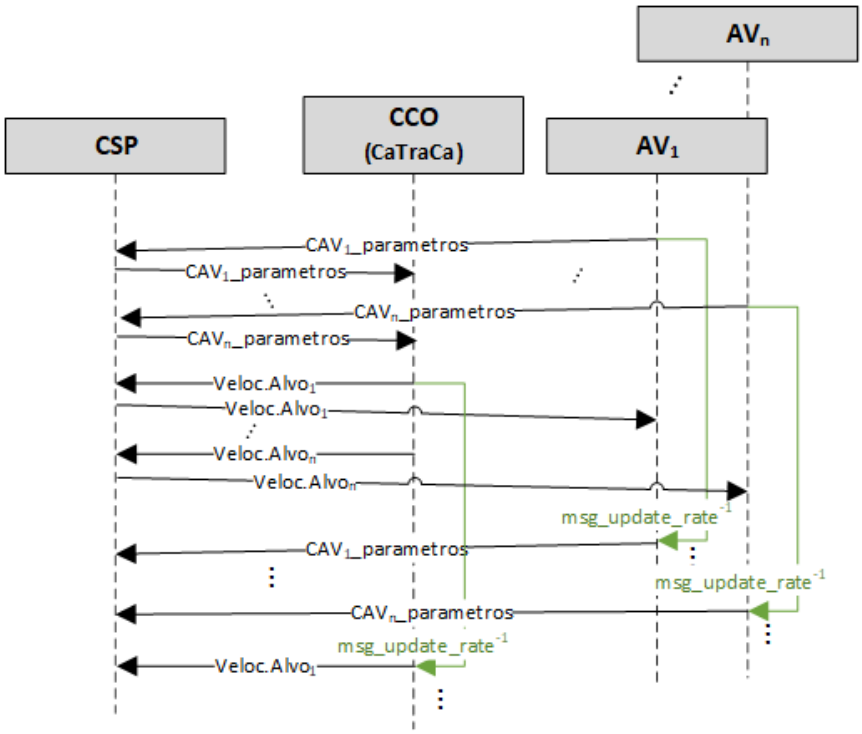


Figura 27 – Fluxo temporal de dados trocados entre elementos do cenário-base

No presente estudo de caso, tanto as mensagens de consciência situacional da região de controle (RCt) enviadas pelos CAV e recebidas pelo CCO (CaTraCa) quanto as velocidades-alvo enviadas pelo CCO e recebidas pelos AVs são **transportadas pelo CSP**. Portanto, CCO (**CaTraCa**) e CAV são usuários finais (EU – *End Users*) deste sistema crítico e devem ser capazes de lidar com problemas de desempenho nos serviços de comunicação fornecidos pelo CSP. Sobretudo os CAV, elementos ciberfísicos do sistema que expõe usuários e ambiente à perdas e danos.

Comumente, os CSP empregados em sistemas baseadas em comunicação são de **propósito geral**. Ou seja, são desenvolvidos e operados por terceiros com o objetivo de transportar dados entre usuários/elementos de um sistema distribuído, independentemente do tipo de aplicação que utilizará seu serviço de comunicação. Portanto, ao contratar os serviços

de comunicação de um CSP, o contratante (sistema/aplicação crítica) e a contratada (CSP) estabelecem um Acordo de Nível de Serviço (SLA – *Service Level Agreement*).

Em um SLA, o **contratante** (neste caso, a aplicação crítica) estabelece os **requisitos mínimos de desempenho de comunicação** que o CSP deve garantir para que sua aplicação opere conforme esperado. Entre estes requisitos, definem-se a latência máxima fim-a-fim, a taxa mínima de transmissão de mensagens, a taxa de perda de pacotes, a taxa de erro de bits (BER – Bit Error Rate), a disponibilidade mínima, entre outros.

No presente estudo de caso, adota-se que a comunicação V2X provida pelo CSP é baseada em **tecnologia móvel celular (LTE/4G ou 5G)**. Assim, os requisitos de serviços de comunicação C-V2X definidos pelas especificações técnicas publicadas pelo ETSI (*European Telecommunications Standards Institute*)[41] são adotadas neste estudo de caso, onde:

- A ETSI TS 22.185 (ETSI, 2017) define os requisitos para serviços de comunicação V2X baseados em 4G-LTE. Nesta especificação, estes requisitos são definidos em função da classe de aplicação de serviços ITS – **segurança viária**, eficiência de tráfego e outras aplicações (ETSI, 2009) – e, consequentemente, o tipo de comunicação V2X que as suportam (V2V, V2P, V2I e V2N). Ao se avaliar os casos de uso utilizados para elaboração da TS 22.185 (ETSI, 2015), os **serviços para segurança viária baseados em V2I e V2N[42]** são aderentes ao presente estudo de caso. Nestes serviços V2X, são considerados cenários onde os nós da infraestrutura de rede (como RSU) e servidores de segurança de tráfego (como o CCO) geram e distribuem mensagens relacionadas à segurança de tráfego para os Usuários Finais (como os CAV).

- Na ETSI TS 22.186 (ETSI, 2022), os requisitos para serviços de comunicação V2X – baseados em tecnologia 5G – estão divididos em ‘cenários relacionados à segurança crítica’ (como a Direção Avançada ou Remota e o *Platooning* de Veículos) e ‘cenários não relacionados à segurança crítica’ (como Entretenimento e Atualização Dinâmica de Mapas). Nos cenários relacionados à segurança crítica, os serviços V2X habilitam a condução semi-automatizada ou completamente automatizada (autônomas),

[41] https://www.etsi.org/

[42] A diferença entre serviços V2I e V2N é o tipo de nó de rede intermediário. No V2I, a comunicação ocorre entre AV e RSU. No V2N, os AVs se comunicam com servidores de aplicação, comunicando-se pela rede 4G-LTE. Contudo, um RSU pode ser implementado tanto por um elemento estacionário (dedicado ao ITS, instalado na via) quanto por um eNB (ETSI, 2015).

compartilhando dados entre si e permitindo a coordenação de trajetórias, permitindo **prevenindo colisões** e promovendo maior segurança para os usuários e maior eficiência para o tráfego. Desta forma, o cenário adotado neste estudo de caso (serviço CaTraCa) pode ser classificado como um cenário relacionado à segurança crítica.

Requisitos gerais e específicos de desempenho da comunicação são definidos tanto na ETSI TS 22.185 – 4G/LTE quanto na ETSI TS 22.186 – 5G. A diferença entre ambas as especificações está na forma como os requisitos específicos são apresentados. Na ETSI TS 22.185, os requisitos específicos estão agrupados em requisitos de desempenho da comunicação: Latência/Confiabilidade (ms), Tamanho da Mensagem (bytes), Frequência (mensagens/segundo), Alcance (metros) e Velocidade dos AVs (km/h). Dentro de cada grupo, os requisitos são definidos por **tipo de comunicação V2X** utilizada (V2V, V2P, V2I e V2N). Já na ETSI TS 22.186, os requisitos específicos – *Carga Paga* (bytes), *Taxa de Transmissão* (mensagens/s), Máxima Latência Fim-a-Fim (ms), Confiabilidade (%), Taxa de Dados (Mbps) e Alcance Mínimo de Comunicação (m) – estão agrupados por **tipo de cenário suportado pela comunicação V2X** e pelo nível de automação dos veículos.

Para o cenário adotado neste estudo de caso (serviço **CaTraCa**), observa-se nestas especificações técnicas que, para a comunicação V2X baseada em **LTE/4G**, a **latência máxima** deve ser de **100ms** (comunicação V2I entre AV e RSU) e **1.000ms** (comunicação V2N entre AV e Servidor de Aplicação), e a **frequência mínima** deve ser de **10Hz** por AV. Para a comunicação V2X baseada em **5G**, a **Máxima Latência Fim-a-Fim** deve estar entre **3ms** (Direção Avançada – alinhamento de trajetórias de emergência) ou **10ms** (maior parte dos cenários críticos em segurança) e **100ms** (Direção Avançada – compartilhamento de informação para veículos com alto nível de automação); e a **Taxa de Transmissão mínima** deve estar entre **10 Hz** (maior parte dos cenários) e **100Hz** (Direção Avançada – prevenção de colisão entre AVs).

Portanto, estas especificações técnicas podem ser utilizadas como base de definição dos requisitos de comunicação mínimos demandados pelos usuários (aplicação crítica) – baseados na ETSI 22.185 (ETSI, 2022) – e estabelecidos no SLA. Consequentemente, envelopes operacionais relacionados aos requisitos de comunicação são estabelecidos para a aplicação crítica (usuários). Estes envelopes operacionais representam as condições nas quais o serviço de comunicação está disponível à aplicação crítica. Caso os parâmetros de comunicação não

estejam aderentes ao envelope de operação (SLA), a comunicação ficará indisponível e, neste momento, os elementos do sistema/aplicação crítica (EU) são responsáveis por garantir a segurança crítica (por exemplo, levando a aplicação a um estado seguro).

A Figura 28 ilustra um exemplo de envelope operacional e as condições onde o serviço é considerado disponível e indisponível.

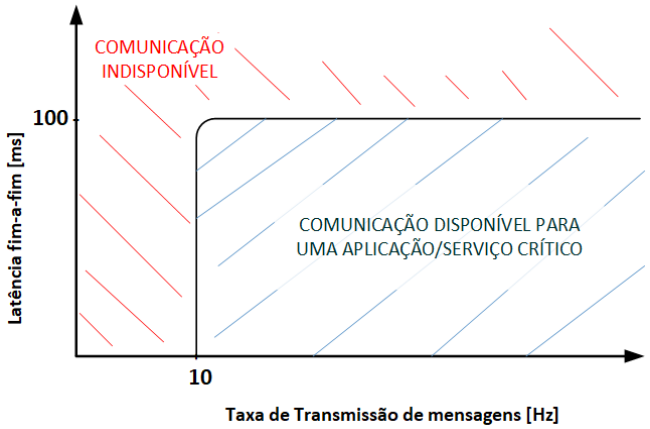


Figura 28 – Exemplo de envelope operacional de comunicação em aplicações críticas

A Figura 29 ilustra os envelopes operacionais para dois cenários de Direção Avançada utilizando comunicação V2X baseada em 5G (ETSI, 2022): Cenário 1 (troca de informação entre elementos para fins de direção automatizada), que demanda requisitos **menos restritivos** – permite operar com maiores latências e menores taxas de transmissão; e Cenário 2 (prevenção cooperativa de colisões), que demanda requisitos de comunicação **mais restritivos** – opera apenas com baixas latências e altas taxas de transmissão.

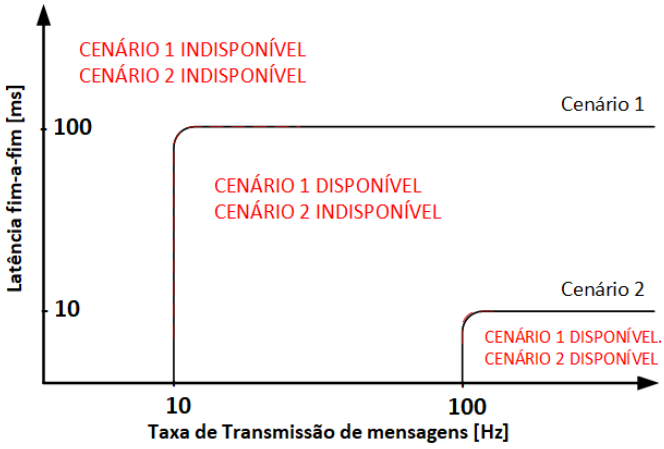


Figura 29 – Envelopes operacionais (Máxima Latência Fim-a-Fim [ms] x Taxa de Transmissão mínima [Hz]) para dois cenários de direção avançada utilizando comunicação V2X

No presente estudo de caso, o cenário-base é utilizado na avaliação do impacto do desempenho da comunicação V2X (realizada pelo CSP) sobre o nível de segurança crítica da aplicação. O desempenho da comunicação é modelado e avaliado por meio de dois parâmetros: Latência Fim-a-Fim e Taxa de Transmissão. Devem ser avaliadas tanto as **condições** nas quais os requisitos de desempenho estabelecidos no SLA são atendidos quanto as condições de desempenho de comunicação não atendem ao envelope operacional (SLA). Desta forma, três cenários de desempenho da comunicação V2X são considerados, conforme ilustrado na Tabela 4:

- Comunicação V2X atende aos requisitos de comunicação (condições menos e mais restritivas);
- Comunicação V2X atende ao requisito de Latência menos restritivo, e a Taxa de Transmissão está degradada (3 níveis de degradação);
- Comunicação V2X atende ao requisito de Taxa de Transmissão menos restritivo, e a Latência está degradada (3 níveis de degradação).

Tabela 4 – Cenário de Desempenho da Comunicação V2X a serem avaliados (ETSI, 2022)

<table>
<tr><th colspan="2">Cenário de Teste</th><th>Latência Fim-a-Fim</th><th>Taxa de Transmissão</th></tr>
<tr><td rowspan="2">V2X atende às especificações.</td><td>Mais restritivo</td><td>3 ms</td><td>100 Hz</td></tr>
<tr><td>Menos restritivo</td><td>100 ms</td><td>10 Hz</td></tr>
<tr><td rowspan="3">Latência atende a especificação.
Taxa de Transmissão degradada.</td><td rowspan="3">Menos restritivo</td><td rowspan="3">100 ms</td><td>0,1 Hz</td></tr>
<tr><td>0,5 Hz</td></tr>
<tr><td>1 Hz</td></tr>
<tr><td rowspan="3">Latência degradada.
Taxa de Transmissão atende a especificação.</td><td rowspan="3">Menos restritivo</td><td>500 ms</td><td rowspan="3">10 Hz</td></tr>
<tr><td>1.000 ms</td></tr>
<tr><td>5.000 ms</td></tr>
</table>

O nível de segurança crítica da aplicação é avaliado por métricas relacionadas ao risco de colisão entre AVs na região de cruzamento (RCz) das vias 1 e 2 durante a prestação do serviço de Semáforo Virtual provido pelo CCO (CaTraCa), considerando cada um dos cenários de desempenho de comunicação V2X definidos na Tabela 4.

Uma **métrica direta** de risco de colisão é a "**Taxa de Colisões**" (quantidade de colisões por unidade de tempo). Contudo, colisões (acidentes) tendem a ser eventos de baixa frequência, difíceis de estimar ou obter. Assim, métricas indiretas (também definidas como variáveis

'proxy'[43]) – relacionadas à eventos de maior frequência de ocorrência (como incidentes) – podem ser úteis na estimativa de risco de colisão.

O risco de colisão em sistemas de tráfego está diretamente relacionado às **distâncias mínimas** (dmin(i,j)) aplicadas entre pares de veículos ({i, j}). Colisões (**acidente**) são eventos cuja distância mínima observada é zero (dmin(i,j) → 0). Quanto menor a distância mínima observada entre dois veículos ao longo do tempo, maior o risco de colisão e, consequentemente, **menor** será o nível de segurança observado.

Desta forma, pode-se estimar métricas indiretas de risco de colisão utilizando estatísticas das distâncias mínimas entre veículos. Estimadores de posição – como a Esperança (E[dmin(i,j)]) ou média (μ[dmin(i,j)]) – e de dispersão – como a Variância (VAR[dmin(i,j)] ou desvio-padrão (σ^2[dmin(i,j)] – podem ser estatísticas empregadas para estimativa de risco de colisão. Além da possibilidade de medição contínua ao longo do tempo, gerando mais dados e produzindo estimadores mais confiáveis, estatísticas sobre dmin(i,j) são implementáveis em sistemas reais por meio do monitoramento dos veículos ao longo do tempo.

Portanto, as **métricas de segurança crítica** obtidas com o cenário-base para avaliar a relação entre desempenho da comunicação V2X e nível de segurança crítica da aplicação são:

- Taxa de Colisão entre AVs = [nº de colisões]/[período de tempo].
- {μ[dmin(i,j)], σ[dmin(i,j)][44]}: média/desvio-padrão das distâncias mínimas (dmin(i,j)) observadas entre todos os pares de veículos (i,j) que cruzam a RCz.

Com este estudo de caso, pretende-se obter métricas de risco observadas na aplicação em função dos cenários de desempenho da comunicação V2X aos quais o cenário-base é submetido.

{Taxa_Colisão; **S**[dmin(i,j)]} = **F**(Latência_Fim-a-Fim; Taxa_Transmissão)

Onde:

[43] "Variáveis 'Proxy' são variáveis utilizadas para substituir outra variável de difícil mensuração e que se presume guardar com ela relação de pertinência" (ABNT, 2011). É uma variável que possa ser observada e que esteja fortemente correlacionada com a variável de interesse, cuja a observação direta é difícil (ou impossível) de se obter.

[44] A justificativa para se utilizar esta métrica relacionada à segurança é apresentada, a seguir, ao longo da descrição do cenário-base com suporte da Abordagem Proposta.

S[dmin()] representam as estatísticas de média e desvio-padrão de dmin(i,j).

F(.) representa a função Risco de Segurança em função dos parâmetros do sistema.

Por fim, vale frisar que, neste *Estudo de Caso*, o CSP (sistema de comunicação) é considerado disponível pela aplicação crítica mesmo quando não atende ao SLA. Assim, os dados obtidos por meio de modelagem e simulação computacional – detalhados a seguir – são utilizados visando alcançar os objetivos deste estudo de caso. Além disso, é possível verificar se os requisitos de desempenho de comunicação para a aplicação crítica, definidos no SLA, são suficientes (promovem uma operação segura) ou estão sub/superdimensionados para aquela aplicação.

4.1.2. Implementação do cenário-base no ambiente computacional virtual

Após sua definição, o cenário-base do estudo de caso é modelado, implementado, executado e avaliado por meio do **ambiente computacional virtual (simulado)** (VISMARI *et al.*, 2018b)). Esse ambiente computacional permite modelar e simular, tanto em tempo real quanto em tempo acelerado, cenários de sistemas de transporte rodoviário (RTS – *Road Transportation Systems*). Nos cenários RTS, é possível considerar as características das vias, ambiente, veículos e condutores – incluindo algoritmos de condução automática ou autônoma – interagindo entre si.

Desta forma, esse ambiente permite avaliar o comportamento de veículos autônomos e em cenários de tráfego, bem como o impacto de novos conceitos (como o Semáforo Virtual), tecnologias (como a comunicação V2X cooperativa e a Inteligência Artificial) e procedimentos sobre sua segurança de aplicações relacionadas aos RTS baseados em Sistemas de Transporte Inteligente (ITS).

A Figura 30 ilustra a arquitetura de alto nível do ambiente computacional utilizado neste trabalho. Fazendo uso de elementos RTS (via, veículo, condutor) – elemento ‘1’, elabora-se o Caso de Uso em RTS a ser avaliado, constituído de sua arquitetura, especificações e procedimentos de análise – elemento ‘2’. Na sequência, este Caso de Uso é implementado e executado por meio de simulação em tempo acelerado (FTS – *Fast Time Simulation*) – elemento ‘3’ – e/ou simulação em tempo real (ReTS – *Real Time Simulation*) – elemento ‘4’. Para estes fins, foram adotadas ferramentas computacionais de padrão aberto e adaptadas. Ao final, os

resultados obtidos com as simulações podem ser avaliados de acordo com os procedimentos de análise adotados.

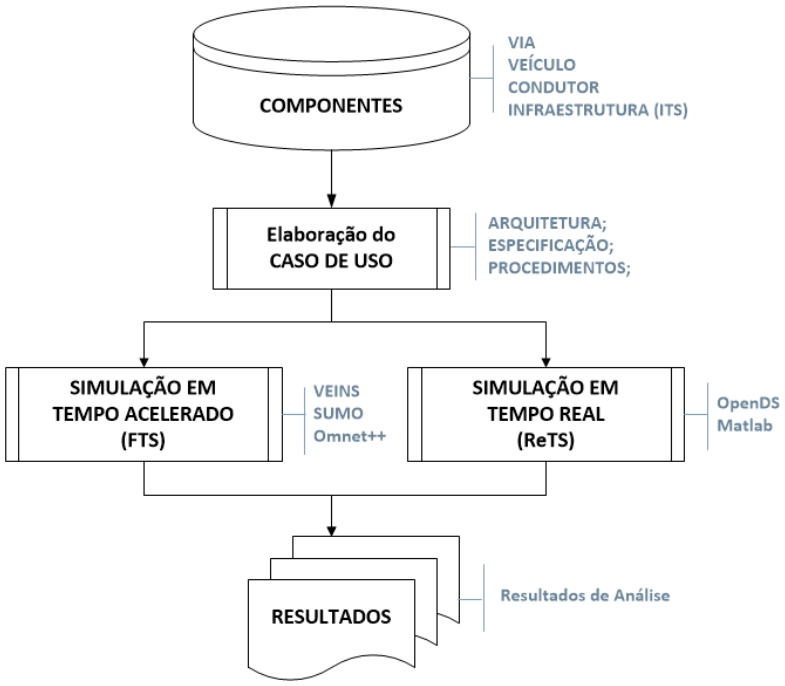


Figura 30 – Arquitetura do ambiente computacional para análise de segurança de RTS.
Fonte: (VISMARI *et al.*, 2018b)

A implementação do presente estudo de caso inicia com o detalhamento da modelagem do cenário-base (ilustrado na Figura 24), conforme ilustrado na Figura 31. Definem-se as especificações detalhadas do ambiente do cenário-base, como dimensões físicas (largura e comprimentos) das vias 1 e 2, dimensões (largura e comprimento) e características dinâmicas dos AVs (taxas de aceleração e desaceleração), dimensões físicas da Região de Controle (RCt), localização e dimensões da Região de Colisão (RC), taxa de ingresso de veículos na RCt, velocidades máximas dos veículos dentro e fora da RCt, entre outros.

Em seguida, este modelo é implementado e simulado utilizando a abordagem de simulação por tempo acelerado (FTS). A Figura 32 apresenta a arquitetura detalhada do módulo FTS (elemento ‘3’), com destaque para as ferramentas:

- SUMO[45] (*Simulation of Urban MObility*), um simulador de tráfego rodoviário. Esta ferramenta é utilizada para modelar e simular tanto os veículos (AVs) quanto as vias (1 e 2) do cenário-base, bem como a dinâmica de tráfego;

- OMNet++[46] (*Objective Modular Network Testbed in C++*), um simulador de rede baseada em eventos. Esta ferramenta permite modelar e simular a comunicação V2X, considerando os AVs (e demais componentes relacionados à comunicação, como o RSU) como nós de uma rede de comunicação de dados que implementa protocolos específicos de comunicação. O comportamento dos condutores dos AVs (algoritmos relacionados à supervisão e controle do cenário-base) também pode ser modelado no OMNET++; e

- VEINS[47] (*Vehicles in Network Simulation*), que integra o SUMO e o OMNet++, permitindo a simulação de redes de comunicação V2X em ambientes de tráfego rodoviário.

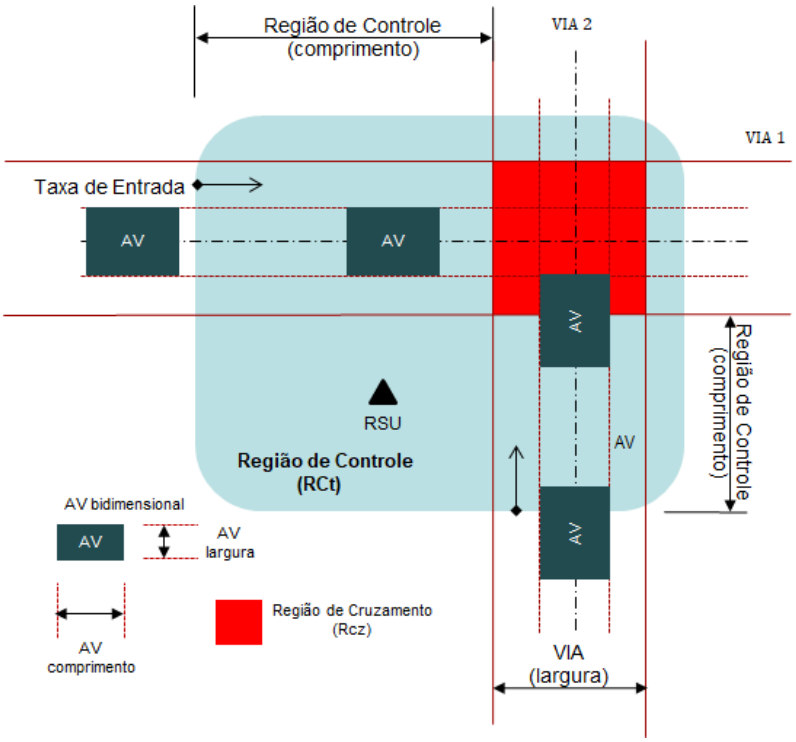


Figura 31 – Modelo detalhado do cenário-base do estudo de caso

[45] https://sourceforge.net/projects/sumo/
[46] https://omnetpp.org/
[47] https://veins.car2x.org/

Além das ferramentas de modelagem e simulação, a Figura 32 ilustra um componente de automação do processo de simulação desenvolvido para este estudo. Assim, pode ser considerado como parte integrante do Módulo FTS, conforme detalhado adiante.

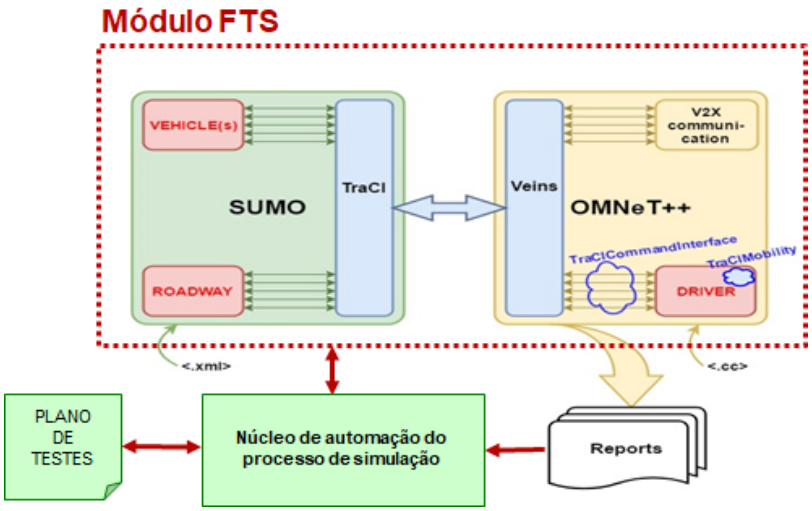


Figura 32 – Componentes principais do Módulo FTS e componente de automação.
Fonte: adaptado de VISMARI (2018b)

Os serviços de comunicação V2X são implementados pelos recursos da ferramenta OMNet++, conforme ilustrado na Figura 32. Na topologia de rede implementada neste cenário-base, os **AVs** e o **RSU** (*Road Side Unit*) – que emula o **CCO** e embarca o algoritmo do serviço CaTraCa (semáforo virtual) – trocam mensagens usando comunicação do tipo V2I (Veículo para Infraestrutura) suportada pelo protocolo 4G/LTE. Os AVs enviam ao RSU, por radiodifusão, o conteúdo de **Mensagens Básicas de Segurança** (**BSM** – *Basic Safety Message*)[48] (SAE, 2022). A RSU envia as velocidades-alvo para cada AV dentro de sua área de cobertura (Região de Controle - RCt).

O SUMO, ferramenta aplicada na modelagem e simulação dos veículos (AVs), das vias (1 e 2) e da dinâmica de tráfego do cenário-base, é uma ferramenta de modelagem **unidimensional**. Assim, tanto vias quanto veículos são modelados apenas por seu comprimento, conforme ilustrado na Figura 33. Métricas de risco de colisão obtidas por meio da modelagem de elementos unidimensionais podem não ser representativas para cenários de tráfego reais, bidimensionais. Com isso, produzem-se estimativas pouco conservadoras, onde

[48] O conteúdo contido nas mensagens BSM dos AVs modelados neste trabalho são *timestamp, position, maximum speed, actual speed, heading, length, maximum acceleration, actual acceleration*.

os riscos estimados são menores do que os esperados. Desta forma, a modelagem e simulação do ambiente do cenário-base bidimensional – largura/comprimento dos veículos (AVs) e das vias (1 e 2) – demanda por aperfeiçoamentos quando o SUMO é utilizado.

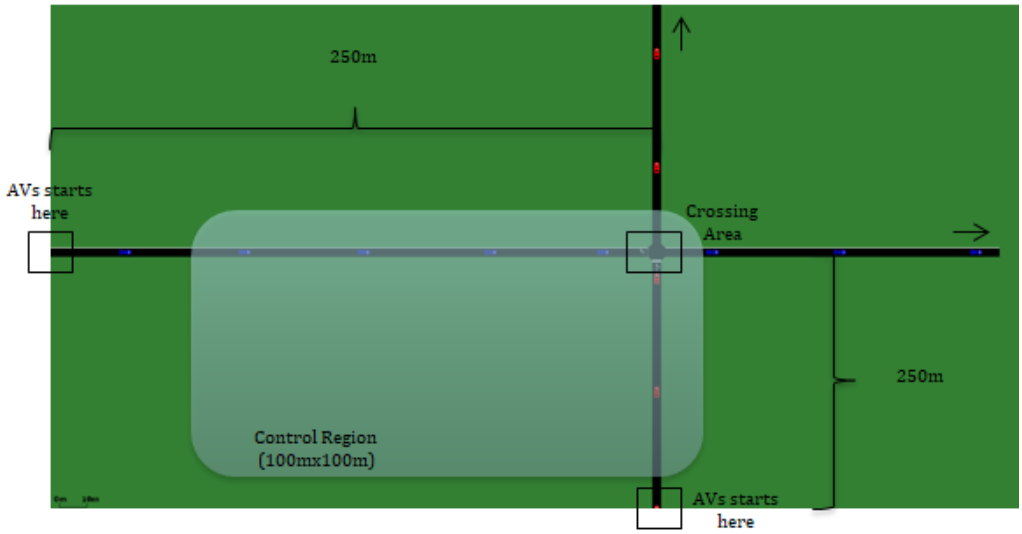


Figura 33 – Exemplo de implementação do cenário-base no SUMO (unidimensional)

Conforme discutido anteriormente, obter estimativas de riscos de colisão entre objetos (neste caso, AVs) demanda estimar a dmin(i,j) entre todos os AVs ao longo do tempo e identificar as situações onde dmin(i,j)→0, para todos os pares de AVs (i,j). Contudo, o cálculo de dmin(i,j) para objetos bidimensionais deve considerar os pontos de maior proximidade entre objetos (AVs), o que pode ser **oneroso em termos computacionais**.

Outra abordagem para estimar dmin(i,j) seria identificar as regiões de interseção (Iij) entre as trajetórias dos objetos (veículos) e estimar se estas regiões estarão ocupadas, ao menos em partes, por dois (ou mais) objetos (‘i’ e ‘j’) em um mesmo instante/intervalo de tempo. A Figura 34 ilustra um exemplo de cenário onde estão identificadas cinco regiões de interseção (Iij) formadas pelas trajetórias de quatro objetos bidimensionais. Contudo, em cenários cujas trajetórias dos objetos podem ser modificadas em mais de um grau de liberdade, as regiões Iij se alteram ao longo do tempo, em localização e/ou forma. Assim, estimativas da frequência de colisões – e, consequentemente, riscos – conteriam incertezas proporcionais ao horizonte de tempo no qual são estimadas, além de também serem computacionalmente onerosas.

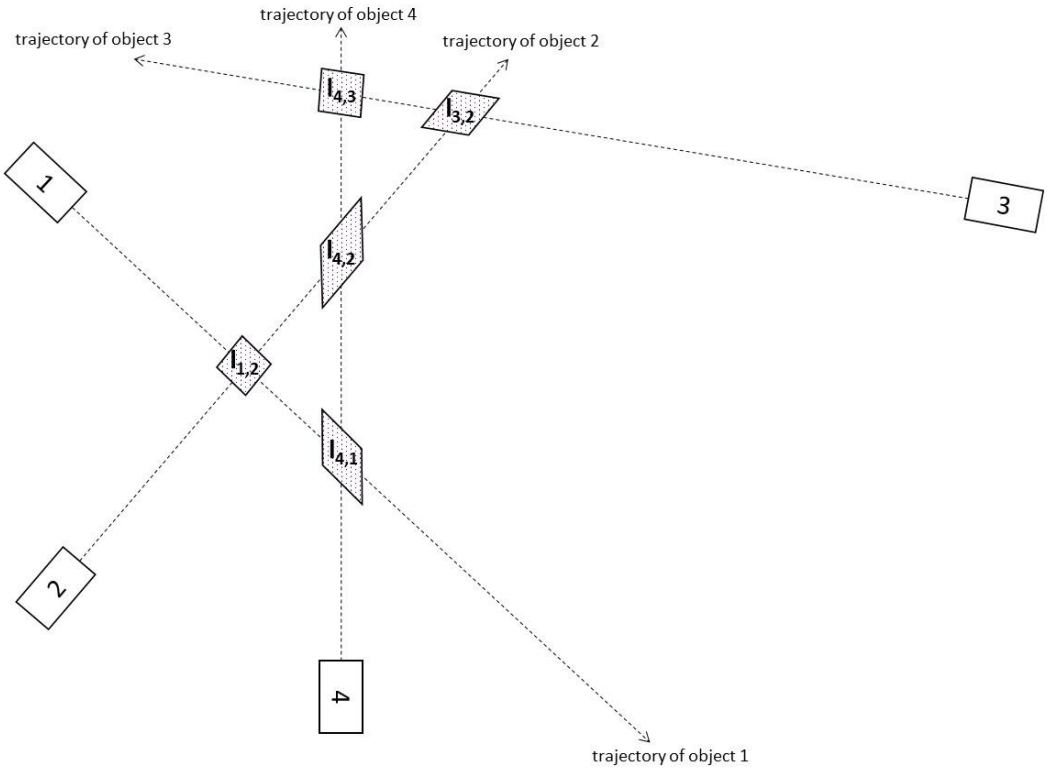


Figura 34 – Exemplo de regiões de interseção entre trajetórias

Por outro lado, em ambientes de **tráfego guiado**, onde os objetos se deslocam com apenas um grau de liberdade dentro de trajetórias bem definidas, as regiões Iij não se alteram ao longo do tempo, sendo possível simplificar os cálculos de dmin(i,j). Desta forma, foi desenvolvida uma **técnica de propósito geral para estimar riscos de colisão em ambientes de tráfego guiado**. Ela se baseia na geometria das vias (previamente conhecidas) e no monitoramento da distância entre objetos e Iij, estimadas com base nos dados de posição e velocidade enviados pelos AVs por meio de comunicação V2X. Esta técnica é apresentada, em detalhes, no **APÊNDICE I**.

A implementação desta técnica no estudo de caso é ilustrada na Figura 35. Considera-se o ambiente de tráfego do cenário-base como um sistema de tráfego guiado, cuja trajetória é formada pelas vias 1 e 2. O ponto central da região de interseção entre vias ($I_{1,2}$) é formado pela interseção ortogonal das vias (unidimensionais) 1 e 2. Para obter uma região de interseção bidimensional, que corresponde a **Região de Colisão (RC)** contida dentro da Região de Cruzamento (RCz), adiciona-se a largura dos AVs – incluindo uma margem de tolerância para incorporar possíveis desvios laterais dos AVs – centrada no ponto de interseção e ortogonal às trajetórias das vias 1 e 2.

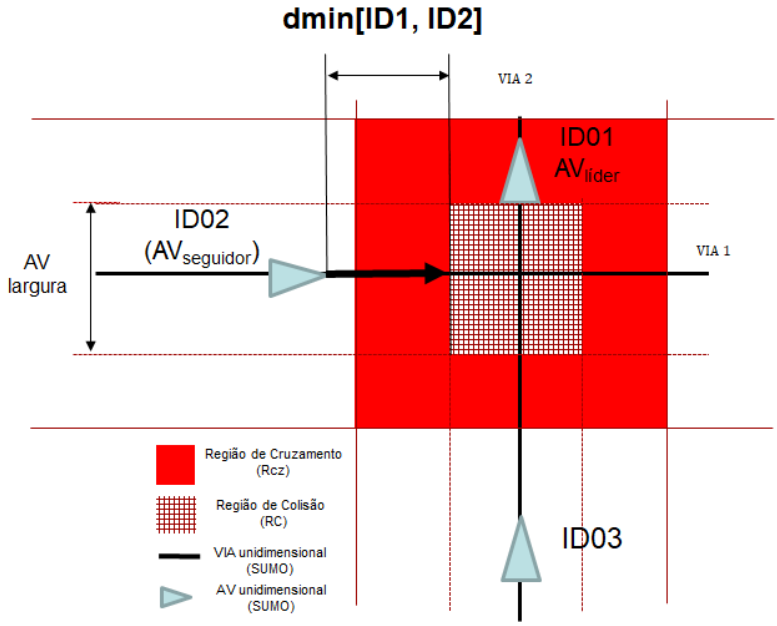


Figura 35 – Exemplo de regiões de interseção e colisão entre vias 1 e 2

Para calcular o **risco de colisão** baseado em distâncias mínimas, obtêm-se continuamente a distância entre o ponto de entrada na região de colisão da via e a parte frontal do AV trafegando por esta via e mais próximo deste ponto, e que ainda não esteja na Região de Colisão (RC). Quando um veículo alcança esta região, a **distância mínima (dmin)** observada para um par de AVs (dmin(i,j)) é obtida pelo menor valor de distância monitorada em cada via no instante em que o primeiro veículo ('j') sai da RC. Considera-se uma **colisão** quando alguma das distâncias monitoradas atinja zero antes do primeiro veículo não ter saído da RC.

Após a implementação da modelo computacional do cenário-base, obtêm-se as **métricas de risco de colisão** executando uma **campanha** de simulações orientada pelos **parâmetros de entrada** definidos pelos cenários de desempenho da comunicação V2X (sobretudo posição dos AVs) e pelas condições configuradas para a campanha de simulação (CS). Para cada **ciclo** da **campanha de simulação**, um par 'n' de parâmetros de desempenho de comunicação ($\{\text{Latência, Taxa\_Transmissão}\}_n$) é configurado. A cada ciclo de simulação (com uma duração pré-definida), os AVs ingressam na RCt (tanto na Via 1 quanto na Via 2) a uma taxa 'm' configurada. O serviço de CATraCa é provido a estes AVs desde seus ingressos na RCt até suas saídas das RCz.

Sempre que um AV finaliza o cruzamento da Região de Colisão (RC), independentemente da via que trafegou, registra-se o número (ID) deste AV (denominado como

"**AV líder**"), o ID do **AV seguidor** (AV mais próximo da RC trafegando na outra via), o status daquela interação ({Colisão, Não Colisão}) e a distância mínima (dmin(AV_lider, AV_seguidor)) observada neste evento de cruzamento. A estrutura de dados deste registro é ilustrada na Figura 36.

AV_seguidor (#AV)	AV_líder (#AV)	Status do AV_lider {colisão, não colisão}	dmin (m)

Figura 36 – Estrutura de dados gerado a cada evento de cruzamento de AVs em um ciclo de simulação

Ao final de um ciclo de simulação, obtém-se uma **base de dados** contendo o **conjunto de registros** gerados por todos os eventos de cruzamento realizados por AVs (controlados pelo CCO) naquele ciclo de simulação. Desta base de dados, extrai-se o **número total de pares de AVs** que interagiram na RCz (igual à quantidade de registros da base), o **Tempo Total** da simulação e o **número de colisões** observadas na campanha. Com estes valores são utilizados para estimar a **Taxa de Colisões** (nº colisões/Tempo Total). Além disso, estima-se a média **(μ[dmin])** e o desvio-padrão **(σ[dmin])** do conjunto de **distâncias mínimas** (dmin(i,j)), observados no ciclo de simulação. Estes dados, indexados pelo par {Latência, Taxa_Transmissão}, formam um registro da base de dados da campanha de simulação. Sua estrutura de dados é apresentada na Figura 37.

Entradas			Saídas		
Latência	Taxa_Transmissão	CS	Taxa_Colisão	**μ** [dmin]	**σ** [dmin(i,j)]

Figura 37 – Estrutura de dados do registro gerado por um ciclo de simulação

Ao final de uma campanha de simulações, o cenário-base terá sido submetido a todos os "K" cenários de desempenho de comunicação considerados na avaliação ('K' pares {Latência, Taxa_Transmissão}), produzindo uma base de dados com "K" registros. Esta base é utilizada para as avaliações e cumprimentos de objetivos definidos para este estudo de caso. Sobretudo, implementação de prova de conceito e avaliação da instanciação da Abordagem Proposta, detalhada a seguir.

4.2. A Abordagem Proposta para o cenário-base

Esta seção trata da aplicação do conceito da abordagem proposta, como ferramenta, em um contexto de tráfego baseado em C-ITS. Por se tratar de um conceito dependente de contexto, o desenvolvimento da Abordagem Proposta instanciada neste estudo de caso é orientado pelas características do cenário-base definidas anteriormente. A abordagem desenvolvida é implementada no modelo computacional do cenário-base. Por meio de simulação computacional, obtêm-se os dados para avaliação da sua efetividade sobre o cenário-base.

4.2.1. Definição da Abordagem Proposta orientada pelo cenário-base

No cenário-base, o CATraCa é um serviço provido por um sistema crítico em segurança baseado em comunicação. Embarcado no CCO, deve gerenciar o fluxo de veículos na região de cruzamento (RCz) e **manter um nível de risco de colisão aceitável** em todas as situações esperadas (normais e degradadas) de tráfego e de sistema crítico – incluindo variações de demanda, degradações de desempenho e falhas em elementos.

Conforme apresentado anteriormente, nos atuais sistemas críticos baseados em comunicação, a comunicação não é considerada um elemento crítico em segurança. Consequentemente, em caso de degradação do serviço de comunicação – comumente realizada por terceiros (Provedores de Serviços de Comunicação – CSP), os elementos do sistema crítico são os responsáveis por garantir a segurança da aplicação. Assim, a disponibilidade dos sistemas críticos tende a ser negativamente impactada quando o CSP não garante os requisitos operacionais especificados nos SLAs.

Para que seja possível garantir o atendimento aos requisitos acordados no SLA, os recursos de infraestrutura de comunicação alocados àqueles usuários críticos tendem a ser superdimensionados pelo CSP. Contudo, para que não se comprometa a qualidade de serviço (QoS) da comunicação provida aos demais usuários, incluindo usuários de sistemas não críticos em segurança, toda a infraestrutura de rede deve ser superdimensionada, acarretando em aumento em custos e complexidade.

Conforme discutido na seção 3, além de sua aplicação no processo de garantia de segurança de uma aplicação crítica, a abordagem proposta também pode ser aplicada ao gerenciamento e otimização de recursos do sistema orientado aos riscos da aplicação. Para que

seja possível, e conforme discutido na seção anterior, a Abordagem Proposta deve **monitorar métricas de risco da aplicação durante a operação do sistema.** Desta forma, dado que a implementação da Abordagem Proposta sobre o cenário-base deste estudo de caso é modelada e avaliada computacionalmente, as **métricas de risco de segurança** (sobretudo **dmin()**) estão disponíveis durante a simulação da operação do sistema.

Portanto, neste estudo de caso, também é possível avaliar se a abordagem proposta possibilita otimizar, em tempo de execução, a configuração dos serviços de comunicação utilizados pela aplicação crítica. Isso corresponde a alocar dinamicamente – em função tanto do risco de colisão observado na aplicação (cruzamento entre vias) quanto da qualidade de serviço (QoS) observada no sistema de comunicação (Latência e Taxa de Transmissão) – os recursos de comunicação de forma que sejam suficientes para manter a segurança crítica. Desta forma, espera-se promover uma ampliação do envelope operacional dos serviços de comunicação definidos nos SLA, reduzindo a demanda por superdimensionamento de recursos e o impacto na disponibilidade do sistema.

Considerando os parâmetros Latência e Taxa de Transmissão para o sistema de comunicação (CSP), a Figura 38 ilustra o **envelope operacional especificado** para o cenário-base sem suporte da Abordagem Proposta, bem como o **envelope operacional esperado** quando a Abordagem Proposta é aplicada sobre o cenário-base.

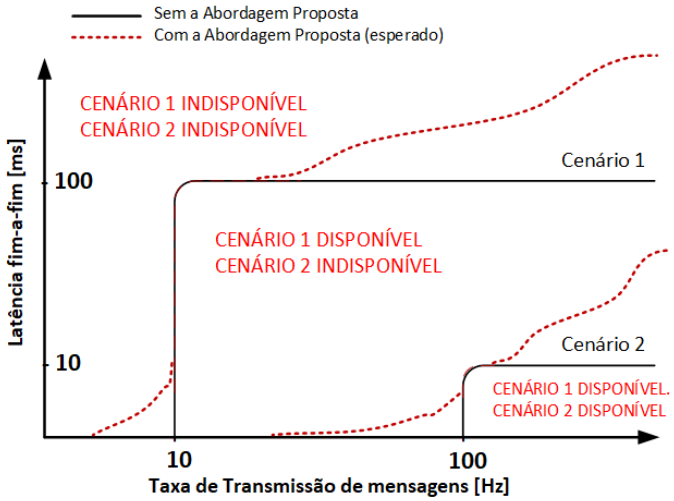


Figura 38 – Ampliação esperada dos Envelopes operacionais para dois cenários de direção avançada utilizando comunicação V2X

Por princípio, os serviços de comunicação são **descontinuados** quando os parâmetros definidos no SLA (envelope operacional) não são atendidos. Assim, uma Latência fim-a-fim maior que 100ms [3ms] e/ou uma Taxa de Transmissão de mensagens menor que 10Hz [100Hz] levariam à indisponibilidade do serviço de comunicação e, consequentemente, interrupção do serviço crítico. Por outro lado, quando a Abordagem Proposta é aplicada sobre o cenário-base, espera-se que os serviços providos pelo sistema crítico continuem disponíveis durante eventos disruptivos de comunicação, onde o envelope Latência *vs.* Taxa de Transmissão definidos no SLA não sejam atendidos.

Conforme apresentado na definição do cenário-base, todas as informações trocadas entre o CCO e os Veículos Autônomos (AVs) trafegam pelo CSP (serviços de comunicação). Desta forma, o **monitoramento contínuo** do nível de risco de colisão na aplicação crítica e da QoS observada no sistema de comunicação, bem como a **reconfiguração de parâmetros dos serviços de comunicação**, pode ser realizado diretamente no CSP. A Figura 39 ilustra o relacionamento entre a Abordagem Proposta e os demais elementos do cenário-base adotado neste estudo de caso.

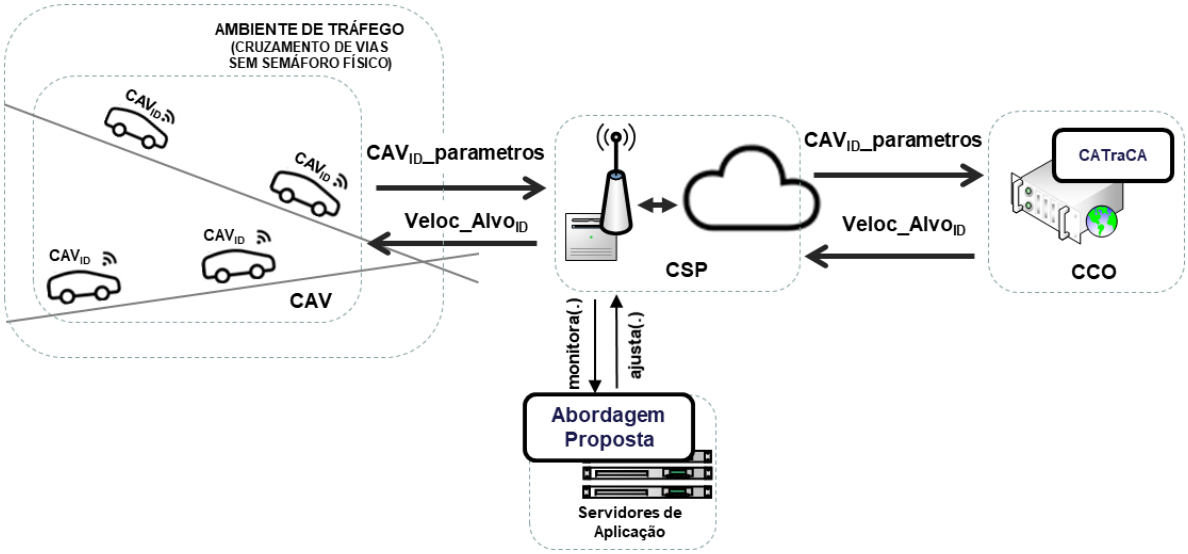


Figura 39 – Relacionamento entre elementos do cenário-base e a Abordagem Proposta

Neste estudo de caso, a arquitetura adotada para a implementação da Abordagem Proposta no cenário-base é ilustrado na Figura 40. Nesta arquitetura, observa-se o relacionamento entre o **Gerador de Dados**, o processo de **Desenvolvimento do Modelo Executivo** e o **Modelo Executivo** da Abordagem Proposta, bem como os dados de entrada (monitora(.)) e saída (ajusta(.)) destes elementos, apresentados a seguir.

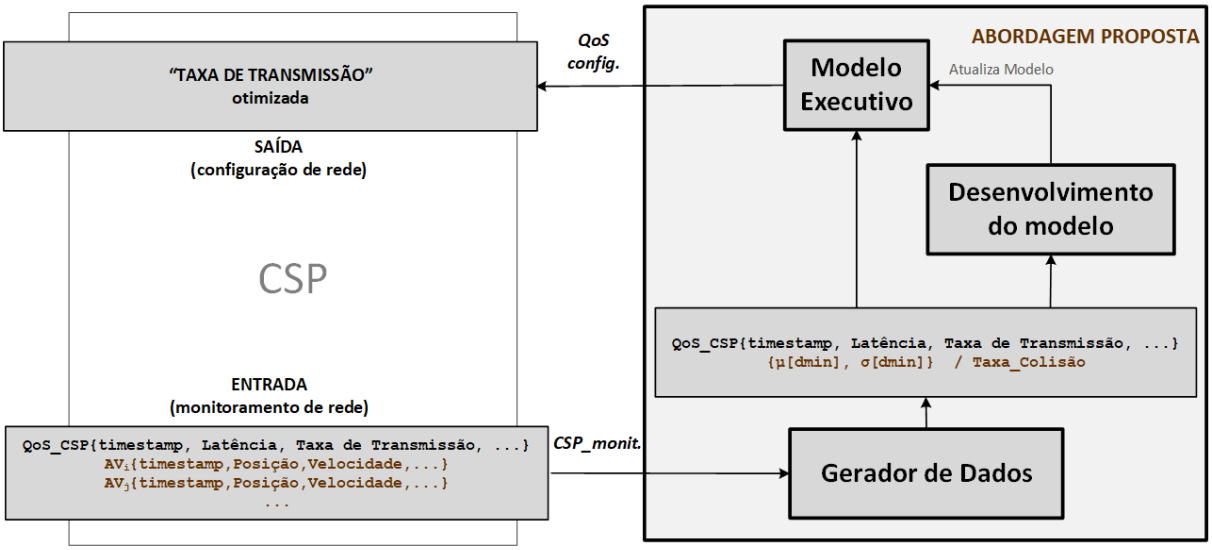


Figura 40 – Arquitetura da Abordagem Proposta para o Estudo de Caso

Com relação aos **dados de entrada** na implementação da Abordagem Proposta (monitorados no elemento de comunicação – CSP), obtêm-se:

- Do **Sistema de Comunicação** (dados de QoS), a Latência fim-a-fim e a Taxa de Transmissão de mensagens observados no CSP em um determinado instante de tempo (*timestamp*).

- Da **Aplicação Crítica**, os dados abertos[49] das mensagens transmitidas pelos AVs dentro da Região de Controle e trafegadas pelo sistema de comunicação: posição, velocidade atual, aceleração atual, direção, comprimento, velocidade máxima e aceleração máxima de cada AV em um determinado instante de tempo (*timestamp*).

Os dados monitorados do CSP relacionados à QoS – valores de latência (em ms) e taxa de transmissão (em Hz) – podem ser utilizados diretamente pelo **Modelo Executivo**[50] da implementação da **Abordagem Proposta**, sem necessidade de pré-processamento. Por outro lado, os dados obtidos da aplicação crítica, como posição e velocidade dos AVs, precisam ser processados (para obtenção de estimativas do nível de risco de colisão) antes de ser aplicado na

[49] Neste estudo de caso, adotam-se as **Mensagens Básicas de Segurança** (**BSM** – *Basic Safety Message*) (SAE, 2022)

[50] O Modelo Executivo instancia a atividade de tomada de decisão (**atividade 4 – 'compensar'**) da abordagem proposta. Neste estudo de caso, o modelo executivo monitora as entradas (Latência e Taxa de Atualização), aplica as regras de decisão implementadas e retorna o valor controlado.

Abordagem Proposta. Desta forma, a implementação da Abordagem Proposta deve considerar um **Gerador de Dados** que processe e produza os dados[51] a serem utilizados por seu **Modelo Executivo**.

O conjunto de dados produzido pelo ***Gerador de Dados*** é utilizado tanto no processo de **desenvolvimento** (*offline*) quanto na **execução** (*online*) do **Modelo Executivo**, onde:

- Para o **desenvolvimento** do Modelo Executivo, os dados produzidos continuamente pelo Gerador de Dados são armazenados, criando as bases de dados utilizadas para **identificar padrões de comportamento** na relação entre o desempenho do sistema crítico (neste caso, sistema de comunicação) e os níveis de risco de colisão observados sobre o cenário da aplicação. Os **padrões de comportamento** identificados são utilizados para implementar os **critérios de tomada de decisão** (regras/conhecimento) que o Modelo Executivo deve seguir.

- Para a **execução** (operação) do Modelo Executivo, os dados produzidos continuamente pelo **Gerador de Dados** são avaliados pelo Modelo Executivo implementado. Baseado nas regras implementadas durante seu desenvolvimento, o Modelo Executivo decide qual a **taxa de transmissão de mensagens** (parâmetro controlável no cenário-base) deve ser configurada no serviço de comunicação (CSP) utilizado pela aplicação crítica (CCO).

Conforme relatado anteriormente, o **Gerador de Dados** processa os dados de consciência situacional dos AVs para estimar o nível de risco de colisão da aplicação. A consciência situacional é obtida por meio dos dados recebidos por meio de mensagens BSM (posição, velocidade, entre outros). Assim, o nível de risco de colisão é estimado **indiretamente**, com base nas *distâncias mínimas entre veículos* (**dmin()**) e suas estatísticas relacionadas (**μ[dmin])** e **σ[dmin]**)).

O processo utilizado pelo **Gerador de Dados** para estimar o nível de risco de colisão monitorada na aplicação é ilustrado na Figura 41. O componente "[1]" utiliza a mesma técnica aplicada anteriormente (e detalhes apresentados no **APÊNDICE I**) para calcular a menor

[51] Devido às características de comunicação dos AVs, onde as mensagens de consciência situacional são transmitidas por radiodifusão, os valores monitorados na rede podem estar em uma sequência diferente da gerada por seus nós (AVs). Assim, é necessário organizar estes dados de forma que se possa estimar o risco em função do tempo.

distância ocorrida entre pares de veículos (Veículo[i], Veículo[j], ...) próximos à região de cruzamento em um determinado instante de tempo (dmin(i,j)[t_n]). Esses valores são utilizados pelo componente "[2]", onde são calculadas as estatísticas (média e o desvio-padrão) do conjunto de **distâncias mínimas** (dmin(i,j)) para todos os pares (i,j) de AVs monitorados.

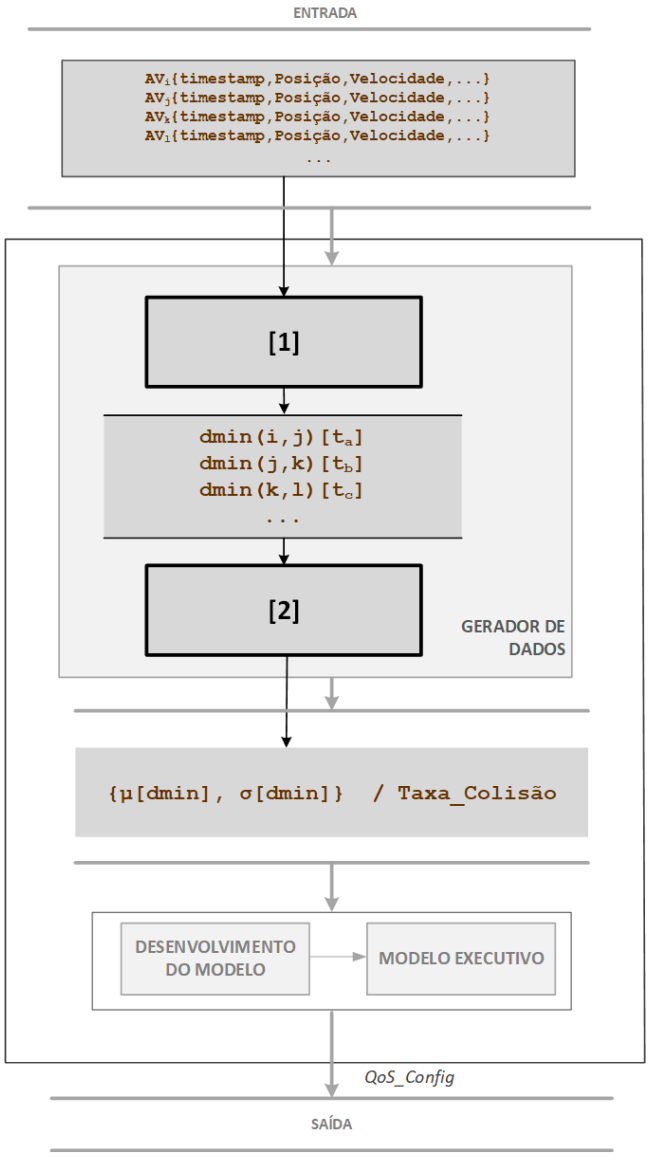


Figura 41 – Detalhe do processo de geração de estatísticas relacionadas à "distância mínima entre veículos"

O Modelo Executivo é o núcleo funcional da implementação da Abordagem Proposta. Este modelo deve contemplar todas as situações esperadas, incluindo as piores condições de operação da aplicação e de desempenho do sistema. Além disso, este modelo precisa ser desenvolvido, implementado e validado (atividades '1' a '3' definidas na Atividade Proposta) antes que possa ser executado (atividade '4' da Abordagem Proposta).

O Modelo Executivo poderia ser desenvolvido por meio de um **processo tradicional de engenharia**, onde seria concebido, especificado, projetado, implementado, verificado, validado e comissionado seguindo ciclo de vida, requisitos, técnicas e procedimentos bem estabelecidos. Além disso, por se tratar de um sistema crítico, um processo de garantia de segurança crítica deveria ser adotado para seu desenvolvimento – sobretudo bases normativas, conforme discutido na seção 2.1.

Contudo, os critérios de tomada de decisão (regras/conhecimento) implementados no Modelo Executivo são baseados nos padrões de comportamento identificados nas bases de dados obtidas com o monitoramento do cenário-base. A literatura atual reconhece que o Aprendizado de Máquina (ML – *Machine Learning*) é eficiente na tarefa de identificação/extração de padrões em bases de dados (SARKER, 2022), (MURDOCH *et al.*, 2019), (MAZHAR RATHORE *et al.*, 2021).

Desta forma, o desenvolvimento do Modelo Executivo da Abordagem Proposta utilizando técnicas de ML em Inteligência Artificial (AI) é realizado por um processo de clássico de treino, teste/validação e implementação, similar à abordagem apresentada por XU e GOODARE (2018). Detalhes do desenvolvimento do Modelo Executivo utilizando AI/ML são apresentados no **ANEXO I**.

Para fins de **verificação e validação**, a Abordagem Proposta é implementada no cenário-base do estudo de caso, conforme ilustrado pela Figura 42. O Modelo Executivo recebe, continuamente, os dados de entrada da Abordagem Proposta e decide a taxa de transmissão de mensagens que deve ser aplicada ao sistema de comunicação (CSP). Contudo, esta taxa é efetivamente configurada no CSP apenas quando a saída da Abordagem Proposta (Modelo Executivo) é conectada ao sistema de comunicação.

Com esta implementação, a **efetividade do Modelo Executivo** pode ser analisada observando sua capacidade em garantir os níveis de segurança da aplicação crítica. Esta avaliação é realizada pelo seguinte procedimento:

- Sem o suporte da Abordagem Proposta (Modelo Executivo com saída não conectada ao CSP), o cenário-base é exposto a situações degradadas do sistema de comunicação. Durante esta operação, observa-se o nível de risco de colisão produzido pela degradação do sistema; **(Situação-Controle)**

- Quando o risco de colisão mensurado no cenário-base estiver alto, ocorrendo muitas colisões devido ao sistema crítico não conseguir garantir a operação segura da aplicação, a saída do Modelo Executivo é conectada ao CSP. A partir deste instante, observa-se se o nível de risco de colisão é reduzido em relação à **situação-controle**. Caso seja significativamente reduzido, pode-se concluir que o **Modelo Executivo** – e a Abordagem Proposta – cumpre o objetivo esperado.

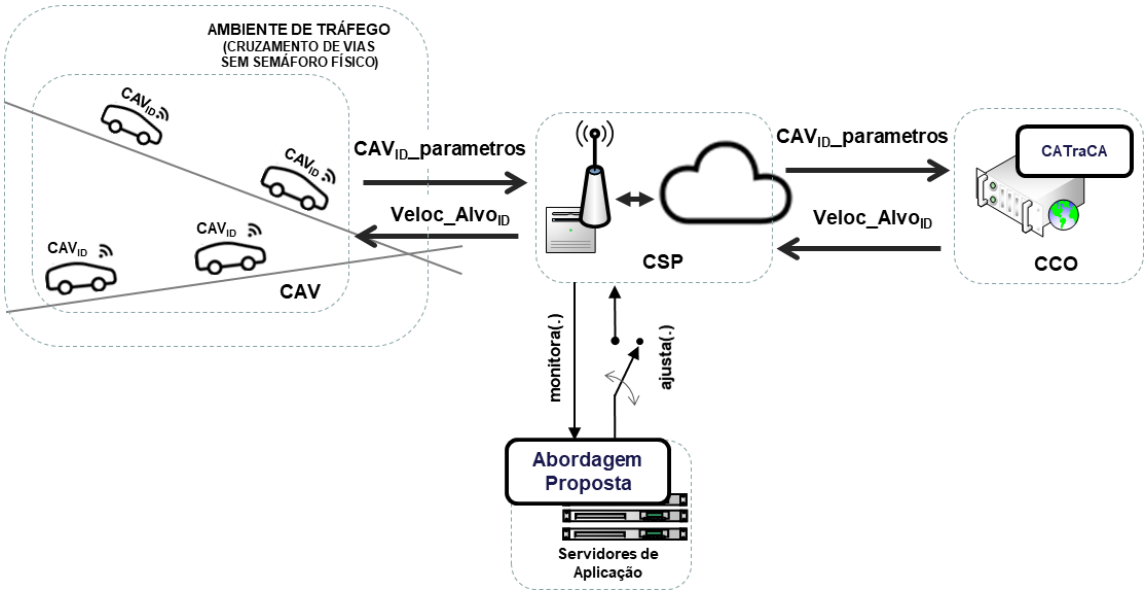


Figura 42 – Implementação da Abordagem Proposta no cenário-base

Por fim, a Figura 43 ilustra o diagrama de tempo e o fluxo das informações entre os elementos do cenário-base (AVs, CSP e CCO) sendo suportado pela Abordagem Proposta. Comparado ao fluxo de dados entre elementos do cenário-base, observa-se que os dados de entrada do Modelo Executivo são obtidos do CSP por meio do Gerador de Dados da Abordagem Proposta. O **Gerador de Dados** processa e disponibiliza os dados de entrada ("CSP_monit") tanto para a **execução** do modelo quanto para o **desenvolvimento** do modelo (***Dados_GD***). Assim, é possível suportar o cenário-base e, paralelamente ao longo do tempo, desenvolver modelos executivos atualizados.

Neste estudo de caso, o cenário-base é implementado e avaliado por meio do **ambiente computacional virtual (simulado)**. Desta forma, tanto o desenvolvimento (geração de base de dados) quanto a operação da Abordagem Proposta são realizadas no mesmo ambiente computacional. Os detalhes desta implementação, execução e validação do estudo de caso em ambiente computacional simulado são apresentados a seguir.

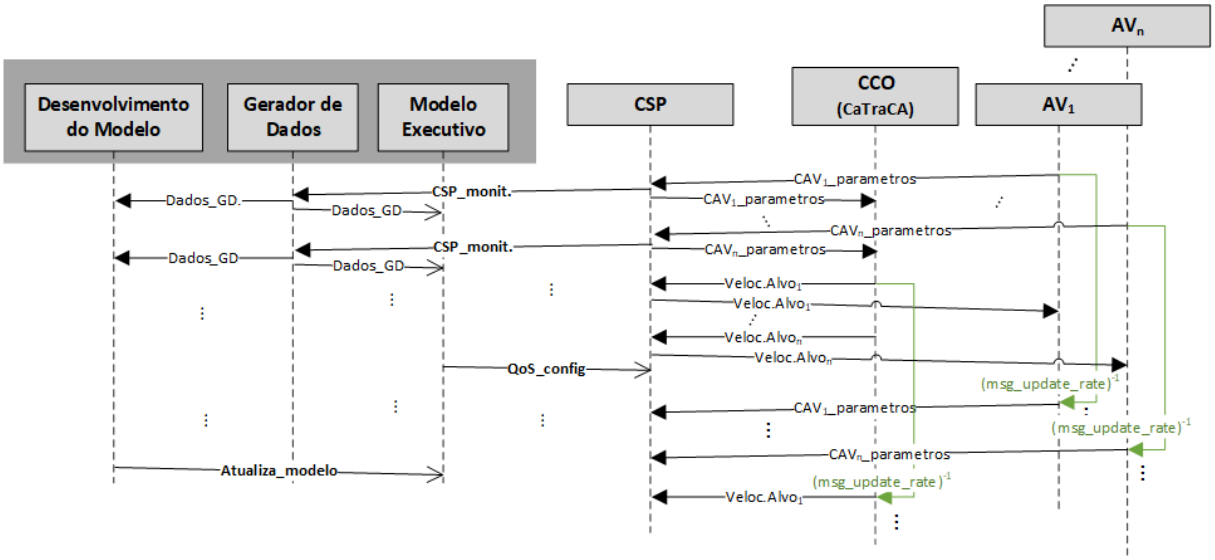


Figura 43 – Fluxo temporal de dados entre cenário-base e Abordagem Proposta

4.2.2. Implementação e avaliação da Abordagem Proposta

A implementação do cenário-base inicia com a apresentação do **ambiente computacional virtual (simulado)** utilizado para modelagem e simulação do presente estudo de caso. Em seguida, detalha-se a implementação (modelagem computacional) do **cenário-base** por meio das ferramentas de software (SUMO, OMNet++ e Veins) disponíveis no Módulo FTS (simulação por tempo acelerado) deste ambiente computacional (Figura 36). Por fim, apresenta-se o processo de simulação do cenário-base, com destaque às estruturas dos registros dos dados produzidos a cada evento de cruzamento de AVs ocorrido em um ciclo de simulação (Figura 36), ao final de cada ciclo de simulação (Figura 37) e ao final de uma campanha de simulação.

Este processo de simulação é ilustrado na Figura 44. Inicialmente, um ciclo de simulação é realizado pelo Ambiente Computacional para a configuração ‘Ti’ definida pelo **Plano de Testes**. As variáveis de entrada definidas para a Abordagem Proposta (obtidas pelo **Gerador de Dados**) são amostradas, continuamente, a cada **passo** do ciclo de simulação. Quando ocorre um evento de cruzamento de AV na Região de Colisão (RC), um registro de dados ‘$\mathbf{e_m}$’ (sobretudo o **dmin()** do evento) é registrado na Base de Dados ‘**< E >**’.

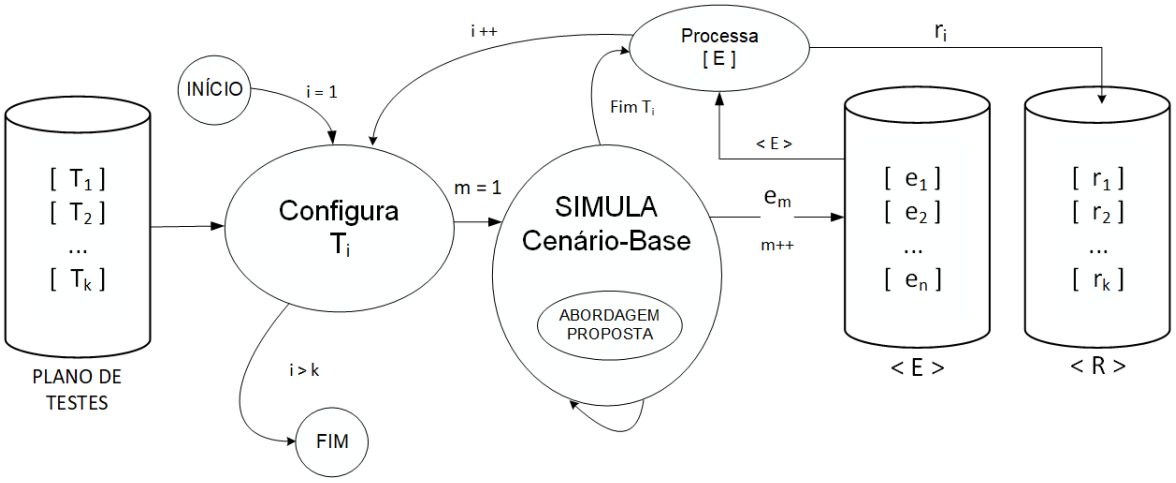


Figura 44 – Processo de simulação do cenário-base

Ao final do **ciclo** de simulação considerando '**T_i**', estimam-se as **estatísticas** relacionadas a **dmin()** (**μ[dmin]** e **σ[dmin]**) por meio de '**<E>**', gerando um registro de dados '**r_i**' na base '**<R>**' cuja estrutura é apresentada na Figura 37. Em seguida, um novo ciclo de simulação 'T_{i+1}' é iniciado. Estes ciclos se repetem até que o Plano de Testes esteja completamente executado, produzindo uma **base de dados '<R>'** relacionada à campanha de simulação. A base de dados **'<R>**' é utilizada como entrada do processo de **desenvolvimento** do **Modelo Executivo** da Abordagem Proposta, conforme ilustrado na Figura 44.

Conforme mencionado anteriormente, o processo de **Desenvolvimento do Modelo Executivo** é realizado por meio da identificação de **padrões de comportamento** sobre a Base de Dados de desenvolvimento '**<R>**' utilizando técnicas de AI/ML. Vale ressaltar que a **geração de modelos por AI/ML demanda uma grande quantidade de dados** e, consequentemente, de simulações. Por isso, **a execução das simulações é automatizada**, conforme ilustrado na Figura 32, pelo "Núcleo de automação do processo de simulação".

Inicia-se a execução do Plano de Teste definido para a campanha de simulação. Avaliam-se as diferenças dos resultados obtidos entre dois ciclos de simulações consecutivas (r_i e r_{i+1}). Caso esta diferença seja elevada, inclui-se um novo cenário no Plano de Teste ($T_{(i+1/2)}$), **interpolando as entradas** ({Latência, Taxa_Transmissão}) configuradas nos cenários de teste T_i e T_{i+1}. Este novo cenário é configurado e simulado, e os dados $r_{(i+1/2)}$ são coletados e armazenados em <R>. A Figura 45 ilustra o Algoritmo de Automação implementado.

Ao final da execução do Plano de Testes, obtém-se em '<R>' uma base de dados detalhada, utilizada como entrada do processo de desenvolvimento do Modelo Executivo

baseado em técnicas de AI/ML. Este processo inicia com a validação da consistência e da integridade desta base de dados, reduzindo a possibilidade de vieses e resultados inválidos com técnicas de AI/ML. Em seguida, são realizadas atividades de desenvolvimento (treinamento e validação) do Modelo Executivo, aplicando técnicas de AI/ML sobre a base de dados validada.

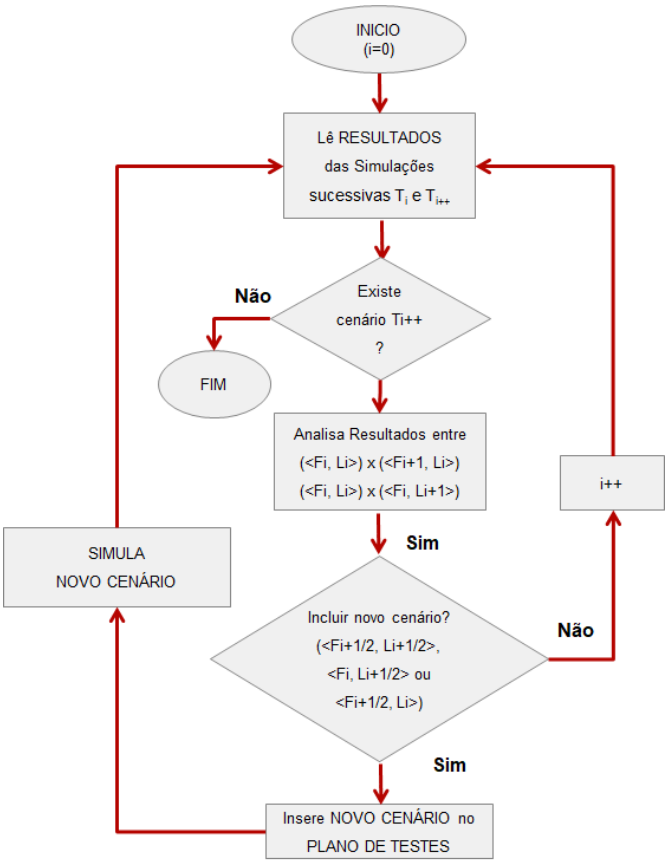


Figura 45 – Algoritmo de automação do processo de simulação

Aprendizado de Máquina (ML) é uma área da Inteligência Artificial que ensina aos computadores a aprender com dados e/ou com a experiência (SARKER, 2022). Algoritmos de ML usam métodos computacionais para aprender informações diretamente de dados, sem a necessidade de seguir equações ou modelos pré-definidos (MATHWORKS, 2023). Além disso, o desempenho dos algoritmos de ML evolui de forma adaptativa e proporcional à quantidade de dados disponíveis para o aprendizado.

Técnicas de ML podem implementar duas abordagens principais de aprendizado: **i. *Aprendizado Supervisionado***, onde um **modelo preditivo** é desenvolvido (treinado) com base em dados de entrada e saída conhecidos; e **ii. *Aprendizado Não Supervisionado***, utilizada para

identificar padrões (ocultos) ou estruturas intrínsecas nos dados de entrada sem que se conheça a relação entre as entradas e saídas (RUSSELL; NORVIG, 2021).

O **Modelo Executivo** da Abordagem Proposta é um **modelo preditivo**, que deve definir a Taxa de Transmissão que o sistema de comunicação deve implementar em função dos valores monitorados (recebido) pelo modelo. Além disso, as relações entre a ENTRADA (Latência, Taxa de Transmissão) e a SAÍDA (nível de risco) das bases de dados de desenvolvimento são conhecidas. Portanto, é ***recomendável*** que o Modelo Executivo da Abordagem Proposta seja desenvolvido por meio de técnicas de **Aprendizado Supervisionado**.

Classificação e **Regressão** são os dois grupos de técnicas de ML relacionados à abordagem de Aprendizado Supervisionado. Técnicas de Classificação predizem respostas discretas, classificando os dados em categorias. Necessitam que os dados de aprendizado possam ser agrupados ou classificados para serem aplicadas no desenvolvimento de modelos preditivos. Por outro lado, **Técnicas de Regressão** predizem respostas contínuas, baseando-se em dados que representem valores quantitativos. A Figura 46, ilustra os principais tipos de técnicas aplicados em aprendizado de máquina, agrupados de acordo com o tipo de aprendizado realizado.

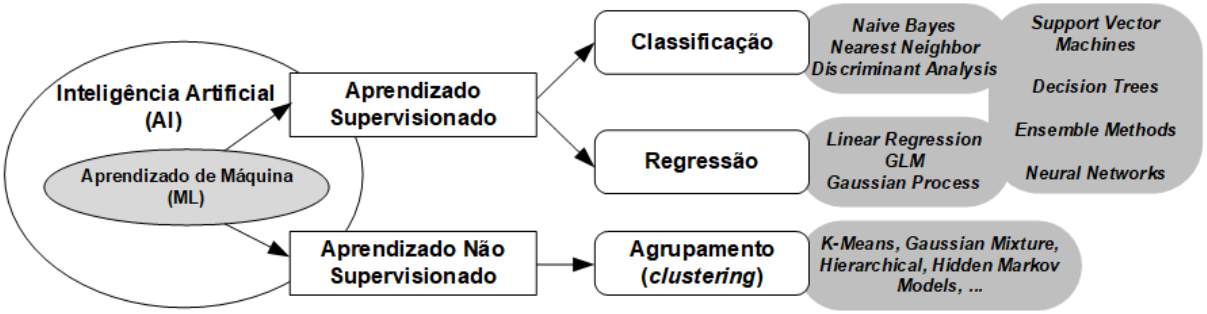


Figura 46 – Tipos e exemplos de técnicas de ML (baseado em MATHWORKS (2023))

A base de dados de desenvolvimento (<R>) precisa ser manipulada conforme a técnica de ML aplicada. Técnicas de **Classificação** exigem que saídas quantitativas (Taxa_Colisão, estatísticas de **dmin()**) sejam agrupadas em classes. Técnicas de **Regressão** não demandam manipulações adicionais. Portanto, a base de dados de desenvolvimento (<R>) é utilizada para gerar três versões de bases de treinamento: ‘2 classes’, ‘3 classes’ e ‘numérica’, onde:

- Utilizando as versões ‘2 classes’ e ‘3 classes’, 19 **técnicas de classificação** geram 19 versões de modelos executivos. Estes modelos são usados para classificar configurações

de desempenho da comunicação (<Latência; Taxa de Atualização>) em classes de risco de segurança (níveis {Alto, Médio, Baixo} – 3 classes – ou {Alto, Baixo} – 2 classes).

- Utilizando a versão 'numérica', 29 **técnicas de regressão** geram 29 versões modelos executivos. Estes modelos são usados para calcular a probabilidade de colisão para uma determinada configuração de desempenho da comunicação (<Latência; Taxa de Atualização>).

O **ANEXO I** apresenta um estudo realizado para avaliar a eficiência das técnicas de classificação e de regressão no desenvolvimento do Modelo Executivo da Abordagem Proposta. A ferramenta computacional Weka[52], versões 3.7 e 3.8, foi utilizada para indução e validação dos modelos. Diversas estratégias de **Dataset Splitting** foram aplicadas no processo treino/teste dos modelos – entre 50/50 a 90/10, e utilizado uma estratégia *k-fold* de validação cruzada, com k igual a 10. Métricas de desempenho foram utilizadas para avaliação dos modelos gerados.

Como principal achado daquele estudo, concluiu-se que as técnicas de **Árvore de Decisão** são a opção mais recomendada para o desenvolvimento do Modelo Executivo da Abordagem Proposta. Além do bom desempenho e consistência dos modelos produzidos, esta categoria de técnica de ML permite compreender e explicar as saídas dos modelos com base nas suas entradas. Essa característica de "***Explicabilidade***" – intrínseca aos modelos gerados por técnicas de Árvore de Decisão – produz modelos "caixa-branca". No processo de garantia de segurança crítica, a utilização de modelos caixa-branca é mandatória para permitir a validação e certificação de segurança de sistemas críticos (BESOLD; UCKELMAN, 2018), (ADADI; BERRADA, 2018).

Para cada técnica de AI/ML aplicada sobre <R> (neste caso, técnicas de Árvore de Decisão), obtêm-se um **Modelo Executivo** com suas **métricas de desempenho**. Os modelos executivos que apresentam os melhores desempenhos são selecionados para avaliação. Vale mencionar que, utilizando técnicas de Árvore de Decisão, os modelos executivos baseados em técnicas de Classificação decidem se um determinado valor de entrada (<Latência$_i$; Taxa de Atualização$_j$>) pode ser considerado 'seguro' ou 'inseguro'. Por outro lado, modelos executivos baseados em técnicas de Regressão estimam a probabilidade de colisão dado um determinado valor de entrada (<Latência$_i$; Taxa de Atualização$_j$>). Ou seja, tem-se:

[52] https://www.cs.waikato.ac.nz/ml/weka/

Classificação: <Latência$_i$; Taxa de Atualização$_j$> ➔ {SEGURO, INSEGURO};

Regressão: Pr(colisão | <Latência$_i$; Taxa de Atualização$_j$>).

Por fim, realiza-se uma **avaliação de efetividade dos Modelos Executivos** na garantia da segurança da aplicação crítica do cenário-base. Para fins de avaliação, o **Modelo Executivo** é incorporado à modelagem do cenário-base por meio de um algoritmo, ilustrado na Figura 47, que emula o processo de execução (operação) da Abordagem Proposta. Este algoritmo é embarcado no módulo de execução de comunicação V2X do OMNeT++ da modelagem do cenário-base.

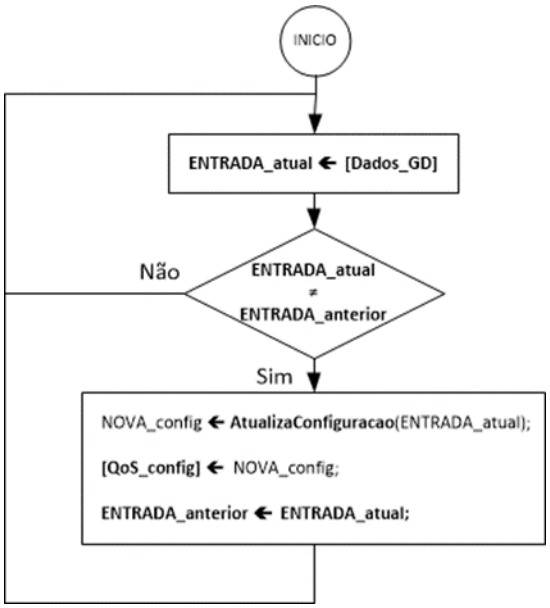


Figura 47 – Algoritmo de emulação da Abordagem Proposta

Enquanto o cenário-base é simulado, este algoritmo executa, em laço infinito, o monitoramento de variáveis na comunicação entre o CCO e os AVs (**[DadosGD]**, fornecidos pelo **Gerador de Dados**, conforme ilustrado na Figura 40). Quando detecta alguma alteração nos valores monitorados (**ENTRADA_atual** ≠ **ENTRADA_anterior**), o algoritmo obtém – por meio da função **AtualizaConfiguracao()** e suportado pela função **Modelo_Executivo()** (Figura 48) – o novo valor de configuração (**[QoS_config]**) a ser utilizada pelo sistema. Neste estudo de caso, o parâmetro "Taxa de Transmissão" é configurado pela Abordagem Proposta.

O Modelo Executivo em avaliação é embarcado em função homônima **(Modelo_Executivo(Latência, Taxa de Transmissão))**. Dado o valor de Latência monitorado pela Abordagem Proposta, esta função identifica o menor valor de Taxa de Transmissão

considerada segura. Uma vez identificada, a nova Taxa de Atualização é retornada ao algoritmo principal (Figura 47) e disponibilizada na saída da implementação da Abordagem Proposta (**[QoS_config]**). Caso a saída do Modelo Executivo esteja acoplado no sistema do cenário-base (Figura 42), o valor presente em **[QoS_config]** é utilizado para atualizar a Taxa de Atualização provida Provedor de Serviços de Comunicação (CSP).

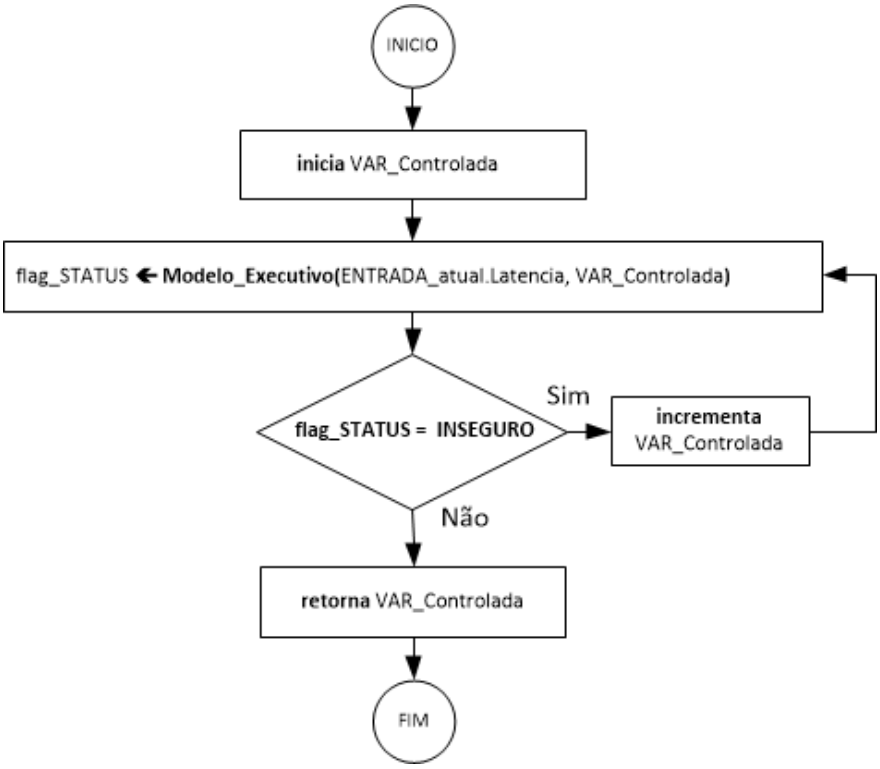


Figura 48 – Algoritmo de AtualizaConfiguracao()

Portanto, para **avaliação da efetividade do Modelo Executivo** na garantia da segurança da aplicação crítica no cenário-base, realiza-se o seguinte procedimento:

1. O **cenário-base é simulado em situações-limite** (valores degradados de latência e/ou Taxa de Transmissão) e **sem** a atuação do Modelo Executivo (**saída desacoplada**). Os resultados obtidos são utilizados como referência (amostra controle) para avaliação de desempenho do Modelo Executivo em avaliação.

2. O **Modelo Executivo** é **acoplado ao cenário-base**, que é simulado nas mesmas situações-limite realizadas anteriormente. Os resultados obtidos são comparados aos valores de referência, avaliando se as métricas de segurança do cenário-base foram impactadas pelo suporte da Abordagem Proposta.

Após evidenciar sua efetividade, realiza-se um **teste de inserção do Modelo Executivo** no cenário-base. Neste teste, a **simulação do cenário-base é iniciada em situações-limite** e **sem a atuação** do Modelo Executivo. Ao decorrer desta simulação, e após observar a perda de capacidade do CCO em manter o fluxo de tráfego seguro (alta taxa de colisões), acopla-se o Modelo Executivo ao cenário-base. A partir do acoplamento do Modelo Executivo, observa-se se o CCO consegue recuperar sua capacidade de controle do tráfego do cruzamento. Uma redução consistente na taxa de colisões evidencia que a Abordagem Proposta é capaz de recuperar os níveis de segurança do sistema.

A próxima seção apresenta os resultados obtidos por meio deste estudo de caso, incluindo o detalhamento do processo de desenvolvimento do Modelo Executivo e a avaliação de seu desempenho no cenário-base.

5. RESULTADOS OBTIDOS E DISCUSSÃO

> "It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts."
> The adventures of Sherlock Holmes
> Sir Arthur Conan Doyle

Na seção anterior, um cenário-base de tráfego rodoviário em Sistemas de Transporte Inteligente Cooperativos (C-ITS) foi desenvolvido e implementado para apoiar o desenvolvimento, análise e demonstração da abordagem proposta nesta pesquisa. Tanto a modelagem do **cenário-base** quando do algoritmo de emulação da **abordagem proposta** foram implementadas por meio da ferramenta computacional anteriormente apresentada. As configurações de desempenho da comunicação V2X, previamente definidas para avaliação nas campanhas de simulação, foram definidas e apresentadas na Tabela 4.

Nesta seção, os resultados obtidos por meio de simulação computacional dos modelos implementados são apresentados. Ao final desta seção, os resultados obtidos são analisados e discutidos em função dos seguintes objetivos:

(i.) Testar a hipótese da garantia das propriedades emergentes no nível da aplicação – neste caso, segurança crítica de tráfego – por meio da compensação funcional entre parâmetros dos elementos no nível de sistema
(ii.) Gerar as bases de dados de desenvolvimento e implementar o **Modelo Executivo** da **Abordagem Proposta**, modelada na seção 4;
(iii.) Avaliar a efetividade da Abordagem Proposta na garantia da segurança da aplicação crítica modelada pelo estudo de caso.

5.1. Resultados obtidos

Para realização das simulações dos cenários-base, faz-se necessário definir o conjunto de **configurações de simulação (CS)** – sobretudo, cenário de tráfego, com o comprimento da RCt e a taxa de entrada de CAVs na Região de Controle (RCt) – que devem ser adotadas nas campanhas de simulação.

Em seguida, é simulado e avaliado o comportamento da aplicação crítica em função de diferentes cenários de desempenho da comunicação V2X, buscando testar a hipótese da

compensação funcional (objetivo ‘i.’). Esta avaliação também permite refinar os parâmetros das CS e de desempenho de comunicação possíveis de serem simulados no contexto deste estudo de caso, bem como as métricas mais adequadas para implementação (treinamento) do Modelo Executivo (objetivo ‘ii.’) e avaliação da abordagem proposta (objetivo ‘iii.’).

5.1.1. Definição das Configurações de Simulação (CS)

Para fins de avaliação, foram definidos os seguintes cenários de **desempenho da comunicação V2X**:

1. Obtenção de dados para cenários atendendo aos requisitos de comunicação, considerando seus valores mais/menos restritivos. Foram adotados:
    a. Cenário mais restritivo (pior caso): **3ms** (Latência) e **100Hz** (Taxa de Transmissão)
    b. Cenário menos restritivo (melhor caso): **100ms** e **10Hz**
2. Obtenção de dados para cenários onde um dos requisitos de comunicação atende ao seu valor menos restritivo de especificação, e o outro requisito está degradado (fora de especificação). Foram adotados:
    a. Latência fim-a-fim de **100ms (**atende à condição **menos restritiva**) e Taxas de Transmissão degradadas de **0,1Hz (1/10 Hz), 0,5Hz (1/2 Hz)** e **1Hz.**
    b. Taxa de Transmissão de **10Hz** (atende à condição **menos restritiva**) e Latências fim-a-fim degradadas de **500ms, 1000ms** e **5000ms.**

Além do **desempenho da comunicação V2X**, é necessário definir as demais **Configuração da Simulação** (**CS**) para as campanhas de simulação. No cenário-base modelado, é necessário configurar as **dimensões físicas das vias 1 e 2**, sobretudo dentro da RCt, dimensões e características dinâmicas dos AVs, localização e dimensões da Região de Colisão (RC), **taxa de ingresso** e **velocidades máximas dos veículos** na RCt. A duração do ciclo de simulação e o passo de simulação utilizado também devem ser definidos.

Foram fixadas as seguintes configurações em todas as campanhas de simulação:

- Dimensões dos veículos (AVs): **4m** de comprimento e **1,8m** de largura.
- Velocidade máxima dos AVs dentro da RCt: **15 m/s** (54 km/h).

As demais configurações de simulação foram definidas por meio da avaliação dos seus efeitos sobre as métricas de risco da aplicação. Assim, por meio de campanhas de simulação, os valores de configuração foram definidas se avaliando como a variação do comprimento das vias na Região de Controle (RCt), da taxa de entrada de AVs na RCt e da duração do ciclo de simulação contribui para a ocorrência de colisões. Foram definidos os seguintes valores de parâmetros:

- Comprimento da Região de Controle (RCt): **100m** e **200m**;
- Taxa de entrada de (o mesmo que o 'intervalo de tempo entre') AVs na RCt: **2s** e **3s.**

Três **campanhas de simulação** foram realizadas, considerando as seguintes configurações para os parâmetros {intervalo de tempo entre AVs', comprimento da RCt}:

1. **Campanha 1: {2s, 200m}**, utilizado como grupo-controle (para comparação);
2. **Campanha 2: {3s, 200m}**, redução da taxa de entrada de entrada de Avs na RCt;
3. **Campanha 3: {2s, 100m}**, redução do comprimento das vias na RCt.

Considerando o objetivo destas simulações, foram adotadas as situações limítrofes e degradadas de desempenho de comunicação V2X. Assim, foram adotados as seguintes combinações de valores para {Latência} e {Taxa de Transmissão}:

a. **{100ms}** x **{1/10Hz; 1/2Hz; 1Hz; 10Hz}**;
b. **{500ms; 1.000ms; 5.000ms}** x **{10Hz}**;

Os resultados das 3 campanhas de simulação são agupadas, em 2 pares, para avaliação. A Tabela 5 ilustra, de forma comparativa, os resultados obtidos pelas campanhas de simulação 1 e 2. A Tabela 5 ilustra, de forma comparativa, os resultados obtidos pelas campanhas de simulação 1 e 3.

Ao comparar os resultados de simulação entre as **campanha 1 e 2** (Tabela 5), é possível inferir os efeitos da **densidade de tráfego na RCt** sobre a segurança da aplicação. Desta forma, Mantido em **200m** o comprimento das vias da RCt, observa-se que o **nível de risco é diretamente proporcional à densidade de tráfego**. Ou seja, o nível de risco reduziu – deixando de apresentar colisões durante o período simulado (1200s) – quando a densidade de tráfego foi **reduzida em 50%** (aumento no intervalo de tempo entre AVs entre as campanhas 1 e 2). De forma análoga, o **nível de risco aumentou** – iniciando a ocorrer colisões na situação

mais degradada de desempenho de **Taxa de Transmissão** na comunicação V2X (100ms, 1/10Hz) – quando a densidade de tráfego aumentou **50%.**

Por hipótese, ao aumentar a densidade de tráfego, reduz-se o espaçamento entre veículos e, consequentemente, aumenta a probabilidade de colisões entre veículos. Assim, é necessário um monitoramento/controle mais frequente dos veículos como forma de gerenciar este risco. Dado que um AV percorre 200m em menos de 15s, serviços de comunicação com taxas de transmissão muito baixas (neste caso, 1/10Hz) não permitem que o CCO realize um ciclo completo de monitoramento/controle de um veículo.

Tabela 5 – Resultados das campanhas de simulação 1 (2s, 200m) e 2 (3s, 200m)

	V2X		Nível de Segurança		Tráfego	
	Latência [ms]	Taxa Tx [Hz]	#Colisões	Acidentes/h	Taxa de entrada de AVs (tempo entre AVs [s])	Comprimento RCt [m]
Campanha 1	100	0,1	38	114	2	200
		0,5	0	0		
		1	0	0		
		10	0	0		
	500	10	0	0		
	1000	10	0	0		
	5000	10	0	0		
Campanha 2	100	0,1	0	0	3	200
		0,5	0	0		
		1	0	0		
		10	0	0		
	500	10	0	0		
	1000	10	0	0		
	5000	10	0	0		

Ao se comparar os resultados de simulação entre as **campanhas 1 e 3** (Tabela 6), é possível inferir os efeitos **do comprimento das vias da RCt** sobre a segurança da aplicação. Desta forma, considerando a maior densidade de tráfego avaliada (intervalo de 2 segundos entre AVs), observa-se que o nível de risco é inversamente proporcional ao comprimento das vias controladas. Ou seja, o nível de risco aumentou – iniciando a ocorrer colisões na situação mais degradada de desempenho de **Latência** na comunicação V2X (5000ms, 10Hz) - quando o comprimento das vias foi reduzido em 50%. De forma análoga, o **nível de risco reduziu** quando o comprimento das vias foi aumentado em 50%.

Tabela 6 – Resultados das campanhas de simulação 1 (2s, 200m) e 3 (2s, 100m)

	V2X		Nível de Segurança		Tráfego	
	Latência [ms]	Taxa Tx [Hz]	#Colisões	Acidentes/h	Taxa de entrada de AVs (tempo entre AVs [s])	Comprimento RCt [m]
Campanha 1	100	0,1	38	114	2	200
		0,5	0	0		
		1	0	0		
		10	0	0		
	500	10	0	0		
	1000	10	0	0		
	5000	10	0	0		
Campanha 3	100	0,1	50	1800	2	100
		0,5	0	0		
		1	0	0		
		10	0	0		
	500	10	0	0		
	1000	10	0	0		
	5000	10	50	1800		

Por hipótese, espera-se que uma via mais longa na RCt permite ao CCO um tempo maior para coordenar os AVs antes deles entrarem na RCz e, consequentemente, reduzindo a quantidade de colisões. Dado que um AV percorre 100m em aproximadamente 7s, serviços de comunicação com latências muito altas (neste caso, 5s) não permitem que o CCO realize um ciclo completo de monitoramento/controle de um veículo.

Com base nos resultados obtidos, adotou-se um **intervalo de tempo de 3s** para a entrada de AVs e **100m de comprimento de via na RCt**. Essa configuração combina uma densidade de tráfego menos restritiva e um comprimento de via na RCt mais restritivo. Dado que a Taxa de Atualização é o parâmetro a ser controlado pelo Modelo Executivo, essa configuração permite avaliar a sensibilidade da Taxa de Atualização sobre diversos parâmetors de Latência.

Durante as campanhas de simulação, foram avaliadas as influências dos **tempos de simulação** sobre os resultados. Para os ciclos de simulação que produziam colisão de todos os pares de AVs, adotou-se um tempo de simulação de **100s**. Nos demais casos, adotou-se tempos de simulação de 1.200s (20 minutos). Devido a esta diferença de tempos de simulação – que produziram quantidades diferentes de pares de AVs simulados, adotou-se a "taxa de colisões por hora" (**Acidentes/h**) ao invés de "percentual de colisões" (**%Colisões**). Contudo, para as campanhas seguintes, a métrica de risco de colisão utilizada será **%Colisões**, com um **tempo de simulação de 3.600s** (1h) fixo todos os ciclos de simulação, produzindo 1.200 pares de veículos neste período.

Por fim, no decorrer destas campanhas de simulação, observou-se que as ferramentas computacionais adotadas (especificamente a OMNET++) não permitiam simular cenários utilizando os parâmetros mais restritivos de comunicação – latência (3ms) e taxa de transmissão (100Hz). Baseado nos tempos de simulação demandados pelos cenários menos restritivos, estimativas otimistas indicaram a necessidade de centenas de horas para concluir os ciclos de simulação mais restritivos utilizando os recursos computacionais disponíveis. Desta forma, ficou estabelecido que cenários com Taxas de Transmissão (F) **maiores que 40Hz** e com Latências Fim-a-Fim (L) **menores que 10ms** não seriam simulados.

Portanto, as **Configuração da Simulação** (**CS**) definidas para as campanhas de simulação deste estudo de caso são:

- Comprimento da Região de Controle (RCt): **100m**;
- Taxa de entrada de (o mesmo que o 'intervalo de tempo entre') AVs na RCt: **3s.**
- Dimensões dos veículos (AVs): **4m** de comprimento e **1,8m** de largura.
- Velocidade máxima dos AVs dentro da RCt: **15 m/s** (54 km/h).
- Tempo de simulação: 3.600s (1h).
- Passo de simulação: 1ms.

Quanto aos cenários de **Desempenho da Comunicação V2X**, estão definidas para as campanhas de simulação deste estudo de caso as seguintes configurações:

1. Obtenção de dados para cenários atendendo aos requisitos de comunicação, considerando seus valores mais/menos restritivos. Foram adotados:

    a. Cenário mais restritivo (pior caso): **10ms** (Latência) e **40Hz** (Taxa de Transmissão)

    b. Cenário menos restritivo (melhor caso): **100ms** e **10Hz**

2. Obtenção de dados para cenários onde um dos requisitos de comunicação atende ao seu valor menos restritivo de especificação, e o outro requisito está degradado (fora de especificação). Foram adotados:

a. Latência fim-a-fim de **100ms,** atendendo à condição **menos restritiva**, e Taxas de Transmissão degradada de **0,1Hz (1/10 Hz), 0,5Hz (1/2 Hz)** e **1Hz.**

b. Taxa de Transmissão de **10Hz,** atendendo à condição **menos restritiva**, e Latências fim-a-fim degradada de **500ms, 1000ms** e **5000ms.**

Concluindo, os parâmetros de Configuração de Simulação (CS) e de Desempenho de Comunicação definidos nesta seção são adotados nas demais simulações realizadas neste estudo de caso.

5.1.2. Comportamento da segurança crítica *vs.* comunicação V2X

Como forma de confirmar a hipótese de que é possível **garantir propriedades emergentes no nível da aplicação** (como segurança de tráfego) por meio da **compensação funcional entre elementos (parâmetros) no nível de sistema**, uma campanha de simulação **(CAMPANHA I.)** foi realizada utilizando os parâmetros definidos anteriormente.

Em cada ciclo de simulação (<Taxa de Transmissão; Latência>), **1.200** pares de AVs trafegaram pela RCz, representando 1h de tráfego rodoviário. Em cada ciclo, foram obtidos o número absoluto de colisões (**#Colisões**), a percentual de colisões (**%Colisões = #Colisões/1.200**), μ[dmin] (**Média(dmin)**)[53] e σ[dmin] (**DP(dmin**))[54]. As Latências simuladas foram 10, 50, 100, 500, 1.000 e 5.000 ms. As Taxas de Transmissão simuladas foram 1/10, 1/2, 1, 10, 20 e 40Hz. A Tabela 7 apresenta os resultados obtidos com a CAMPANHA I de simulação.

Os resultados da Tabela 7 estão tabulados no formato {**Latência [ms], Taxa de Transmissão [Hz]}** ➔ **%Colisões**, conforme tabulação de formato ilustrada na Tabela 8. Os resultados podem ser agrupados em três (3) **cenários** distintos, definido anteriormente nos Planos de Testes:

[53] μ[dmin] = Média(dmin) = $= \frac{1}{N} \sum_{i=0}^{N-1} dmin(i)$, onde 'N' é a quantidade de amostras dmin obtidas no ciclo de simulação.

[54] σ^2[dmin] = (DP(dmin))^2 = $= \frac{1}{n-1} \sum_{i=0}^{n-1} (dmin(i) - \mu[\text{dmin}])^2$.

A. V2X atende às especificações (latência < 100ms; taxa de transmissão ≥ 10Hz)

B. Latência atende a especificação (100ms) e Taxa de Transmissão degradada (< 10Hz)

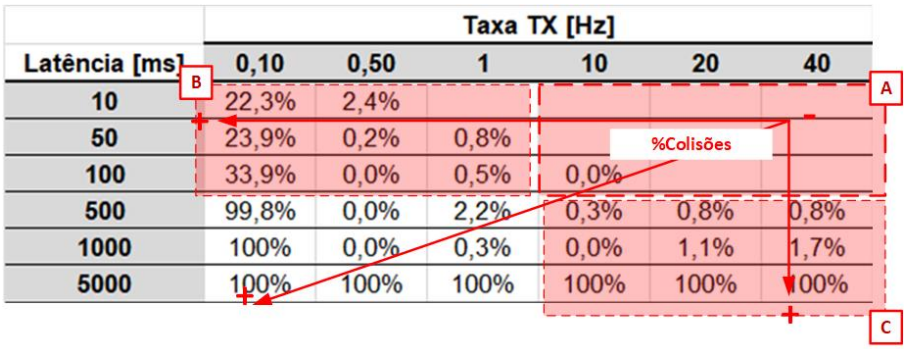
C. Latência degradada (>100ms) e Taxa de Transmissão atende a especificação (10Hz).

Tabela 7 – Resultados da simulação (CAMPANHA I.)

Taxa Tx [Hz]	Latência [ms]	#Colisões	%Colisões	Média(dmin) [m]	DP(dmin) [m]
0,1	10	267	22,26%	1,25	9,42
	50	287	23,90%	12,09	11,50
	100	407	33,89%	0,85	9,66
	500	1199	99,83%	0,00	0,00
	1000	1200	100,00%	0,00	0,00
	5000	1200	100,00%	0,00	0,00
0,5	10	29	2,40%	6,76	6,88
	50	2	0,17%	6,30	6,69
	100	0	0,00%	5,71	6,77
	500	0	0,00%	5,83	7,09
	1000	0	0,00%	8,97	7,31
	5000	1200	100,00%	0,00	0,00
1	50	9	0,77%	6,33	6,91
	100	6	0,50%	5,73	6,81
	500	26	2,17%	6,16	7,16
	1000	3	0,25%	9,02	7,53
	5000	1200	100,00%	0,00	0,00
10	100	0	0,00%	9,43	4,12
	500	3	0,29%	9,33	5,78
	1000	0	0,00%	9,44	4,62
	5000	1200	100,00%	0,00	0,00
20	500	9	0,76%	5,98	6,40
	1000	13	1,05%	5,33	6,95
	5000	1200	100,00%	0,00	0,00
40	500	10	0,81%	5,55	7,42
	1000	20	1,65%	5,09	7,37
	5000	1200	100,00%	0,00	0,00

Tabela 8 – Resultados da CAMPANHA I – Cenários A, B e C

Latência [ms]	Taxa TX [Hz] 0,10	0,50	1	10	20	40
10	22,3%	2,4%				
50	23,9%	0,2%	0,8%			
100	33,9%	0,0%	0,5%	0,0%		
500	99,8%	0,0%	2,2%	0,3%	0,8%	0,8%
1000	100%	0,0%	0,3%	0,0%	1,1%	1,7%
5000	100%	100%	100%	100%	100%	100%



Não foram observadas colisões no Cenário **"A"**[55]. Neste grupo, foi simulado apenas a pior configuração de desempenho da comunicação que ainda atende à especificação (maior Latência, menor Taxa de Transmissão), não sendo considerados outras configurações pertencentes ao envelope de operação. Nos Cenários "**B**" e "**C**", pode-se observar uma **tendência de aumento** na porcentagem de colisões, proporcional à **redução** da Taxa de Transmissão e/ou ao **aumento** da Latência**.**

Com objetivo de reforçar esta inferência, foi realizada uma segunda campanha de simulação (**CAMPANHA II.**). Nesta campanha, foram incorporadas às configurações de desempenho de comunicação não simulados anteriomente. Além disso, para fins de análise, foram medidas as durações de cada ciclo de simulação. A Tabela 9 apresenta os resultados obtidos. Estes resultados estão tabulados no formato **Latência [ms] x Taxa de Transmissão [Hz] ➔ %Colisões**, conforme ilustrada na Tabela 10.

Analogamente à campanha de simulação anterior, os resultados da segunda campanha de simulação são agrupados nos mesmos três (3) **cenários (A, B e C)**. No grupo 'A', são observadas taxas de colisão baixas e uniformes de colisão são observadas (menores que 1%, com exceção do ciclo de simulação de duração mais longa). Nos cenários "**B**" e "**C**", pode-se observar uma **tendência de aumento** na porcentagem de colisões, proporcional à **redução** da Taxa de Transmissão e/ou ao **aumento** da Latência**.**

Em função dos resultados obtidos, considera-se que há bons indícios de que é possível melhorar (ou mesmo manter) a segurança de tráfego (considerando a métrica "%Colisões") por

[55] Neste conjunto de dados, apenas foi simulada a pior configuração de desempenho da comunicação que ainda atende à especificação (maior Latência – 100ms, menor Taxa de Transmissão – 10Hz).

meio da compensação funcional entre parâmetros (antagônicos e independentes) de desempenho no nível do serviço de comunicação V2X.

Tabela 9 – Resultados da simulação (CAMPANHA II.)

Taxa Tx [Hz]	Latência [ms]	#Colisões	%Colisões	Média(dmin) [m]	DP(dmin) [m]	Duração [hh:mm]
0,1	10	267	22,26%	12,29	11,03	00:37
	50	289	24,04%	3,50	11,32	00:38
	100	260	21,64%	12,09	11,32	00:38
	500	1199	99,92%	---	---	---
	1000	1200	100,00%	0,00	0,00	00:32
	5000	1200	100,00%	0,00	0,00	00:28
0,5	10	37	3,06%	8,61	6,25	00:47
	50	0	0,00%	0,00	6,70	00:48
	100	0	0,00%	0,00	0,59	00:48
	500	0	0,00%	---	---	---
	1000	0	0,00%	2,63	7,22	00:45
	5000	1200	100,00%	0,00	0,00	00:28
1	10	86	7,18%	8,316	5,982	00:53
	50	19	1,57%	6,23	6,95	01:19
	100	5	0,38%	5,88	5,32	00:52
	500	26	2,17%	---	---	---
	1000	1	0,08%	2,59	7,74	00:49
	5000	1200	100,00%	0,00	0,00	00:29
10	10	11	0,93%	5,74	7,53	01:05
	50	7	0,59%	5,72	6,84	01:10
	100	10	0,80%	3,09	7,68	01:09
	500	7	0,58%	---	---	---
	1000	0	0,00%	4,73	6,93	01:08
	5000	1200	100,00%	0,00	0,00	00:43
20	10	13	1,05%	5,864	7,486	**05:10**
	50	7	0,59%	1,83	7,687	**05:21**
	100	12	0,97%	5,669	7,619	**05:20**
	500	9	0,75%	---	---	---
	1000	0	0,00%	4,80	5,83	01:30
	5000	1200	100,00%	0,00	0,00	**02:49**
40	10	40	3,32%	5,689	7,102	**10h+**
	50	9	0,71%	1,73	7,735	**10h+**
	100	11	0,88%	5,626	7,584	**10h+**
	500	24	2,00%	---	---	---
	1000	0	0,00%	5,10	7,36	**07:31**
	5000	1200	100,00%	0,00	0,00	**04:10**

Tabela 10 – Resultados da CAMPANHA II – Cenários A, B e C

Latência [ms]	Taxa TX [Hz] 0,10	0,50	1	10	20	40
10	22,3%	3,1%	7,2%	0,9%	1,1%	3,3%
50	24,0%	0,0%	1,6%	0,6%	0,6%	0,7%
100	21,6%	0,0%	0,4%	0,8%	1,0%	0,9%
500	99,9%	0,0%	2,2%	0,6%	0,8%	2,0%
1000	100%	0,0%	0,1%	0,0%	0,0%	0,0%
5000	100%	100%	100%	100%	100%	100%

Com relação às métricas relacionadas às **distâncias mínimas (dmin())**, ao se analisar os resultados de μ[dmin] (Média(dmin)) e σ[dmin] (DP(dmin)), não foi possível identificar uma relação causal {Latência [ms] x Taxa de Transmissão [Hz]} ➔ {μ[dmin], σ[dmin]}. A Figura 49 ilustra a relação entre μ[dmin] e a faixa de Latência (10ms a 5.000ms) para cada uma das 6 configurações (1/10Hz, 1/2Hz e 1Hz na parte superior da figura; 10Hz, 20Hz e 40Hz na sua parte inferior) de Taxa de Transmissão simuladas. Neste caso, esperava-se que, para qualquer valor de Latência, a média da distância mínima aumentaria quanto maior o valor da Taxa de Transmissão.

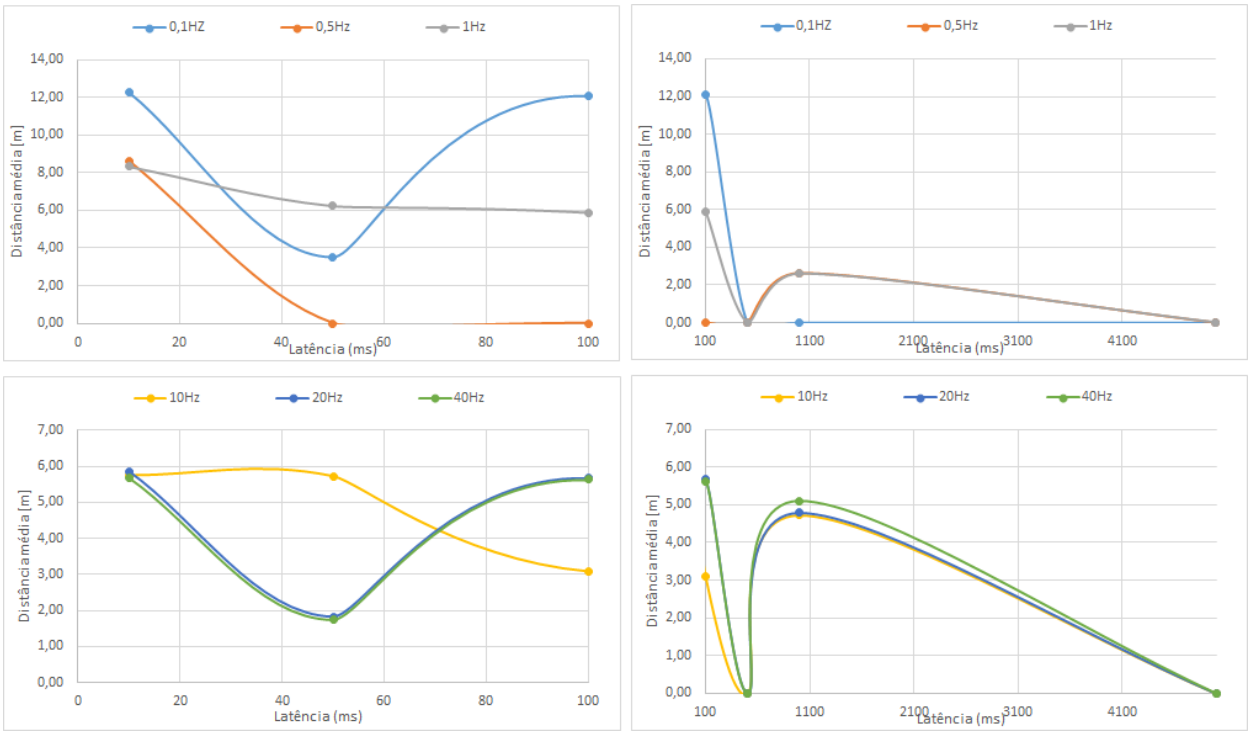


Figura 49 – Resultados da CAMPANHA II (μ[dmin] x Latência | Taxa de Transmissão)

Analogamente à Figura 49, a Figura 50 ilustra a relação entre μ[dmin] e as Taxas de Transmissão (1/10Hz a 40Hz) para cada uma das configurações (10ms, 50ms e 100ms na parte superior da figura; 100ms, 1.000ms e 5.000ms na sua parte inferior) de Latência simuladas. Neste caso, esperava-se que, para qualquer valor de Taxa de Transmissão, a média da distância mínima aumentaria quando o valor da Latência fosse reduzido.

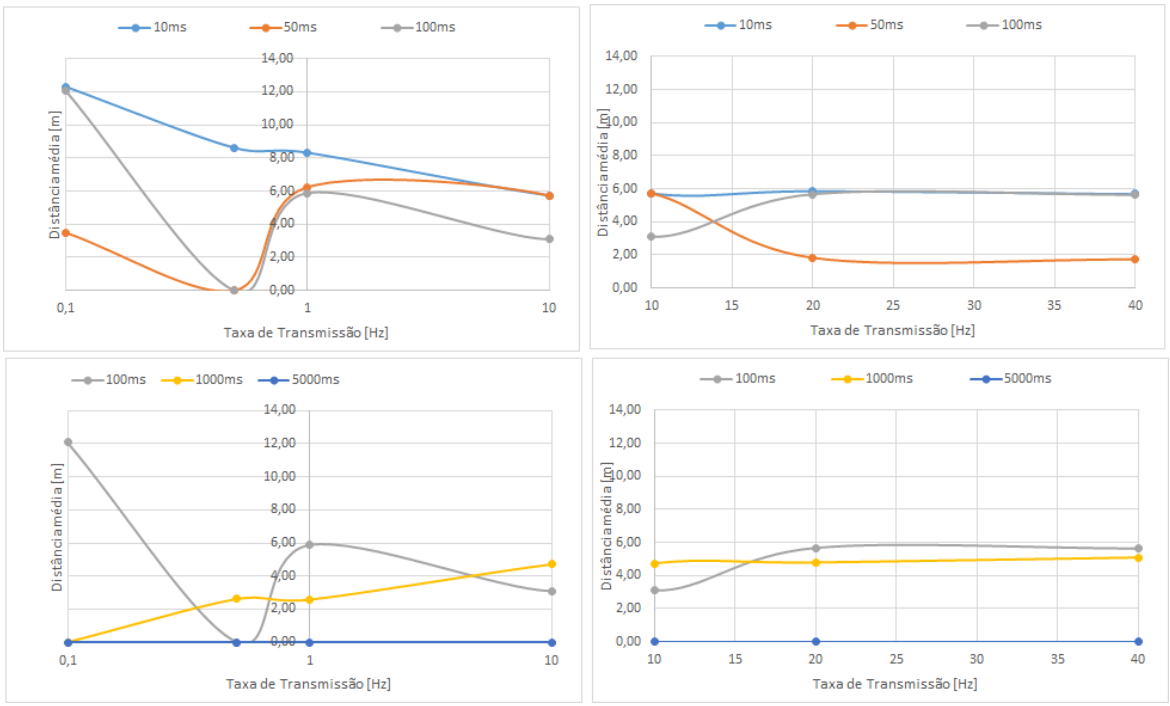


Figura 50 – Resultados da CAMPANHA II (μ[dmin] x Taxa de Transmissão | Latência)

No caso das variâncias das distâncias mínimas, esperava-se que seus valores reduzissem quanto menor o valor de Latência e/ou maior o valor da Taxa de Transmissão. Contudo, σ[dmin] se mantém praticamente constante, independentemente da Taxa de Transmissão (com exceção de valores abaixo de 1Hz) e Latência, conforme pode ser observado na Figura 51.

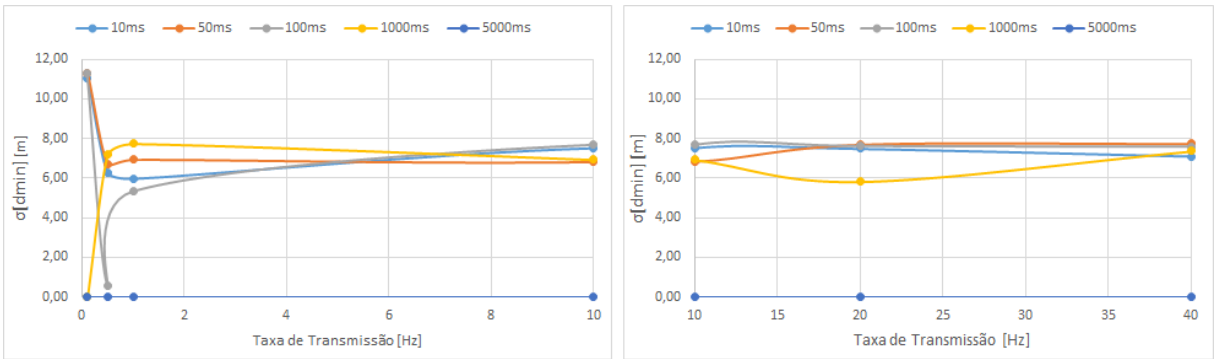


Figura 51 – Resultados da CAMPANHA II (σ[dmin] x Taxa de Transmissão | Latência)

Portanto, dada a falta de relação aparente entre as métricas citadas, optou-se por não utilizá-las para a próxima etapa do estudo de caso, onde são geradas as bases de dados necessárias para o desenvolvimento do **Modelo Executivo** da **Abordagem Proposta.**

Ao analisar as **durações de cada ciclo de simulação**, pode-se observar que são **diretamente** proporcionais a Taxa de Transmissão e **inversamente** proporcionais à Latência. Para Taxas de Transmissão altas (maiores que 10Hz) e Latências baixas (menores que 100ms) – configurações que representam o envelope operacional da comunicação, a duração de cada ciclo de simulação foi superior a 10h. No total, esta campanha de simulação demandou mais de 100h de simulação. E aproximadamente 60% destas horas foram utilizadas nos ciclos simulação com altas taxas de transmissão e latências menores (simulação do envelope operacional).

Conforme ilustrado na Figura 52, a duração dos ciclos de simulação e taxa de transmissão têm relação **linear** quando a latência é mantida constante. Além disso, **taxa de variação** da duração de cada ciclo é inversamente proporcional à latência, estabilizando em valores menores ou iguais a 100ms.

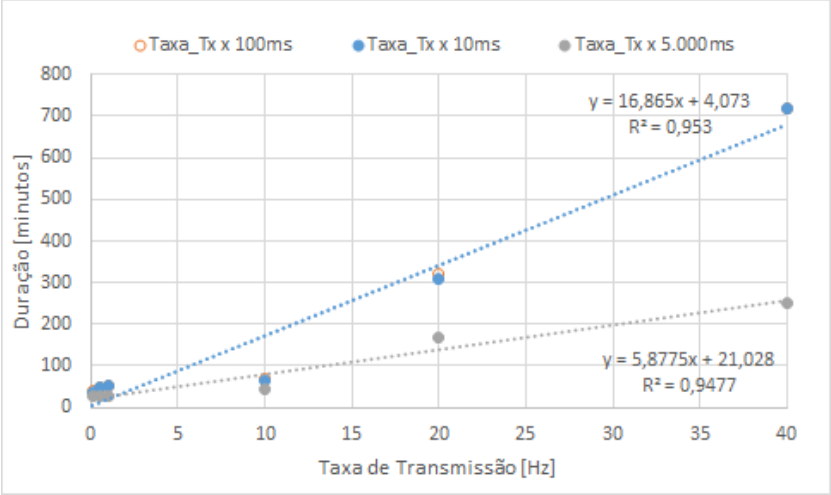


Figura 52 – Resultados da CAMPANHA II (Duração x Taxa de Transmissão)

Observa-se também, conforme ilustrado na Figura 53, que a duração dos ciclos de simulação e a latência têm relação **linear** e **com taxa de variação constante** quando a Taxa de Atualização está degradada e a latência é mantida constante. Para taxas de atualização mais altas (atendendo ao envelope operacional), a duração dos ciclos de simulação e a latência têm relação logaritmica.

Portanto, não considerar nas campanhas de simulação Taxas de Transmissão **maiores que 40Hz** e Latências Fim-a-Fim **menores que 10ms** é justificável, tanto pela demanda computacional elevada para executá-los quanto pelas baixas taxas de colisão observadas, durante as simulações, para valores aderentes as envelope operacional. Mesmo assim, a latência de 1ms foi incluída na campanha de simulação. Contudo, pode-se esperar que os valores obtidos neste ciclo de simulação não sejam confiáveis, dado que o passo de simulação adotado tem o mesmo valor da latência (1ms.)

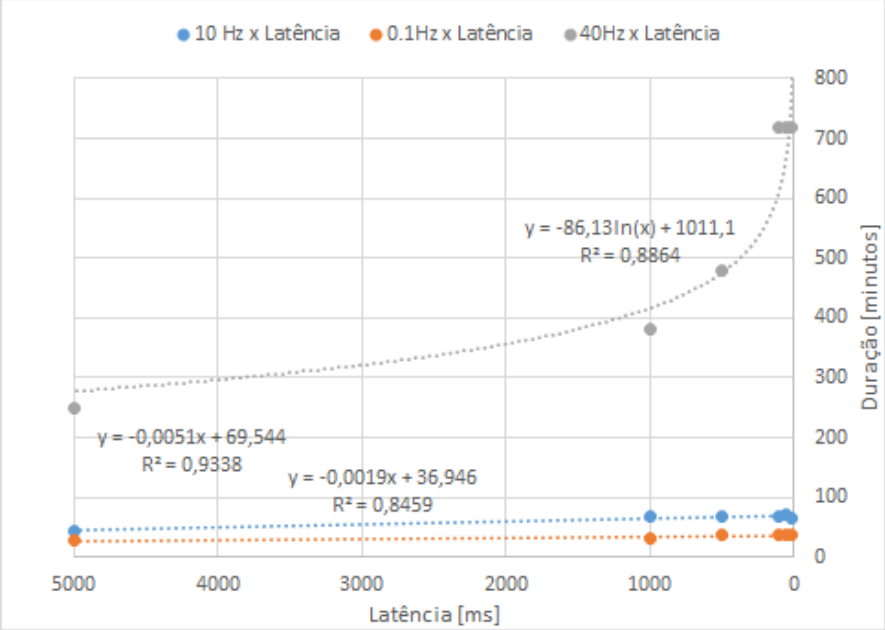


Figura 53 – Resultados da CAMPANHA II (Duração x Latência)

5.1.3. Desenvolvimento do Modelo Executivo

Para a obtenção da **base de dados de desenvolvimento** dos **modelos executivos** da Abordagem Proposta, uma campanha de simulação (**CAMPANHA III.**) foi realizada utilizando tanto os parâmetros de Plano de Testes – configurações de simulação e os cenários de desempenho de comunicação – quanto o algoritmo de automação da simulação definidos anteriormente. Os resultados obtidos são apresentados na Tabela 11.

Tabela 11 – Resultados da simulação (CAMPANHA III.)

	Taxa TX [Hz]					
Latência [ms]	**0,10**	**0,50**	**1**	**10**	**20**	**40**
1	24,75%	14,20%	23,28%	3,79%	0,84%	0,90%
10	22,26%	3,06%	7,18%	0,93%	1,05%	3,32%
20	21,64%	0,50%	1,22%	0,95%	0,65%	0,50%
30	23,14%	0,06%	0,71%	0,45%	0,88%	0,97%
40	24,22%	0,13%	1,27%	0,76%	1,01%	0,54%
50	24,04%	0%	1,57%	0,59%	0,59%	0,71%
100	21,64%	0%	0,38%	0,80%	0,97%	0,88%
200	22,34%	0%	0%	0%	0,93%	0,67%
500	99,80%	0%	2,20%	0,30%	0,80%	2,00%
1000	100%	0%	0,08%	0%	0%	0%
5000	100%	100%	100%	100%	100%	100%

Em relação aos parâmetros de desempenho de comunicação V2X originalmente definidos, observa-se que esta base de dados incluiu cinco (5) novos valores de latência (1ms, 20ms, 30ms, 40ms e 200ms), resultando em trinta (30) novos ciclos de simulação e, consequentemente, 30 resultados adicionais. Estes resultados são apresentados, na íntegra, na Figura 54 (**%Colisões** em função da **Latência**, apresentando uma curva para cada **Taxa de Atualização**) e, na Figura 55, o trecho destacado na Figura 54. Pode-se observar que as curvas atendem ao comportamento previamente avaliado e evidenciado, no qual o risco de colisão pode ser reduzido com a compensação entre parâmetros do sistema (por exemplo, aumentando a taxa de transmissão).

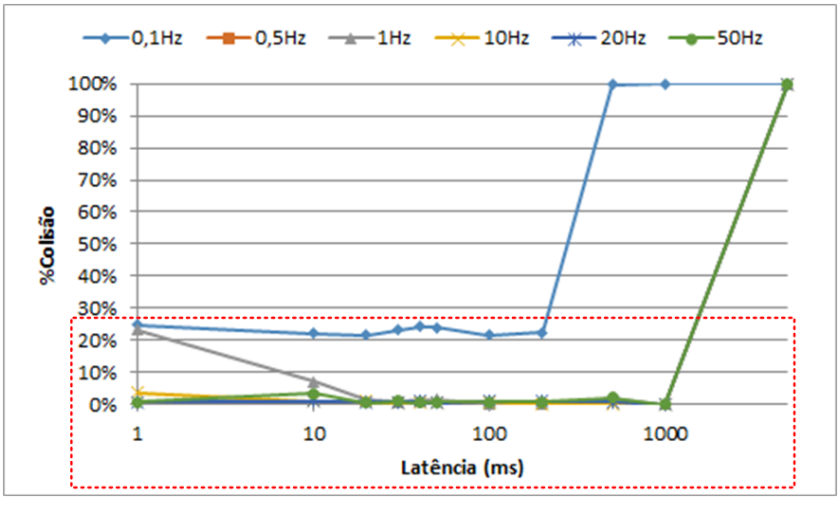


Figura 54 – Resultados da CAMPANHA III (%Colisão x Latência | Taxa de Transmissão)

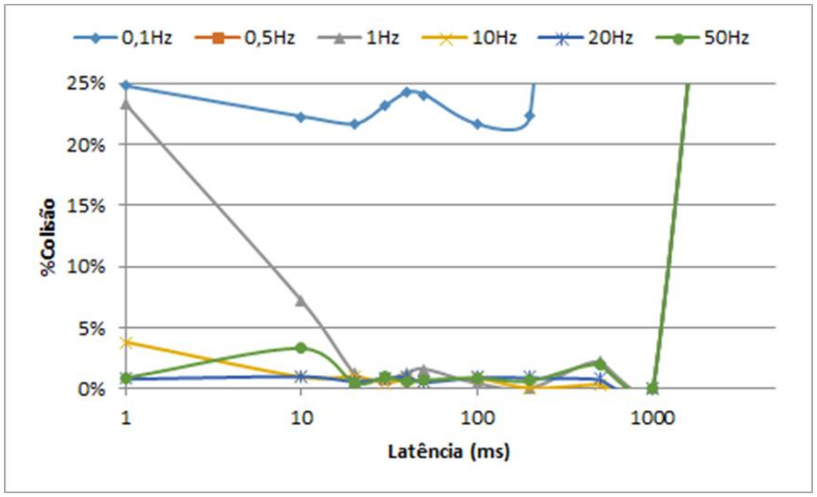


Figura 55 – Resultados da CAMPANHA III (%Colisão x Latência | Taxa de Transmissão) – destaque

A partir desta base de dados, prossegue-se no **processo de desenvolvimento do Modelo Executivo**. Assim, aplicando técnicas de AI/ML sobre a base de dados de desenvolvimento (Tabela 11), os modelos executivos são treinados e validados. Conforme justificado anteriormente, os modelos executivos deste estudo de caso utilizam técnicas de AI/ML baseadas em **Árvore de Decisão**, gerando **modelos de classificação**. Estes modelos executivos decidem se uma configuração de Latência e Taxa de Atualização aplicada em sua entrada é considerado 'SEGURO' (risco baixo) ou 'INSEGURO' (risco alto).

Ao se desenvolver modelos de classificação, é necessário classificar os resultados numéricos (%Colisão) apresentados na base de dados de desenvolvimento (Tabela 11), mapeando cada um dos valores em um grupo pré-definido. Neste estudo de caso, adotou-se modelos de classificação baseados em duas classes: **SEGURO** e **INSEGURO**. O **limiar** entre as duas classes foi identificado de forma experimental, visando minimizar o desbalanceamento naturalmente imposto pela natureza dos experimentos e pela baixa quantidade de resultados apresentados na base de dados de desenvolvimento. O resultado deste processo é apresentado na Tabela 12 e detalhado no **ANEXO I**.

Tabela 12 – Definição dos limiares de classes para aplicação de AI/ML

Classe de risco	Limiar	Quantidade de amostras
SEGURO	0.0093	35
INSEGURO	1	31
Total		66

Os resultados numéricos da base de dados de desenvolvimento (Tabela 11) foram classificados de acordo com as classes definidas para o Modelo Executivo. O resultado desta atividade está apresentado na Tabela 13.

Tabela 13 – Resultados da CAMPANHA III – classificação (**S**eguro; **I**nseguro)

	Taxa TX [Hz]					
Latência [ms]	**0,10**	**0,50**	**1**	**10**	**20**	**40**
1	I	I	I	I	S	S
10	I	I	I	S	I	I
20	I	S	I	I	S	S
30	I	S	S	S	S	S
40	I	S	I	S	I	S
50	I	S	I	S	S	S
100	I	S	S	S	I	S
200	I	S	S	S	S	S
500	I	S	I	S	S	I
1000	I	S	S	S	S	S
5000	I	I	I	I	I	I

Em seguida, técnicas de AI/ML baseadas em **Árvore de Decisão** foram aplicadas sobre a base de dados classificada (Tabela 13). Após treinamento, testes e validação de todos os modelos gerados, dois (2) **modelos executivos** foram escolhidos para a **avaliação de efetividade da abordagem proposta** na garantia da segurança da aplicação crítica no cenário-base.

A Figura 56 ilustra as regras de tomada de decisão induzidas pelo processo de ML/IA para o Modelo #1 (à esquerda) e para o Modelo #2 (à direita). Nota-se que tanto o Modelo #1 quanto o Modelo #2 produziram 9 regiões de decisão. No Modelo #1, estas regiões são delimitadas pelas configurações {0Hz; 0,3Hz; 5,5Hz} x {0ms; 5,5ms; 3.000ms}. No Modelo #2, as regiões são delimitadas pelas combinações de configuração {0Hz; 0,1Hz; 10Hz}x{0ms; 1ms; 1.000ms}.

As regras de tomada de decisão inferidas aos Modelos #1 e #2 foram codificadas e embarcadas na função **Modelo_Executivo(ENTRADA_atual_Latencia, VAR_Controlada)**, representando funcionalmente o Modelo Executivo durante a simulação do cenário-base. Os pseudo-algoritmos dos Modelo #1 e Modelo #2 são apresentados a seguir.

```
/* MODELO #1 (início) */
SE
  (ENTRADA_atual_Latencia < 5,5ms E VAR_Controlada < 5,5Hz) OU
  (5,5ms ≤ ENTRADA_atual_Latencia ≤ 3.000ms E VAR_Controlada<0,3Hz) OU
  (ENTRADA_atual_Latencia ≥ 3.000ms))
      retorna INSEGURO;
SENAO
      retorna SEGURO;
FIM
/* MODELO #1 (fim) */


/* MODELO #2 (início) */
 SE
   (ENTRADA_atual_Latencia ≤ 1ms E 0,1Hz < VAR_Controlada ≤ 10Hz) OU
   (VAR_Controlada ≤ 0,1Hz) OU
   (ENTRADA_atual_Latencia > 1.000ms))
       retorna INSEGURO;
 SENAO
       retorna SEGURO;
 FIM
/* MODELO #2 (fim) */
```

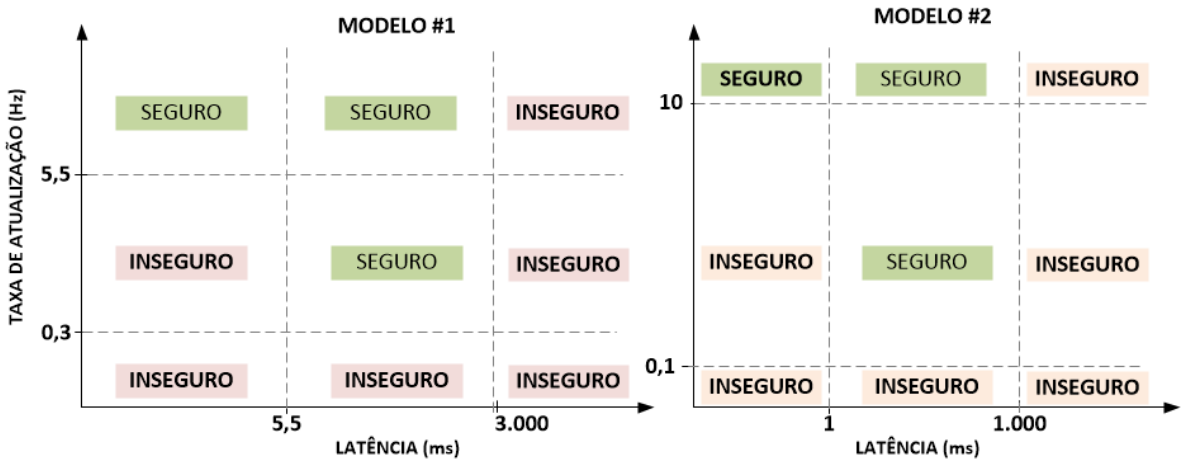


Figura 56 – Regras de decisão induzidas nos Modelos 1 e 2

Para a **avaliação de efetividade dos modelos executivos #1 e #2**, o cenário-base é simulado em condições-limite – neste caso, Latência de 10ms e Taxa de Atualização de 0,25Hz – considerando as situações **(1.) Sem o suporte** da abordagem proposta; **(2.) Com o suporte do Modelo Executivo #1** da abordagem proposta; e **(2.) Com o suporte do Modelo Executivo #2** da abordagem proposta. Os resultados obtidos são apresentados na Tabela 14.

Tabela 14 – Avaliação dos Modelos #1 e #2 (resultados)

SIMULAÇÃO DO CENÁRIO-BASE (Latência = **10ms**; Taxa de Atualização = **1/4 Hz**)					
(1.) Sem suporte		**(2.) Suporte do Modelo #1**		**(3.) Suporte do Modelo #2**	
30	**2,5**	**19**	**1,6**	**30**	**2,5**
Colisões/h	**%Colisões**	**Colisões/h**	**%Colisões**	**Colisões/h**	**%Colisões**

A **situação "1."** é utilizada como referência na avaliação do desempenho do **Modelo #1** e do **Modelo #2**. Desta forma, pode-se observar que o **Modelo #1 apresentou um impacto positivo sobre os riscos de segurança**, promovendo uma redução na taxa de acidentes superior a 1/3 (de 30 para 19 colisões por hora – 36,6 % menos colisões) em relação à situação sem suporte da Abordagem Proposta. No entanto, o **Modelo #2** não apresentou efeito sobre os riscos de segurança, obtendo um desempenho semelhante à simulação do cenário-base sem suporte da Abordagem Proposta ("1.").

5.1.4. Verificação da efetividade do Modelo Executivo no cenário-base

Na sequência, são realizados **testes de inserção do Modelo #1** no cenário-base. A **simulação do cenário-base é iniciada em situações-limite** e **sem a atuação** do Modelo Executivo. Após **180 segundos** desde o início da simulação, acopla-se o Modelo Executivo ao cenário-base. Após acoplamento, dá-se prosseguimento na simulação, finalizando em **520 segundos**.

Foram simuladas três (3) situações-limite:

A. **Latência** de **1.000ms; Taxa de Atualização** de **1/5Hz**;
B. **Latência** de **2.000ms; Taxa de Atualização** de **1/5Hz**;
C. **Latência** de **2.500ms; Taxa de Atualização** de **1/5Hz**;

Os resultados dos testes são apresentados na Figura 57 **(situação-limite A)**, Figura 58 **(situação-limite B)** e Figura 59 **(situação-limite C)**. São apresentadas as distâncias mínimas obtidas entre pares de veículos (**dmin(i,j)**) ao longo do tempo, o instante de acoplamento do **Modelo Executivo** ao cenário-base (**status** é alterado de 0 para 1), a **média** e o **desvio-padrão** (DP) de **dmin(i,j)** para cada um dos dois segmentos da simulação (**status 0 e 1** de acoplamento

do Modelo Executivo). A Tabela 15 compila os resultados dos testes de inserção apresentados graficamente.

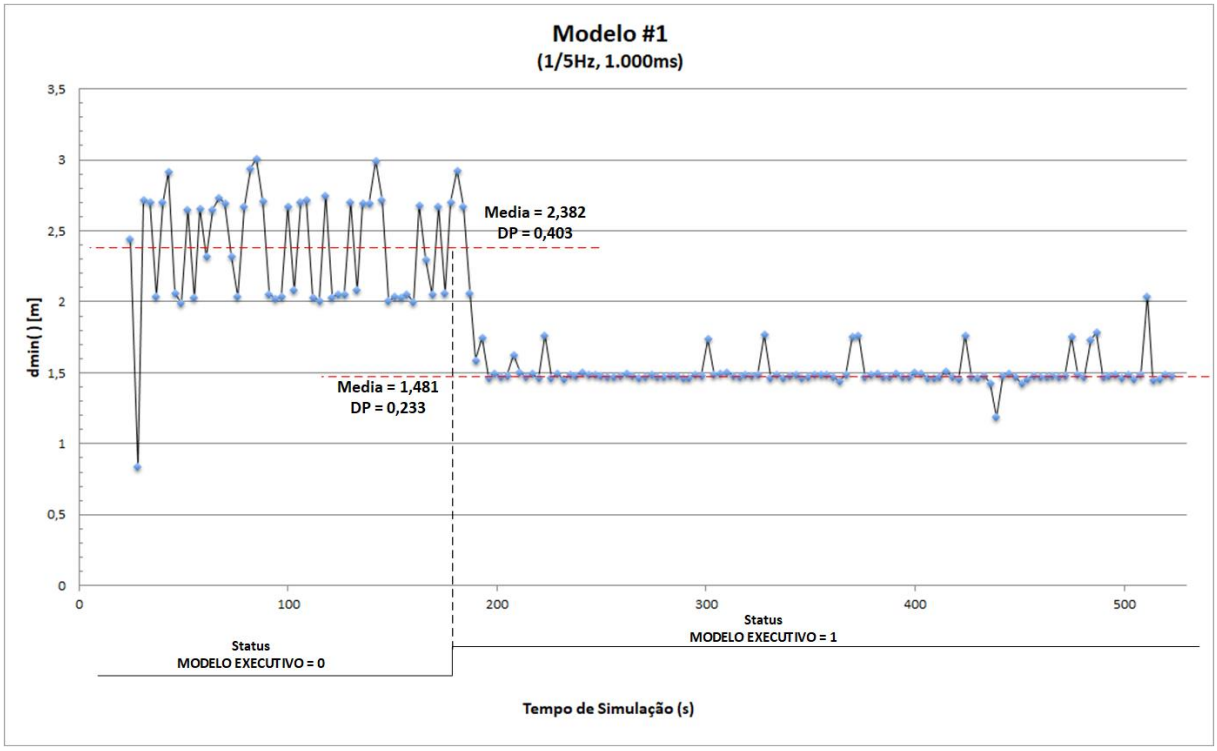


Figura 57 – Teste de inserção do Modelo #1, A (1.000ms; 1/5Hz)

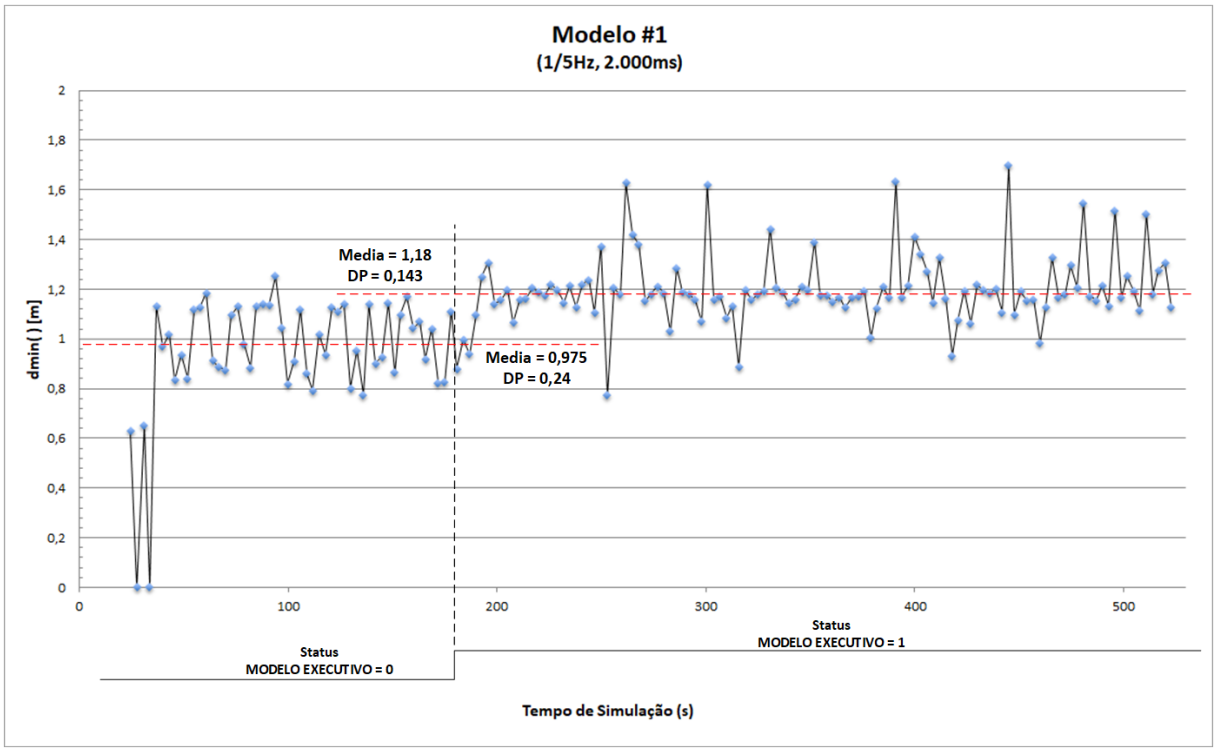


Figura 58 – Teste de inserção do Modelo #1, B (2.000ms; 1/5Hz)

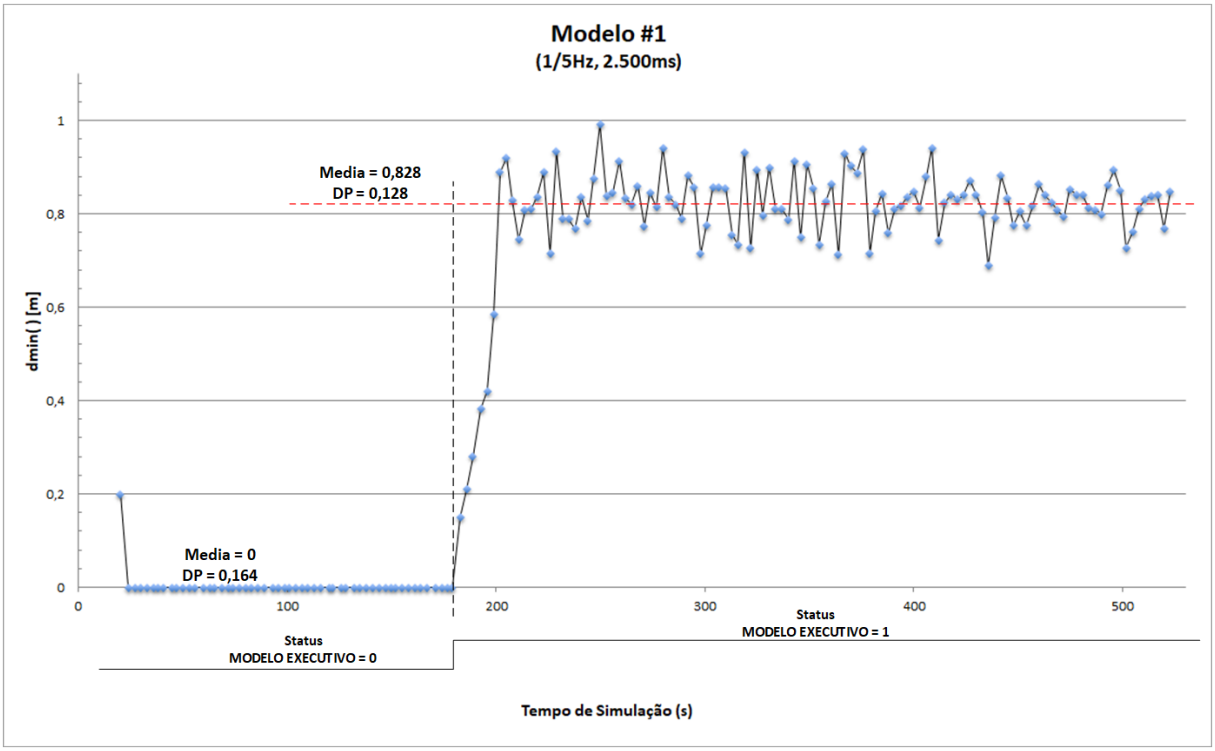


Figura 59 – Teste de inserção do Modelo #1, C (2.500ms; 1/5Hz)

Tabela 15 – Resultados dos testes de inserção, Modelo #1

Status (tempo)	**métricas**	**Taxa de Atualização: 1/5Hz**		
		Latência		
		1.000ms	**2.000ms**	**2.500ms**
0 (180s)	**#colisões**	0	1	51
	(Media; DP) [m]	(**2,382**; 0,403)	(**0,975**; 0,24)	(**0**; 0,164)
1 (342s)	**#colisões**	0	0	2
	(Media; DP) [m]	(**1,481**; 0,233)	(**1,18**; 0,143)	(**0,828**; 0,128)
Tempo de reação após acoplamento (aprox.) [s]		16	25	28

No primeiro teste de inserção (**A**), não foram observadas colisões (#colisões) ao longo de todo o período de simulação. No segundo teste (**B**), observou-se uma colisão no início da simulação, enquanto o Modelo Executivo não estava acoplado. Após acoplamento, não foram observadas colisões. No terceiro teste (**C**), praticamente todos os veículos colidiram no cruzamento enquanto o Modelo Executivo não estava acoplado. Após acoplamento, foram observadas

colisões imediatamente após o acoplamento do Modelo Executivo. Contudo, nenhuma colisão foi observada depois de decorrido o **tempo de reação após acoplamento**.

Tempo de reação após acoplamento é o intervalo entre o acoplamento do Modelo Executivo no cenário-base e o instante no qual o modelo atingiu sua efetividade máxima sobre segurança crítica do cenário-base. Essa efetividade é quantificada pela **média da dmin()** que o Modelo Executivo consegue garantir em um determinado cenário. Vale frisar que, considerando a arquitetura da Abordagem Proposta implementada, o Modelo Executivo monitora continuamente a aplicação. Assim, quando acoplado, o modelo possui uma saída definida, que é configurada no CSP (serviço de comunicação).

Em todos os testes de inserção (A, B e C), ocorreu uma alteração significativa no comportamento das dmin() após o acoplamento do Modelo Executivo, mensurada tanto pelo **valor médio** quanto pelo **desvio-padrão** da **dmin()**. Com relação ao valor do **desvio-padrão** das **dmin()**, observa-se sua **redução** significativa em **todos os testes** (A, B e C) após o acoplamento do Modelo Executivo. Isso é resultado do aumento da capacidade do CCO em manter a uniformidade do fluxo de tráfego dos AVs, obtido com a melhoria do desempenho dos serviços de comunicação.

Com relação ao valor da **média** das **dmin()** após o acoplamento do Modelo Executivo no cenário-base, observam-se duas situações:

- O **aumento** da média de dmin() nos **testes "B"** e **"C"**. Isso é resultado do aumento da capacidade do CCO em manter o distanciamento seguro entre AVs, obtido com a melhoria do desempenho dos serviços de comunicação.

- A **redução** da média de dmin() no **teste "A"**. Diferentemente do **teste B**, o **teste A** não observou colisões. A **média das dmin()** aplicada entre AVs pelo CCO se demonstrou **maior do que a necessária** para manter a segurança do tráfego no cruzamento. Assim, em conjunto com uma menor dispersão das dmin() – redução do DP, pode-se reduzir o distanciamento entre AVs sem impactar a segurança do tráfego. Isso é resultado do aumento da capacidade do CCO em manter a uniformidade do distanciamento entre AVs, obtido com a melhoria do desempenho dos serviços de comunicação.

5.2. Discussão dos resultados

A missão planejada para o conceito da Abordagem Proposta é garantir a segurança da **aplicação** por meio da **compensação funcional** entre **elementos do sistema** em função de **eventos perigosos** (nível de risco de segurança) observados na **aplicação** durante a operação do sistema. A dinâmica de compensação funcional é regida pela **relação entre as características antagônicas dos elementos do sistema e o nível de risco observado na aplicação**.

Para desempenhar esta missão, obteve-se uma Abordagem Proposta formada por **4 atividades**, conforme apresentado na seção 3.3: **(1.) Inferir** comportamento dos parâmetros dos elementos do sistema em relação à métricas de segurança da aplicação (neste estudo, a Taxa de Colisão (CR)); **(2.) Identificar** a tupla de parâmetros (neste estudo, dupla {Latência, Taxa de Transmissão}) não dependentes entre si, com comportamento antagônico em relação às métricas de risco de segurança observadas na aplicação e, ao menos, com um dos parâmetros controlável (neste estudo, Taxa de Transmissão); **(3.) Definir** as funções de transferência (regras de decisão) entre tuplas de parâmetros e métricas de risco de segurança (neste estudo, {(Latencia, Taxa de Transmissão) ➔ CR} e Taxa de Transmissão = g(Latencia, CR)); e **(4.) Compensar** variações nos parâmetros monitorados de forma a manter a segurança no nível da aplicação, reconfigurando os parâmetros controláveis segundo as regras definidas na atividade **3**.

A Abordagem Proposta foi implementada e avaliada com base no cenário-base desenvolvido neste estudo de caso, conforme ilustrado na Figura 39. Neste estudo de caso, a abordagem proposta monitora um CSP de propósito geral que provê serviços de comunicação V2X entre os veículos autônomos colaborativos (CAV) – trafegando em um cenário de cruzamento de tráfego sem semáforo físico – e um serviço de semáforo virtual (CATraCA) provido pelo CCO – aplicação crítica. Durante sua operação, a Abordagem Proposta (Figura 40) monitora, de forma direta, a qualidade de serviço (QoS) fornecida pelo CSP (Latência e Taxa de Transmissão) e as métricas relacionadas ao risco de colisão de veículos no cruzamento, no caso a dmin, obtida indiretamente por meio de dados posicionais transmitidos pelos CAV e transportados pelo CSP. Com base nestas entradas, o Modelo Executivo da Abordagem Proposta **reconfigura o parâmetro de Taxa de Transmissão do CSP** de forma a manter a aplicação crítica (cruzamento de via) com um nível aceitável de risco de colisão.

O desenvolvimento das atividades 1 e 2 da Abordagem Proposta está apresentado na seção 5.1.2. O objetivo daquela seção foi **confirmar a hipótese** de garantir propriedades emergentes no nível da aplicação (segurança crítica) por meio da compensação funcional entre elementos no nível de sistema. Então, foram realizadas campanhas de simulação computacional do modelo do cenário-base, obtendo-se conjuntos de dados cuja estrutura está representada na Figura 37. Sobre os conjuntos de dados obtidos, inferiu-se o comportamento dos parâmetros "Latência" e "Taxa de Transmissão" do elemento de comunicação (CSP) em relação às métricas de risco de segurança (Taxa de Colisão e estatísticas de posição e dispersão de dmin) – **atividade 1**. Consequentemente, esta dupla de parâmetros foi identificada como viável à aplicação no processo de compensação funcional.

Apenas parâmetros do elemento de comunicação – neste caso, um CSP provendo comunicação V2X – foram considerados para as **atividades 1 e 2** da Abordagem Proposta. Porém, por princípio, outros parâmetros do sistema poderiam ser aplicados à inferência realizada pela **atividade 1** para, em seguida, serem considerados na tarefa de identificação de tuplas de parâmetros da **atividade 2**. Contudo, conforme discutido na seção 3.3, a Abordagem Proposta deve utilizar os dados trafegados pelo elemento de comunicação (CSP) de forma a inferir o comportamento da relação entre os parâmetros dos elementos do sistema e do nível de risco de segurança da aplicação (**atividade 1**). E, dado que deve ser monitorado, os parâmetros do elemento comunicação também podem ser considerados nas atividades 1 e 2. Assim, optou-se por lidar apenas com parâmetros de comunicação, sobretudo devido **ao estudo de caso se propor a avaliar o impacto do desempenho da comunicação V2X sobre a segurança do tráfego no contexto C-ITS**.

O desenvolvimento da **atividade 3** está apresentado na seção 5.1.3. O Modelo Executivo instancia a função de transferência (regras de decisão) entre *tuplas* de parâmetros dos elementos do sistema (Latência e Taxa de Transmissão) e as métricas de risco de segurança da aplicação (%Colisão, Média(dmin), DP(dmin)). Neste estudo de caso, estas regras de decisão foram definidas por meio de um processo de desenvolvimento baseado em técnicas de AI/ML (seção 4.2.2), utilizando resultados da campanha III de simulação (Tabela 11) e gerando dois modelos de classificação baseados em árvores de decisão (Figura 56). Ao final, estas regras de decisão foram convertidas em algoritmos e embarcadas no modelo do cenário-base de forma a emular a Abordagem Proposta realizando a **atividade 4** (Figura 47 e Figura 48).

Devido à falta de relação aparente observada entre as estatísticas de dmin e a dupla (Latência, Taxa de Transmissão) nos conjuntos de dados obtidos para inferência dos modelos (seção 5.1.2), optou-se por não utilizar **dmin** na **atividade 3** da Abordagem Proposta. Consequentemente, métricas de risco monitoradas em tempo de execução do modelo (representando a operação do sistema) não foram utilizadas nas tomadas de decisão realizadas na **atividade 4**. Portanto, o Modelo Executivo foi definido no formato:

(Latência, Taxa de Transmissão) ➔ {SEGURO, INSEGURO}

Taxa_Transmissão = g(Latência,SEGURO), onde:

SEGURO => (%Colisão ≤ *Limiar*) e "*Limiar*" (≈ 0,1%), definida na Tabela 12.

Ao executar a **atividade 4**, o Modelo Executivo obtido na **atividade 3** avalia continuamente os valores monitorados durante a operação (Latência, Taxa de Transmissão). Caso identifique um condição insegura – ou seja, o par (Latência, Taxa de Transmissão) ➔ INSEGURO, um novo valor de "Taxa de Transmissão" é reconfigurado segundo a relação g(Latência, SEGURO), buscando pelo melhor **desempenho** da Taxa de Transmissão (que poderia ser representado pela alocação adicional de recursos) que possa levar o sistema a um nível aceitável de risco (%Colisão ≈ *Limiar*), mitigando a condição insegura.

Este comportamento é confirmado pela avaliação do Modelo Executivo #1, cujos resultados estão apresentados na seção 5.1.4. Durante a **atividade 4**, o Modelo Executivo #1 foi capaz de identificar condições inseguras – representadas pelas situações-limite impostas ao desempenho da comunicação – e definir um valor mínimo de desempenho para a Taxa de Transmissão considerando a Latência observada e o nível de risco aceitável (g(Latência,SEGURO)). Ao reconfigurar o valor da Taxa de Transmissão, observou-se a resolução da condição insegura por meio da métrica **dmin** (distância mínima observada entre veículos), com a melhoria da uniformidade do tráfego (redução da dispersão de dmin) e readequação da distância segura entre veículos (média de dmin).

Quanto à readequação da distância segura, observou-se um aumento na média de dmin nas situações mais críticas de degradação de desempenho de comunicão, onde foram observadas maiores taxas de eventos inseguros e acidentes (situações-limite 'B.' e 'C.'). Em contrapartida, observou-se uma redução na média de dmin na situação-limite menos crítica de

degradação de desempenho de comunicão (‘situação-limite ‘A.’). Nas três situações-limite, não foram observadas condições inseguras após a reconfiguração do elemento de comunicação.

Conforme discutido na seção 3.3, a implementação da Abordagem Proposta neste estudo de caso ficou limitada ao processo de **garantia de segurança da aplicação crítica**, pois as métricas de segurança monitoradas não foram utilizadas durante a operação no processo de definição da Taxa de Transmissão (**atividade 4**) da comunicação do sistema. Caso fossem utilizadas, a Abordagem Proposta também possibilitaria otimizar recursos do sistema orientado aos riscos da aplicação. Consequentemente, seria possível **reduzir o desempenho** da Taxa de Transmissão quando observado um nível de risco maior que o aceitável (%Colisão << *Limiar*), representando uma liberação de recursos de comunicação enquanto **mantém a aplicação em uma condição segura**.

Conforme discutido na seção 4, a comunicação não é considerada um elemento crítico em sistemas críticos em segurança suportados por comunicação. Nestes sistemas críticos em segurança, **os usuários finais (EU)** – neste estudo de caso, CCO e CAV – **são os elementos responsáveis por lidar, de forma segura, com as degradações do sistema de comunicação**. Assim, segundo a abordagem tradicional (*top-down)* de Engenharia de Segurança de Sistemas (SSE), as condições e situações inseguras previstas durante a operação do sistema devem ser identificadas durante seu projeto. Então, **requisitos de segurança** são especificados de forma a garantir níveis aceitáveis de riscos de segurança relacionados às situações inseguras identificadas. Em seguida, estes requisitos são detalhados e mapeados sobre os elementos críticos do sistema. Neste processo, **requisitos mínimos de desempenho de comunicação** são definidos para as funcionalidades críticas baseadas em comunicação. Estes requisitos devem ser atendidos para que o sistema não seja exposto a uma condição insegura resultante da degradação dos serviços de comunicação. Portanto, estes requisitos são utilizados na implementação de **mecanismos de detecção e reação à perda de desempenho mínimo da comunicação** pelos **elementos que desempenham estas funcionalidades críticas** em segurança.

Os requisitos mínimos de desempenho de comunicação adotados neste estudo de caso foram definidos com base em normas de serviços de comunicação utilizando tecnologias 4G-LTE e 5G, conforme apresentado na seção 4.1.1. Por serem definidos para categorias gerais de aplicação – e, a princípio, não orientados por uma análise de risco – estes requisitos devem

atender a uma ampla quantidade de cenários de serviços baseados em comunicação V2X. Portanto, entende-se que estes **requisitos são conservadores** (mais restritivos do que um mínimo necessário). **Consequentemente, caso haja meios de monitorar e gerenciar os riscos de segurança durante a operação (neste caso, a Abordagem Proposta), seria possível tornar os requisitos mínimos de desempenho de comunicação menos restritivos para aplicações específicas**, **ampliando o envelope operacional** do sistema sem comprometer sua segurança.

Neste estudo de caso, adotou-se que os EU (sobretudo os CAV[56]) não detectam problemas de comunicação (bem como outras situações potencialmente inseguras) e não mitigam situações perigosas, expondo o sistema a condições inseguras. Desta forma, foi possível identificar os níveis mínimos de desempenho de comunicação abaixo dos quais a aplicação crítica (cenário-base) se comporta de forma insegura por influência direta da comunicação. Caso contrário, os valores identificados estariam limitados pelas situações especificadas nos mecanismos embarcados nos CAV, limitados pelos processos de SSE que as definiram.

A Figura 60 apresenta os valores de requisitos de desempenho mínimo de comunicação obtidos com a modelagem e simulação do cenário-base e a implementação da Abordagem Proposta, comparando-os com o envelope operacional especificado para esta aplicação (baseado em normas ETSI para comunicação V2X utilizando 4G-LTE e 5G). **Conforme esperado, a Abordagem Proposta proporcionou uma ampliação significativa do envelope operacional dos requisitos de comunicação para o cenário-base, permitindo a operação do sistema mesmo em situações nas quais o CSP não atenda a um SLA normativo.**

Conforme mencionado anteriormente, esta ampliação do envelope operacional seria possível devido ao monitoramento e gestão de riscos de segurança proporcionado pela Abordagem Proposta durante a operação do sistema. Contudo, o Modelo Executivo implementado não utiliza métricas de risco em tempo de execução para tomada de decisão. Além disso, considerando o processo de desenvolvimento adotado (Figura 40), um Modelo Executivo em operação está atualizado com as situações inseguras observadas até o momento

[56] Como principal elemento crítico em segurança em um sistema de tráfego, os veículos (neste caso, os CAV) deveriam **lidar com todas as situações de perigo localmente identificadas pelo veículo** (como perda de desempenho da comunicação, falhas em sistemas do veículo, obstáculo inesperado na via, entre outros), levando o veículo para um estado seguro por meio de mecanismos de Segurança Ativa (como frenagem de emergência) ou reduzindo a severidade de um acidente por meio de mecanismos de Segurança Passiva.

em que foi colocado em operação. No caso da ocorrência de situações inseguras não previstas previamente, é necessário atualizar o Modelo Executivo de forma a manter os riscos de segurança gerenciados.

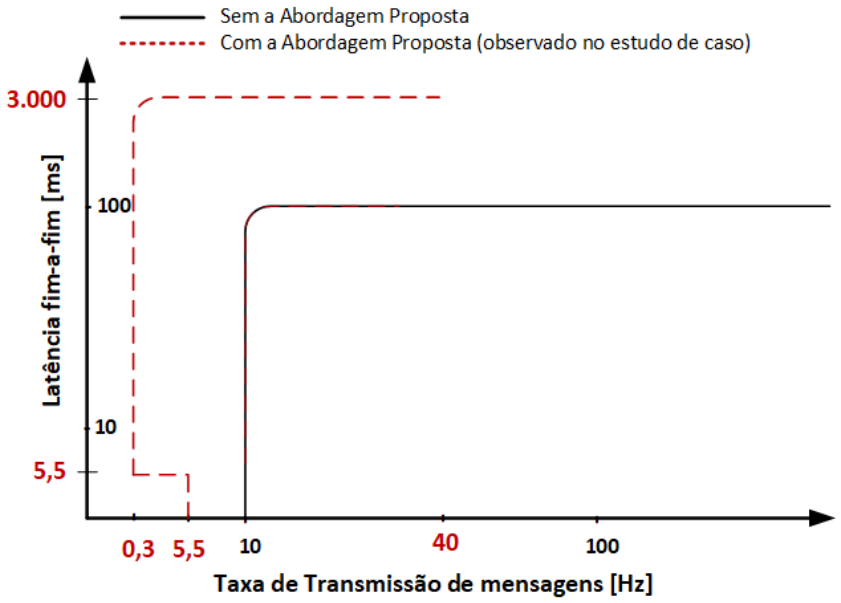


Figura 60 – Ampliação observada para o envelope operacional do cenário-base

Porém, o processo de desenvolvimento do Modelo Executivo recebe continuamente os dados monitorado do sistema e disponibilizados pelo Gerado de Dados (Figura 43). Desta forma, o Modelo Executivo pode **se manter atualizado com novas situações inseguras que venham a ocorrer na operação do sistema**, mas, não tenham sido previstas no projeto. Com isso, a **limitação em se prever situações inseguras em CoES é reduzida**. Contudo, a melhor forma de realizar este processo de atualização do Modelo Executivo não está no escopo deste trabaho.

Neste estudo de caso, vale frisar que a relação entre as características do sistema e o nível de risco da aplicação **(atividade 3)** foi identificada por meio de técnicas de AI/ML e de modelagem e simulação computacional de um sistema de tráfego no contexto do C-ITS. Técnicas de AI/ML foram adotadas neste estudo devido à sua capacidade em reconhecer padrões de comportamento em bases de dados obtidas com o monitoramento do cenário-base. Esta heurística produziu o melhor Modelo Executivo possível, considerando o conjunto de dado obtido segundo as premissas de modelagem adotadas.

Contudo, toda técnica e modelo possui limitações intrínsecas, que podem justificar alguns **resultados obtidos e não esperados.** Por exemplo, com relação a ampliação do

envelope operacional obtido (Figura 60), não seria esperado que a região {Taxa de Transmissão (Hz), Latência fim-a-fim (ms)} = {[0,3; 5,5], [0; 5,5]} fosse excluída das configurações operacionais do sistema, sobretudo devido à configuração na região {[0,3; 5,5], [5,5; 3.000]} ter sido considerada aceitável. Contudo, esta situação pode ter ocorrido por limitações da técnica de AI/ML, que induziram regras de decisão (Figura 56) baseada em um volume pequeno de dados (Tabela 11) considerando o tipo de técnica aplicada para indução do modelo. Além disso, dado que este tipo de técnica demanda por uma base balanceada de treinamento, o **valor de limiar de segurança (*Limiar*) adotada para a métrica de risco de colisão – %Colisão** (Tabela 12) – foi diretamente influenciado pela quantidade de dados disponíveis, gerando anomalias na base de treinamento (Tabela 13).

É importante ressaltar, porém, que o desenvolvimento do Modelo Executivo da implementação da Abordagem Proposta não fica restrito às técnicas de AI/ML. Portanto, outras abordagens de desenvolvimento poderiam ser aplicadas sem comprometimento dos conceitos principais apresentados neste trabalho. Contudo, dado que a Abordagem Proposta deve identificar e tomar decisões baseada em novas situações, a capacidade de aprendizado é uma característica necessária ao desenvolvimento do Modelo Executivo (atividade 3). **Portanto, as limitações técnicas relacionadas à obtenção de bases de dados de treinamento para as abordagens de AI/ML podem ser contornadas com a adoção de técnicas e, sobretudo, recursos de maior poder computacional no processo de simulação dos modelos CoES.**

Ainda com relação **aos resultados obtidos e não esperados**, observou-se que as taxas de colisão obtidas para Latência fim-a-fim de **1ms** distorceram o comportamento antagônico esperado para as relações {*Taxa de Transmissão*, *Latência*} vs *%Colisão*. Devido ao comportamento de compensação funcional, esperavam-se Taxas de Colisão menores que as obtidas para a Latência fim-a-fim de 10 ms nesta mesma região de configuração. Porém, observou-se resultado inverso, principalmente para Taxas de Transmissão entre ½ Hz e 10Hz. Contudo, e conforme discutido na 5.1.2, esta anomalia pode ser justificada pelo fato do passo de simulação adotado neste estudo – por questões de desempenho computacional (seção 5.1.1) – ser igual ao do parâmetro de Latência em questão (1 ms). Desta forma, **esta questão técnica também pode ser contornada com a adoção de técnicas e, sobretudo, recursos de maior poder computacional no processo de simulação dos modelos CoES.**

6. CONCLUSÕES

"Now this is not the end. It is not even the beginning of the end.
But it is, perhaps, the end of the beginning"
(CHURCHILL, 1942)

Esta seção inicia com a apresentação das conclusões deste trabalho de investigação científica em função dos objetivos propostos e dos resultados obtidos. Na sequência, elenca as principais contribuições produzidas. Por fim, sugere tópicos com potencial de continuidade para a investigação.

6.1. Conclusões

Um sistema crítico em segurança pode ser considerado seguro para operar em uma determinada aplicação e em um determinado ambiente operacional caso ele não cause danos não intencionais enquanto operar em condições normais, bem como apresentar um nível aceitável de risco de danos quando operar em condições inseguras previstas. Condições inseguras são ocasionadas por situações nas quais os elementos do sistema (individualmente ou em conjunto) expõem os envolvidos ou a si mesmos a perigos. Desta forma, situações inseguras devem ser sistematicamente identificadas e tratadas tanto na especificação quanto na implementação do sistema de forma a manter os riscos de segurança em níveis aceitáveis e, consequentemente, garantir um sistema seguro.

O paradigma de sistemas críticos em segurança vem incorporando características de complexidade inerente devido a adoção de novas ICT, arquiteturas, abordagens de desenvolvimento e operação, entre outros. Esta complexidade inerente **limita a capacidade de previsão dos comportamentos** destes sistemas complexos artificiais produzidos pelo homem, denominados como Sistemas Complexos de Engenharia (CoES). Consequentemente, as abordagens atuais de Engenharia de Segurança de Sistemas (SSE – *System Safety Engineering*) possuem capacidade limitada em identificar e prever, de forma eficiente e completa, situações inseguras que possam ocorrer **durante a operação** destes sistemas. Portanto, suas

especificações serão incompletas desde os requisitos de segurança de sistema, e falhas inseguras[57] podem causar acidentes mesmo que o sistema tenha sido validado como seguro.

Como solução à esta problemática, este trabalho de investigação científica propôs uma mudança de mentalidade (*mindset*[58]) na Engenharia de Segurança de Sistemas aplicada à Sistemas Complexos de Engenharia: ao invés de gerenciar as **falhas inseguras** com base nos requisitos de segurança – e, consequentemente, procedimentos e mecanismos – especificados e implementados durante o projeto**,** o processo de **especificação de requisitos de segurança do sistema deve ser incorporado à sua operação**. Assim, **a especificação do sistema** é atualizada **durante sua operação** caso seja identificada alguma **condição insegura** não prevista anteriormente, sobretudo no projeto. Ou seja, durante sua operação, buscam-se por **situações normais** – nas quais o sistema atende aos **requisitos de segurança** vigentes (i.e. elementos operam conforme especificado), mas o nível de risco de segurança **observado na aplicação** está abaixo do aceitável.

Quando este cenário é identificado, a **configuração (status) observada nos parâmetros** dos elementos no sistema é definida como uma **falha insegura**. Então, esta falha insegura é incorporada à especificação do sistema, atualizando seus requisitos e, consequentemente, procedimentos e mecanismos de segurança do sistema. Desta forma, caso venha a ocorrer, o sistema tem condições de **gerenciar seus recursos buscando por uma configuração de parâmetros considerada segura**.

Como forma de instanciar esta solução, este trabalho obteve uma **abordagem orientada a riscos para gerenciamento de recursos com a finalidade de garantia da segurança crítica de CoSE. Conceitualmente**, a solução proposta amplia a capacidade em **prever e lidar com condições inseguras do sistema**, independentemente de suas causas, mantendo os níveis de risco aceitáveis durante sua operação. Para que a solução proposta se concretize, esta solução exige o **monitoramento**, durante a operação, de parâmetros (configurações) dos elementos de sistema e o nível de risco da aplicação. Além disso, para ser possível gerenciar recursos do sistema buscando por uma configuração segura, é necessário **compreender** a forma como os

[57] Falhas inseguras são situações nas quais os elementos do sistema expõem os envolvidos ou a si mesmos a condições inseguras devido ao não atendimento das suas especificações.

[58] "(*A*) *mindset is a set of assumptions, methods or notations held by one or more people or groups of people which is so established that it creates a powerful incentive within these people or groups to continue to adopt or accept prior behaviors, choices, or tools.* …" (LEVESON, 2011a)

parâmetros monitorados influenciam os – e são influenciados pelos – riscos de segurança crítica de suas aplicações, bem como ser capaz de implementar a nova configuração no escopo de especificação do sistema.

Ao longo deste trabalho de investigação, observou-se que elementos dos sistemas de tráfego cooperativo no contexto dos C-ITS possuem parâmetros não dependentes entre si e cujos comportamentos são antagônicos com relação aos níveis de risco de segurança. Assim, **ao compreender as relações entre parâmetros e nível de risc**o, é possível gerenciar o nível de risco de segurança crítica da aplicação por meio da **compensação funcional** entre os parâmetros do sistema que atendam a propriedade observada.

Além disso, por meio do estudo de caso, este trabalho demonstrou que é possível se obter as configurações (status) dos parâmetros dos elementos no sistema e inferir o nível de risco da aplicação ao **monitorar**, durante a operação, os dados trafegados pelos **elementos de comunicação** dos Sistemas Ciberfísicos (CPS) distribuídos e críticos em segurança. Além disso, justificou-se que Sistemas Ciberfísicos (CPS) distribuídos e cooperativos e Sistemas de Transporte Cooperativo no contexto dos C-ITS podem ser considerados como Sistemas Complexos de Engenharia críticos em segurança.

Portanto, este trabalho propôs um conjunto de 4 (quatro) atividades – denominada como "Abordagem Proposta" – que possibilitou gerenciar os recursos do sistema orientado ao nível de risco observado na aplicação. Ao identificar uma situação insegura (configuração de parâmetros que expõem o sistema à uma condição insegura), a abordagem buscou por uma configuração de recursos que mantém o sistema em um nível aceitável de risco de segurança. Além disso, a abordagem permitiu reconhecer, durante a operação, situações inseguras não previstas, bem como atualizar as especificações de segurança para que, caso viesse a ocorrer novamente, busca-se por uma configuração de recursos que mantenha a segurança.

Essa abordagem foi implementada e avaliada em um estudo de caso em Sistemas de Transporte Cooperativo no contexto dos C-ITS. Neste estudo de caso, parâmetros do sistema de comunicação foram considerados no processo de garantia de segurança por meio do gerenciamento destes recursos orientado ao risco observado na aplicação. A abordagem foi efetiva ao possibilitar a manutenção de níveis aceitáveis de risco de segurança em uma aplicação de tráfego.

O estudo de caso implementado monitorou, durante a operação, métricas relacionadas à segurança crítica. Entre elas, a distância mínima entre veículos (dmin), justificado como um estimador de risco de segurança recomendável a este tipo de aplicação. Contudo, dmin não foi utilizada na tomada de decisão de gerenciamento de recursos durante a operação. Desta forma, não foi possível avaliar, de forma direta, a capacidade da abordagem proposta em **otimizar recursos** do sistema em momentos onde seja observado uma menor demanda da aplicação quanto à exposição aos riscos. Porém, observou-se que a abordagem proposta contribui com a definição, durante a operação, das configurações realmente necessárias à manutenção da segurança do sistema em função dos cenários de operação. Com isso, o envelope operacional do sistema pôde ser ampliado, e uma quantidade menor de recursos precisa ser alocado para garantir a segurança.

Por fim, conclui-se que os resultados obtidos neste trabalho – avaliados no contexto de C-ITS – possam ser extrapolados para outros contextos de CoES críticos em segurança. Portanto, onde for aplicável, esta generalização pode ser realizada.

6.2. Contribuições

Com base nos resultados e conclusões obtidos neste trabalho de investigação, pode-se elencar suas seguintes contribuições diretas:

- Um **princípio de compensação funcional (seção 3.3)** aplicado a parâmetros não dependentes entre si e cujos comportamentos sejam antagônicos com relação aos níveis de risco de segurança crítica. Observou-se que elementos de Sistemas Complexos de Engenharia (CoES) possuem parâmetros que atendem estas propriedades. Consequentemente, o princípio de compensação funcional poderá fazer parte da base de conceitos que orienta práticas e processos mais apropriados para lidar com o ciclo de Engenharia de Sistema Complexos, colaborando com a inovação da área.

- **Uma abordagem conceitual que contribui com a garantia da segurança crítica de Sistemas Complexos de Engenharia (seção 3.4)** – caracterizados aqui como Sistemas Ciberfísicos (CPS) Distribuídos e Cooperativos – durante situações inseguras não previstas em projeto e que se manifestam durante a operação. Esta abordagem utiliza o **princípio da compensação funcional** e, durante a operação do sistema, monitora e identifica as relações entre as métricas de **risco de segurança** observados na aplicação

crítica e o desempenho de **parâmetros** dos elementos do sistema. Então, aplicando esse conhecimento no gerenciamento dos recursos do sistema de forma a manter um nível aceitável de risco de segurança durante sua operação.

Esta abordagem tem aplicação tanto em **processos de garantia de segurança crítica** – aumentando o desempenho de parâmetros do sistema como forma de reduzir riscos de segurança em situações potencialmente inseguras – quanto em **processos de otimização de recursos** – liberando recursos do sistema em situações nas quais um nível de risco de segurança aceitável seria obtido mesmo com a redução de desempenho de seus parâmetros.

- **Uma patente de invenção** – WO/2022/115009-A1, ***Network parameter for cellular network based on safety*** (A. HATA; R. INAM; M.V. MARQUEZINI; L.F. VISMARI; A. NASCIMENTO; C. MOLINA; J.B. CAMARGO JR; J.RADY; P. CUGNASCA, 2020). Esta patente reivindica as atividades 3 e 4 da Abordagem Proposta como um método – a ser executado por um nó de uma rede móvel celular – que permite determinar e reconfigurar o valor de um parâmetro da rede com base em um conjunto de observações dos dispositivos de comunicação (outros nós da rede móvel), um nível de segurança crítica e um indicador desempenho da rede celular. Assim, é possível realizar uma alocação aprimorada de recursos de rede enquanto mantém o dispositivo móvel operando de forma autônoma de maneira segura. Com isso, a infraestrutura de comunicação do CSP pode ser otimizada com uma quantidade de recursos suficientes às necessidades reais das aplicações críticas em segurança, reduzindo custos e complexidade da infraestrutura e aumentando a **disponibilidade** dos serviços prestados.

- **Um novo método de cálculo de distância mínima entre objetos em sistemas de transporte guiado (APÊNDICE I)**. A Abordagem Proposta neste trabalho de investigação depende do monitoramento dos riscos de segurança observados durante a operação do sistema. Métricas relacionadas aos eventos de colisão entre elementos do ambiente (veículos, pedestres, obstáculos) são representativas dos riscos de segurança no escopo de sistemas que possuam elementos móveis. A distância mínima entre estes elementos (**dmin**) é uma métrica relacionada à exposição ao risco de colisão que pode ser mensurada durante a operação sem a ocorrência de acidentes. Contudo, a obtenção desta métrica é **onerosa em termos computacionais**. Desta forma, um método

desenvolvido (apresentado no **APÊNDICE I**) realiza um cálculo mais eficiente de dmin utilizando dados posicionais dos objetos, permitindo que o monitoramento seja realizado segundo requisitos mais rígidos de tempo real durante a operação do sistema por recursos computacionais limitados.

- **Uma possível abordagem para a área da *Engenharia da Resiliência*** (PROVAN *et al.*, 2020; WOODS, 2015), dado que a Abordagem Proposta neste trabalho permite que o sistema de adapte, de forma segura, a situações e condições emergentes. Ou seja, promove o aumento da resiliência dos sistemas, sobretudo sistemas inerentemente complexos, dado que os serviços providos por estes sistemas continuam disponíveis e seguros mesmo na presença de eventos disruptivos, tanto externos e internos ao sistema.

- **Uma prova de conceito ao desenvolvimento de novas aplicações em C-ITS** que possam fazer proveito dos dados de consciência situacional abertos e disponíveis nos serviços de comunicação V2X. Neste trabalho, evidenciou-se a possibilidade de **inferir o nível de segurança crítica** de um **ambiente de tráfego cooperativo** por meio das **mensagens que trafegam pela infraestrutura de rede dos serviços de comunicação V2X**, sobretudo Mensagens Básicas de Segurança (BSM) (SAE, 2022) e Mensagens de Consciência Cooperativa (CAM) (ETSI, 2019). Recentemente, o Conselho de Pesquisa em Transporte (TRB) publicou um estudo cujo objetivo é orientar as agências de transporte a utilizar os conteúdos das mensagens BSM para extrair métricas de tráfego aplicáveis em sistemas de gerenciamento (VASUDEVAN *et al.*, 2022).

- **Uma semente para a mudança de mentalidade na Engenharia de Segurança de Sistemas aplicada à Sistemas Complexos de Engenharia**. Na abordagem Proposta, as etapas de desenvolvimento e operação de um sistema não ocorrem de forma linear. Atividades como a especificação (e gerência) de requisitos de segurança do sistema é contínuo e incorporado também à operação. Aliás, a Abordagem Proposta pode ser considerada como um elemento de pressão seletiva na engenharia de sistema complexos, o qual guia a evolução do sistema e gerencia mudanças.

6.3. Trabalhos futuros

Com base nos resultados e conclusões obtidos com este trabalho de investigação, pode-se sugerir os seguintes tópicos com maior potencial para continuidade da investigação:

- Desenvolver um estudo de caso cujo Modelo Executivo possa considerar os valores de dmin durante a operação do sistema (simulação do cenário-base), sobretudo como prova de conceito para a utilização da Abordagem Proposta em processos de otimização de recursos orientada a riscos de segurança. Neste trabalho, a implementação da Abordagem Proposta ficou restrita ao processo de garantia de segurança da aplicação crítica, pois a métrica **dmin** mensurada durante a operação – relacionada à exposição ao risco de colisão, mas, não dependente da ocorrência de acidentes – não foi utilizada nas atividades 3 (elaboração do Modelo Executivo) e 4 (definição e reconfiguração da Taxa de Transmissão) da Abordagem Proposta;

- Desenvolver estudos de caso cujos Modelos Executivo considerem outros parâmetros de elementos do sistema durante sua operação (simulação do cenário-base). O estudo de caso apresentado neste trabalho ficou restrito à utilização de parâmetros do elemento de comunicação (Latência, Taxa de Transmissão) na implementação da Abordagem Proposta. Para isso, algumas questões de pesquisa precisam ser desenvolvidas:

    - A atividade 1 (inferência dos comportamentos "parâmetro vs risco") e a atividade 2 (identificação das tuplas de parâmetros que cumpram os requisitos da compensação funcional) são atividades computacionalmente complexas. Sobretudo devido ao grande espaço de combinações a serem avaliadas, o que as torna onerosas em tempo de processamento. Portanto, lidar com este problema computacional na realização das atividades 1 e 2 seria um primeiro passo para se considerar outros parâmetros dos elementos do sistema na Abordagem Proposta.

    - Neste trabalho, obteve-se uma abordagem capaz de gerenciar interações antagônicas entre parâmetros dos elementos do sistema como forma de regular o nível de segurança da aplicação. Em sistemas complexos de engenharia, deve-se considerar as interações entre elementos em domínios diferentes – como sistemas de produção, operação, reguladores, consumidores, entre outros – e em

níveis de abstração diferentes. Portanto, a abordagem proposta poderia ser expandida para este contexto multinível e interdomínio, adotando técnicas e ferramentas de modelagem que possam representar estas características. Metodologias não reducionistas para representação e análise de sistemas críticos em segurança – por exemplo, o STAMP/STPA (HOLLNAGEL; WOODS, 2006)(LEVESON, 2011b, 2011c, 2017) e o FRAM (PATRIARCA *et al.*, 2020) – podem trazer benefícios à Abordagem Proposta.

- Investigar os **aspectos da validação da segurança para sistemas utilizando a Abordagem Proposta utilizada técnicas de AI/ML**. Devido a sua capacidade em realizar tarefas de reconhecimento de padrões em bases de dados, este trabalho fez uso de técnicas Inteligência Artificial e Aprendizado de Máquina (AI/ML) para o desenvolvimento do Modelo Executivo implementado (atividade 3). Contudo, a aplicação de AI/ML em processos de tomada de decisão em sistemas críticos em segurança é uma área em desenvolvimento, e ainda não aceito normativamente em diversos domínios de aplicação crítica. O problema é, sobretudo, a limitação atual em validar que o sistema é seguro e apto a operar. O estudo apresentado no **ANEXO I** buscou contribuir com este tema, avaliando a capacidade de 'explicabilidade' dos modelos gerados por diversas categoria de técnicas de AI/ML. Este estudo indicou os modelos baseados em árvores de decisão – bem como modelos baseados em regras – como opção para lidar com o problema da validação de segurança crítica. Porém, é necessário avançar neste tema como forma de viabiliar o uso desta abordagem em sistemas reais.

- Investigar formas de implementar a Abordagem Proposta sem utilizar técnicas de AI/ML em seu desenvolvimento. Conforme discutido no trabalho, a abordagem de desenvolvimento do Modelo Executivo não fica restrita às técnicas de AI/ML. Portanto, outras abordagens de desenvolvimento, mais determinísticas, podem ser aplicadas – de forma aderente – aos conceitos principais apresentados neste trabalho.

- Investigar e desenvolver **processos eficientes evolução e de atualização do Modelo Executivo durante a operação do sistema**. Na Abordagem Proposta implementada, o Modelo Executivo deve ser continuamente atualizado com os dados obtidos durante a operação como forma de identificar situações inseguras não previstas. Contudo, ele

passa a fazer parte do processo de gerenciamento de recursos apenas quando instanciado. Consequentemente, novas situações inseguras não contempladas pelo Modelo Executivo instanciado serão gerenciadas apenas quando o Modelo Executivo em desenvolvimento (offline) for instanciado. Portanto, existe uma janela de exposição a esta situação inseguras pelo sistema. Desta forma, obter formas eficientes de **evoluir o Modelo Executivo** ao longo da operação são necessárias.

- Investigar os impactos de outras condições de mundo real – como perda de integridade dos dados de comunicação, aspectos de cibersegurança, mau-funcionamento dos componentes do sistema, entre outros – sobre a eficiência da Abordagem Proposta.

6.4. Considerações finais

Em pouco mais de duas décadas, pode-se observar uma revolução tecnológica nos domínios de aplicação crítica em segurança, onde os sistemas se tornaram cada vez mais ciberfísicos e distribuídos. Contudo, na quase totalidade destes domínios de aplicação, ainda não se observa uma característica importante que os tornam sistemas complexos de engenharia: a **tomada de decisão autônoma** – ou, pelo menos, sem um controle centralizado – pelos elementos ciberfísicos distribuídos.

Exceção a este padrão se dá aos Sistemas de Transporte Inteligentes Cooperativos (C-ITS). A tomada de decisão na condução de veículos em sistemas de transporte rodoviário é sempre suportada por mecanismos baseados em inteligência, seja humana (veículos tradicionais) ou de máquina (veículos autônomos). Além disso, diferentemente de sistemas metroviários em uso, o controle não é realizado de forma centralizada. Portanto, o C-ITS é um estudo de caso apropriado para desenvolver pesquisas na área de sistemas complexos de engenharia.

Ao utilizar este domínio de aplicação para desenvolver os conceitos e a proposta obtida neste trabalho, foi possível identificar características e comportamentos em sistemas complexos de engenharia que podem ser generalizados para outros domínios de aplicação, desde que atendidos os critérios necessários. Portanto, espera-se os resultados obtidos neste trabalho de investigação possam, de fato, servir para suportar novas pesquisas e desenvolvimento em engenharia de sistema complexos. E, de forma mais pretenciosa, tenha plantado uma semente para a possibilidade de lidar com a Engenharia de Segurança de Sistemas Complexos por meio de uma mudança completa de mentalidade na abordagem de engenharia de sistema.

REFERÊNCIAS

A. HATA; R. INAM; M.V. MARQUEZINI; L.F. VISMARI; A. NASCIMENTO; C. MOLINA; J.B. CAMARGO JR; J.RADY; P. CUGNASCA. **Network parameter for cellular network based on safety**. WO 2022/115009 A1. 2020.

ABNT. **NBR 14653-2:2011. Avaliação de bens. Parte 2: Imóveis Urbanos**. [s.l: s.n.].

ABNT. **ABNT NBR ISO 31000:2018. Gestão de Riscos - Diretrizes**. [s.l: s.n.].

ADADI, A.; BERRADA, M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). **IEEE Access**, v. 6, p. 52138–52160, 2018. Disponível em: <https://ieeexplore.ieee.org/document/8466590/>.

ADAMS, S. **Dilbert**. Disponível em: <http://dilbert.com/strip/1993-04-23>. Acesso em: 2 fev. 2023.

ALMEIDA JR, J. R. **Segurança em Sistemas Críticos e em Sistemas de Informação – Um Estudo Comparativo**. 2003. Escola Politécnica da USP, 2003.

ARENA, F.; PAU, G.; SEVERINO, A. A Review on IEEE 802.11p for Intelligent Transportation Systems. **Journal of Sensor and Actuator Networks**, v. 9, n. 2, p. 22, 26 abr. 2020. Disponível em: <https://www.mdpi.com/2224-2708/9/2/22>.

AVIZIENIS, A.; LAPRIE, J.-C. Dependable computing: From concepts to design diversity. **Proceedings of the IEEE**, v. 74, n. 5, p. 629–638, 1986. Disponível em: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1457795>.

AVIZIENIS, A.; LAPRIE, J.-C.; RANDELL, B.; LANDWEHR, C. Basic concepts and taxonomy of dependable and secure computing. **IEEE Transactions on Dependable and Secure Computing**, v. 1, n. 1, p. 11–33, jan. 2004. Disponível em: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1335465>. Acesso em: 5 jul. 2011.

BALLINGALL, S.; SARVI, M.; SWEATMAN, P. Safety Assurance Concepts for Automated Driving Systems. In: SAE Int. J. Advances & Curr. Prac. in Mobility 2(3), 2020, [...]. 2020. p. 1528–1537.

BAR-YAM, Y. **Dynamics of Complexity Systems**. 1st. ed. [s.l.] Addison-Wesley, 1997. 865 p.

BEDAU, M. A. Weak emergence. **Philosophical Perspectives**, v. 11, p. 375–399, 1997.

BESOLD, T. R.; UCKELMAN, S. L. The What, the Why, and the How of Artificial Explanations in Automated Decision-Making. 21 ago. 2018. Disponível em: <http://arxiv.org/abs/1808.07074>.

BOCCARA, N. **Modeling Complex Systems**. 2nd. ed. [s.l.] Springer, 2010. 498p p.

CAMARGO JUNIOR, J. B. **Metodologia de Análise de Risco em Sistemas Computacionais de Aplicação Crítica**. 2002. Universidade de São Paulo, 2002.

CHRISTENSEN, A.; CUNNINGHAM, A.; ENGELMAN, J.; GREEN, C.; KAWASHIMA, C.; KIGER, S.; PROKHOROV, D.; TELLIS, L.; WENDLING, B.; BARICKMAN, F. Key Considerations in the Development of Driving Automation Systems. In: 24th Enhanced Safety of Vehicles (ESV) Conference, 2015, Gothenburg, Sweden. [...]. Gothenburg, Sweden:

2015.

CHURCHILL, W. **Winston Churchill's Speech at the Mansion House, 10 November 1942**. Disponível em: <https://www.iwm.org.uk/collections/item/object/1030031903>. Acesso em: 31 maio. 2023.

CLEARY, D. Perspectives on Complex-System Engineering. **SEPO Colaborations**, v. 3, n. 2, 2005. Disponível em: <www.mitre.org/work/sepo/>.

COUNCIL EUROPEAN UNION. **On the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport (DIRECTIVE 2010/40/EU)**Official Journal of the European Union, 2010. Disponível em: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:207:0001:0013:EN:PDF>.

DOD. **MIL-STD-882E - Systems Safety**. [s.l: s.n.].

ERICSON II, C. A. **System Safety Primer**. [s.l.] CreateSpace Independent Publishing Platform, 2011. 152 p.

ETSI. **Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Definitions**. [s.l: s.n.]. Disponível em: <https://www.etsi.org/deliver/etsi_tr/102600_102699/102638/01.01.01_60/tr_102638v010101p.pdf>.

ETSI. **3GPP TR 22.885 V14.0.0 release 14 - LTE; Study on LTE support for Vehicle-to-Everything (V2X) services**. [s.l: s.n.]. Disponível em: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2898>.

ETSI. **3GPP TS 22.185 version 14.3.0 Release 14 - LTE; Service requirements for V2X services**. [s.l: s.n.]. Disponível em: <https://www.etsi.org/deliver/etsi_ts/122100_122199/122185/14.03.00_60/ts_122185v140300p.pdf>.

ETSI. **ETSI EN 302 637-2 v1.4.1. Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service**. [s.l: s.n.]. Disponível em: <https://www.etsi.org/deliver/etsi_en/302600_302699/30263702/01.04.01_60/en_30263702v010401p.pdf>.

ETSI. **3GPP TS 22.186 version 17.0.0 Release 17 - 5G; Service requirements for enhanced V2X scenarios**. [s.l: s.n.]. Disponível em: <https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>. Acesso em: 30 maio. 2023.

FLAVIO VISMARI, L.; CAMARGO JUNIOR, J. B. A safety assessment methodology applied to CNS/ATM-based air traffic control system. **Reliability Engineering & System Safety**, v. 96, n. 7, p. 727–738, jul. 2011. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S095183201100024X>. Acesso em: 18 ago. 2011.

FOSTER, H. D. Resilience theory and system evaluation. *In*: WISE, A.; HOPKIN, V. D.; STAGER, P. **Verification and validation of complex systems: Human factors issues**. Berlin: Springer Verlag, 1993. p. 35–60.

GOODE, H. H.; MACHOL, R. E. **System Engineering: an Introduction to the Design of Large-Scale Systems**. USA: McGraw-Hill, 1957. 551 p.

HARDY, T. L. **The System Safety Skeptic: Lessons Learned in Safety Management and Engineering**. [s.l.] AuthorHouse, 2010. 312 p.

HOLLNAGEL, E.; WOODS, D. D. Resilience Engineering Precepts. *In*: HOLLNAGEL, E.; WOODS, D. D.; LEVESON, N. G. **Resilience engineering: concepts and precepts**. [s.l.] Ashgate Publishing Co., 2006. p. 347–357.

HOLLNAGEL, E.; WOODS, D. D.; LEVESON, N. **Resilience engineering: concepts and precepts**. 1st. ed. [s.l.] Ashgate Publishing Co., 2006. 410p. p.

ICAO. **MANUAL OF TECHNICAL PROVISIONS FOR THE AERONAUTICAL TELECOMMUNICATION NETWORK (ATN)**. 2a. ed. Montreal, CA: ICAO, 1999. 1204 p.

IEC. Functional safety of electrical/electronic/programmable electronic safety-related systems, Parts 1 to 7. **IEC 61508**, 2010.

IEEE. IEEE Recommended Practice for Architectural Description for Software-Intensive Systems. **IEEE Std 1471-2000**, p. 1–30, 2000.

INCOSE. **INCOSE Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities**. 5th. ed. [s.l.] John Wiley & Sons Ltd., 2023. 1569 p.

ISO/IEC/IEEE. ISO/IEC/IEEE International Standard - Systems and software engineering -- Guidelines for the utilization of ISO/IEC/IEEE 15288 in the context of system of systems (SoS). **ISO/IEC/IEEE 21840:2019(E)**, p. 1–68, 2019.

ISO/IEC. **ISO/IEC Guide 51: Safety aspects -- Guidelines for their inclusion in standards**. [s.l: s.n.].

ISO. **ISO Guide 73:2009 - Risk management -- Vocabulary**. [s.l: s.n.]. Disponível em: <https://www.iso.org/standard/44651.html>.

JOHNSON, C. What are emergent properties and how do they affect the engineering of complex systems?☆. **Reliability Engineering & System Safety**, v. 91, n. 12, p. 1475–1481, dez. 2006. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0951832006000330>. Acesso em: 29 ago. 2011.

KHAITAN, S. K.; MCCALLEY, J. D. Design Techniques and Applications of Cyberphysical Systems: A Survey. **IEEE Systems Journal**, v. 9, n. 2, p. 350–365, jun. 2015. Disponível em: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6853346>.

KIM, I.; YOO, H.; YOUNG-HYUN, E.; CHO, S.; JEON, B. An Integrated Communication Message Framework of Inter-Vehicles for Connected Vehicles using Mobile Virtual Fence(MVF). **International Journal of Engineering & Technology**, v. 7, n. 3.33, p. 102, 29 ago. 2018. Disponível em: <https://www.sciencepubco.com/index.php/ijet/article/view/18584>.

KNIGHT, J. Safety Standards – a New Approach. *In*: DALE, C.; ANDERSON, T. **Addressing Systems Safety Challenges: Proceedings of the Twenty-second Safety-critical Systems Symposium, Brighton, UK, 4-6th February 2014**. [s.l.] CreateSpace Independent Publishing Platform, 2014. p. 312.

LAPRIE, J.-C. From Dependability to Resilience. In: 38th IEEE/IFIP Int. Conf. On Dependable Systems and Networks - Sup. Vol., 2008, Anchorage, Alaska. [...]. Anchorage, Alaska: IEEE/IFIP, 2008. p. G8–G9.

LEMES, M. J. R. **Complexidade, acoplamento e criticalidade (C^2A) como indicadores de**

risco em projetos de sistemas. 2011. Universidade de São Paulo, São Paulo, 2011. Disponível em: <http://www.teses.usp.br/teses/disponiveis/3/3141/tde-28052012-114709/>.

LEVESON, N. The Use of Safety Cases in Certification and Regulation. **Journal of System Safety (eEdition)**, v. 47, n. 6, p. 13–25, 2011a. Disponível em: <http://www.system-safety.org/ejss/past/current/>.

LEVESON, N.; DULAC, N.; ZIPKIN, D.; CUTCHER-GERSHENFELD, J.; CARROLL, J.; BARRETT, B. Engineering Resilience into Safety-Critical Systems. *In*: HOLLNAGEL, E.; WOODS, D.; LEVESON, N. **Resilience engineering: concepts and precepts**. [s.l.] Ashgate Publishing Co., 2006. p. 95–123.

LEVESON, N. G. Applying systems thinking to analyze and learn from events. **Safety Science**, v. 49, n. 1, p. 55–64, 2011b. Disponível em: <http://www.sciencedirect.com/science/article/B6VF9-4Y7P9S0-3/2/bb3a9b43a9b551439e4545b0b5e25d17>.

LEVESON, N. G. **Engineering a Safer World: Systems Thinking Applied to Safety**. 1st. ed. [s.l.] MIT Press, 2011c. 608pp p.

LEVESON, N. G. Rasmussen's legacy: A paradigm change in engineering for safety. **Applied Ergonomics**, v. 59, p. 581–591, mar. 2017. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0003687016300151>.

LU, J.; SIBAI, H.; FABRY, E.; FORSYTH, D. NO Need to Worry about Adversarial Examples in Object Detection in Autonomous Vehicles. 11 jul. 2017. Disponível em: <http://arxiv.org/abs/1707.03501>.

MATHWORKS. **What is Machine Learning?** Disponível em: <https://www.mathworks.com/discovery/machine-learning.html>. Acesso em: 31 maio. 2023.

MAZHAR RATHORE, M.; SHAH, S. A.; SHUKLA, D.; BENTAFAT, E.; BAKIRAS, S. The Role of AI, Machine Learning, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities. **IEEE Access**, p. 1–1, 2021. Disponível em: <https://ieeexplore.ieee.org/document/9359733/>.

MICULESCU, D.; KARAMAN, S. Polling-systems-based Autonomous Vehicle Coordination in Traffic Intersections with No Traffic Signals. 26 jul. 2016. Disponível em: <http://arxiv.org/abs/1607.07896>.

MICULESCU, D.; KARAMAN, S. Polling-Systems-Based Autonomous Vehicle Coordination in Traffic Intersections With No Traffic Signals. **IEEE Transactions on Automatic Control**, v. 65, n. 2, p. 680–694, fev. 2020. Disponível em: <https://ieeexplore.ieee.org/document/8732975/>.

MOLINA, C. B. S. T.; ALMEIDA, J. R. de; VISMARI, L. F.; GONZALEZ, R. I. R.; NAUFAL, J. K.; CAMARGO, J. B. Assuring Fully Autonomous Vehicles Safety by Design: The Autonomous Vehicle Control (AVC) Module Strategy. In: 2017 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W), 2017, [...]. IEEE, 2017. p. 16–21.

MOLINA, C. B. S. T.; VISMARI, L. F.; FUJI, T.; CAMARGO JR, J. B.; DE ALMEIDA JR, J. R.; INAM, R.; FERSMAN, E.; HATA, A.; MARQUEZINI, M. V. Enhancing Sensor Capabilities of Open-Source Simulation Tools to Support Autonomous Vehicles Safety Validation. *In*: BARBARA GALLINA, AMUND SKAVHAUG, ERWIN SCHOITSCH, F. B. **Computer Safety, Reliability and Security - SAFECOMP 2018 Workshops -**

ASSURE, DECSoS, SASSUR, STRIVE, and WAISE. [s.l.] Springer International Publishing, 2018. p. 353–364.

MURDOCH, W. J.; SINGH, C.; KUMBIER, K.; ABBASI-ASL, R.; YU, B. Definitions, methods, and applications in interpretable machine learning. **Proceedings of the National Academy of Sciences**, v. 116, n. 44, p. 22071–22080, 29 out. 2019. Disponível em: <http://www.pnas.org/lookup/doi/10.1073/pnas.1900654116>.

MURPHY, N.; ELLIS, G. F. R.; O'CONNOR, T. (ed.). **Downward Causation and the Neurobiology of Free Will**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

NASCIMENTO, A. M.; VISMARI, L. F.; MOLINA, C. B. S. T.; CUGNASCA, P. S.; CAMARGO JR, J. B.; DE ALMEIDA JR, J. R.; INAM, R.; FERSMAN, E.; MARQUEZINI, M. V.; HATA, A. Y. A Systematic Literature Review about the impact of Artificial Intelligence on Autonomous Vehicle Safety. **IEEE Transactions on Intelligent Transportation Systems**, p. 1–19, 4 abr. 2019. Disponível em: <https://ieeexplore.ieee.org/document/8892611/>.

NATELLA, R.; COTRONEO, D.; MADEIRA, H. S. Assessing Dependability with Software Fault Injection. **ACM Computing Surveys**, v. 48, n. 3, p. 1–55, 8 fev. 2016. Disponível em: <https://dl.acm.org/doi/10.1145/2841425>.

NAUFAL, J. K.; CAMARGO, J. B.; VISMARI, L. F.; DE ALMEIDA, J. R.; MOLINA, C.; GONZALEZ, R. I. R.; INAM, R.; FERSMAN, E. A 2 CPS: A Vehicle-Centric Safety Conceptual Framework for Autonomous Transport Systems. **IEEE Transactions on Intelligent Transportation Systems**, v. 19, n. 6, p. 1925–1939, jun. 2018. Disponível em: <http://ieeexplore.ieee.org/document/8054749/>.

NORMAN, D. A. **Living with Complexity**. [s.l.] The MIT Press, 2010. 312 p.

NORMAN, D. O. **Engineering a Complex System: A Study of the AOC**: Mitre Tech Papers 2004. [s.l: s.n.]. Disponível em: <http://www.mitre.org/work/tech_papers/tech_papers_04/04_0527/index.html>.

NORMAN, D. O.; KURAS, M. L. **Engineering Complex Systems**: MITRE Tech Papers 2004. [s.l: s.n.]. Disponível em: <http://www.mitre.org/work/tech_papers/tech_papers_04/norman_engineering/>.

PATRIARCA, R.; DI GRAVIO, G.; WOLTJER, R.; COSTANTINO, F.; PRAETORIUS, G.; FERREIRA, P.; HOLLNAGEL, E. Framing the FRAM: A literature review on the functional resonance analysis method. **Safety Science**, v. 129, 1 set. 2020.

PICARDI, C.; PATERSON, C.; HAWKINS, R.; CALINESCU, R.; HABLI, I. Assurance Argument Patterns and Processes for Machine Learning in Safety-Related Systems. (H. Espinoza, J. Hernández-Orallo, X. C. Chen, S. S. ÓhÉigeartaigh, X. Huang, M. Castillo-Effen, R. Mallah, J. McDermid) In: Proceedings of the Workshop on Artificial Intelligence Safety (SafeAI 2020), 2020, New York, USA, Feb 7, 2020. [...]. New York, USA, Feb 7, 2020.: 2020. p. 23–30.

PROVAN, D. J.; WOODS, D. D.; DEKKER, S. W. A.; RAE, A. J. Safety II professionals: How resilience engineering can transform safety practice. **Reliability Engineering & System Safety**, v. 195, p. 106740, mar. 2020. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0951832018309864>.

RAJABLI, N.; FLAMMINI, F.; NARDONE, R.; VITTORINI, V. Software Verification and Validation of Safe Autonomous Cars: A Systematic Literature Review. **IEEE Access**, 2020.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. 4th. ed. [s.l.] Pearson, 2021.

SAE. **Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems (Ground Vehicle Standard J3016_201401)**. [s.l: s.n.].

SAE. **J2735_202211: V2X Communications Message Set Dictionary - SAE International**. [s.l: s.n.]. Disponível em: <https://www.sae.org/standards/content/j2735_202211/>. Acesso em: 31 maio. 2023.

SARKER, I. H. AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems. **SN Computer Science**, v. 3, n. 2, p. 158, 10 mar. 2022. Disponível em: <https://link.springer.com/10.1007/s42979-022-01043-x>.

SEPULCRE, M.; GOZALVEZ, J. Context-aware heterogeneous V2X communications for connected vehicles. **Computer Networks**, v. 136, p. 13–21, maio 2018. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S1389128618300975>.

SHEARD, S. A.; MOSTASHARI, A. Principles of complex systems for systems engineering. **Systems Engineering**, v. 12, n. 4, p. 295–311, set. 2009. Disponível em: <http://doi.wiley.com/10.1002/sys.20124>. Acesso em: 17 out. 2011.

SINGH, S. **Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey**. [s.l: s.n.]. Disponível em: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812115>.

SOMMERVILLE, I. **SOFTWARE ENGINEERING**. 9th. ed. [s.l.] Addison-Wesley, 2011. 790 p.

SOMMERVILLE, I.; CLIFF, D.; CALINESCU, R.; KEEN, J.; KELLY, T.; KWIATKOWSKA, M.; MCDERMID, J.; PAIGE, R. Large-scale complex IT systems. **Communications of the ACM**, v. 55, n. 7, p. 71, 1 jul. 2012. Disponível em: <http://dl.acm.org/citation.cfm?doid=2209249.2209268>. Acesso em: 12 mar. 2013.

US-DOT. **ARC-IT Physical View**. Disponível em: <https://www.arc-it.net/html/viewpoints/physical.html>. Acesso em: 17 ago. 2023.

VASUDEVAN, M.; O'HARA, J.; TOWNSEND, H.; ASARE, S.; MUHAMMAD, S.; OZBAY, K.; YANG, D.; GAO, J.; KURKCU, A.; ZUO, F. **Algorithms to Convert Basic Safety Messages into Traffic Measures**. Washington, D.C.: Transportation Research Board, 2022. 172 p.

VISMARI, L. F. **Vigilância dependente automática no controle de tráfego aéreo: avaliação de risco baseada em modelagem em redes de Petri fluidas e estocásticas.** 2007. Universidade de São Paulo, São Paulo, 2007. Disponível em: <http://www.teses.usp.br/teses/disponiveis/3/3141/tde-08012008-153718/>.

VISMARI, L. F.; CAMARGO, J. B.; NAUFAL, J. K.; DE ALMEIDA, J. R.; MOLINA, C. B. S. T.; INAM, R.; FERSMAN, E.; MARQUEZINI, M. V. A Fuzzy logic, risk-based autonomous vehicle control approach and its impacts on road transportation safety. In: 2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES), 2018a, Madrid, Spain. [...]. Madrid, Spain: IEEE, 2018. p. 1–7.

VISMARI, L. F.; CAMARGO JUNIOR, J. B. Safety assurance in complex engineered systems: could the current safety standards and regulations be doing our critical systems more unsafe? In: II Brazilian Conference on Critical Embedded Systems (CBSEC) - Student Workshop, 2012, Campinas, Brazil. [...]. Campinas, Brazil: 2012. p. 21–28.

VISMARI, L. F.; MOLINA, C. B. S. T.; JR, J. B. C.; JR, J. R. A.; INAM, R.; FERSMAN, E.; MARQUEZINI, M. V. A simulation-based safety analysis framework for autonomous vehicles – assessing impacts on Road Transport System's safety and efficiency. (J. E. V. Stein Haugen, Anne Barros, Coen van Gulijk, Trond Kongsvik) In: Safety and Reliability – Safe Societies in a Changing World: Proceedings of ESREL 2018, June 17-21, 2018, Trondheim, Norway, 2018b, [...]. CRC PRESS, 2018. p. 2067--2075.

WIBLE, J. What is Complexity? *In*: **Complexity and the History of Economic Thought**. [s.l.] Routledge, 2000.

WOODS, D. D. Four concepts for resilience and the implications for the future of resilience engineering. **Reliability Engineering & System Safety**, v. 141, p. 5–9, set. 2015. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0951832015000848>.

XIONG, J. **New Software Engineering Paradigm Based on Complexity Science: An Introduction to NSE**. 1st. ed. [s.l.] Springer New York, 2011. 746 p. 488 illus. p.

XU, Y.; GOODACRE, R. On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. **Journal of Analysis and Testing**, v. 2, n. 3, p. 249–262, 29 jul. 2018. Disponível em: <http://link.springer.com/10.1007/s41664-018-0068-2>.

APÊNDICE I. MÉTODO DE CÁLCULO DA DISTÂNCIA MÍNIMA ENTRE OBJETOS EM CENÁRIOS DE MOVIMENTO GUIADO

Resumo. Este trabalho propõe um novo método que tem como objetivo apoiar o processo de estimativa de risco em Sistemas Inteligentes de Transporte e outros ambientes críticos com objetos móveis guiados. O método também é aplicável em quaisquer funções relacionadas à segurança crítica em sistemas de transporte guiados que necessitem o rastreamento da distância mínima entre os objetos. Este método é aplicável a cenários reais e simulados. Dado que este método fornece uma maneira mais eficiente de calcular as distâncias mínimas entre objetos, suas vantagens são (i.) reduzir o tempo de conclusão do cálculo, permitindo que sistemas de tomada de decisão baseados em risco atendam a requisitos de tempo real mesmo usando recursos computacionais limitados; e (ii.) redução do consumo de energia, tornando a respectiva aplicação mais sustentável; (iii.) permitir modelar e simular sistemas maiores e mais complexos usando os sistemas computacionais atuais.

1. INTRODUÇÃO

Sistemas críticos em segurança são capazes de expor vidas humanas, o meio ambiente e a propriedade a perdas e danos. Um alto nível de risco pode comprometer a existência de um sistema, inviabilizando-o. Portanto, gerenciar os riscos de segurança crítica, mantendo-os em níveis aceitáveis, é uma tarefa obrigatória em domínios de aplicação críticos. Assim, é imperativo identificar, estimar, avaliar e tratar os riscos de segurança de forma eficiente e econômica e durante todo o ciclo de vida dos sistemas críticos – desde a concepção até a desativação do sistema.

Sistemas de transporte, bem como em outros domínios cujos sistemas possuem objetos móveis (como manufatura automatizada) são sistemas intrinsecamente críticos para a segurança. Nestes sistemas, a maior parte dos riscos de segurança está relacionada à colisão entre objetos em movimento, tanto entre si quanto com pessoas e outros elementos do ambiente. Os riscos de colisão estão diretamente relacionados às distâncias mínimas que os elementos do ambiente são mantidos entre si. Quanto menores as distâncias mínimas entre os elementos, maiores os riscos de colisão e, consequentemente, menos seguro é o sistema.

Portanto, a distância mínima entre os elementos pode ser usada tanto como uma **variável proxy**[59] para gerenciamento de risco quanto como entrada para outras funções relacionadas a riscos em sistemas de transporte. Pode ser aplicado tanto em atividades de avaliação de riscos – calculando a probabilidade de colisão e, consequentemente, estimando os riscos de colisão – quanto como entrada de algoritmos nos níveis operacional (evitar colisões), tático (aprendizagem por reforço) e/ou estratégico (planejamento).

No entanto, calcular distâncias mínimas entre elementos em um ambiente densamente populado pode ser **computacionalmente oneroso**, pois é necessário monitorar **continuamente** a posição de todos os Objetos de Interesse (OoI), armazenar e calcular a **distância** entre os $\binom{n}{2}$ pares de OoI (**dist(OoI$_i$, OoI$_{i+1}$), i = [1,** $\binom{n}{2}$**-1]**), onde 'n' é a quantidade de OoI no ambiente monitorado. Quando um OoI sai do ambiente monitorado no instante de tempo t_i (OoI'), a **distância mínima** entre este OoI' e cada OoI monitorado é obtida a partir da busca do menor valor de mínima **dist(OoI', OoI)** mensurada entre OoI' e os demais 'n-1' elementos nos dados armazenados durante todo o período de tempo monitorado ($dmin(OoI', OoI) = min(dist(OoI', OoI_k)(t)), \forall k, 0 < t < t_i$). Este processo de obtenção das distâncias mínimas entre OoI é uma tarefa computacionalmente onerosa, tanto em termos de obtenção de dados de posicionamente e de cálculo de distâncias .

Como forma de reduzir os custos computacionais, pode-se utilizar estratégias de agrupamento para definir os objetos dentro de um cluster onde o OoI está contido. Assim, o número de distâncias a serem calculadas pode ser reduzido. No entanto, problemas de agrupamento podem variar de soluções de tempo polinomial a *NP-Hard*, dependendo da estratégia de agrupamento adotada, que é influenciada pelos algoritmos de cálculo de distância.

Em resumo, algumas abordagens de cálculo existentes dependem da atualização de todas as distâncias para cada etapa da simulação, tarefa intrinsecamente onerasa em termos computacionais. Outras abordagens exigem a atualização das distâncias dos objetos dentro de uma região ou cluster, reduzindo o número de cálculos de distâncias e, consequentemente, o custo computacional. No entanto, estas abordagens demandam de computação adicional para

[59] "Variáveis 'Proxy' são variáveis utilizadas para substituir outra variável de difícil mensuração e que se presume guardar com ela relação de pertinência" (ABNT, 2011). É uma variável que possa ser observada e que esteja fortemente correlacionada com a variável de interesse, cuja a observação direta é difícil (ou impossível) de se obter.

avaliar a região e os objetos dentro dela ou para atualizar o cluster onde o OoI está localizado e identificar os objetos de cluster que ele inclui. Além disso, em ambas as abordagens, a medição de distância não leva em consideração a dinâmica específica do fenômeno alvo da investigação. Mesmo nos cenários mais simples, onde apenas algumas distâncias precisam ser calculadas. Isso pode reduzir sensivelmente o valor proxy das medições para fins de avaliação de risco de segurança.

Dada sua importância e utilidade, sobretudo para o gerenciamento de risco em sistemas críticos que possuem objetos em movimento (como os sistemas de transporte), faz-se necessária uma forma computacionalmente eficiente de se obter os valores de distância mínima entre objetos. Desta forma, este trabalho propõe um novo método de medição de distância mínima para calcular os riscos de colisão entre objetos em movimento em **ambientes de movimento guiado**, tanto em cenários reais quanto simulados.

Utilizando características intrínsecas de movimentação em sistemas de transporte guiado, o novo método **transfere o foco de cálculo, de todos os pares de OoI para as regiões em que é possível uma colisão** (regiões de conflito), reduzindo a complexidade do cálculo. Consequentemente, esta solução é computacionalmente menos onerosa, mais eficiente e mais rápida, permitindo melhorar a eficiência e reduzir custos nas atividades de gestão de riscos *runtime* (operação do sistema) e *offline* (desenvolvimento do sistema).

A próxima seção apresenta, em detalhes, o desenvolvimento e obtenção deste método.

2. PROPOSTA DE MÉTODO DE $dmin(OoI_i,OoI_j)$ EM TRÁFEGO GUIADO

Nos métodos atuais de medição de distância mínima, cada posição OoI deve ser continuamente monitorada ao longo do tempo, e as distâncias entre todos os pares OoI são calculadas e armazenadas. A distância mínima observada para cada par OoI é obtida a partir de uma busca nos dados armazenados $min(dist(OoI', OoI_k)(t)), \forall k, 0 < t < t_i)$.

Por outro lado, este trabalho propõe um novo método de medição de distância como forma de se obter as distâncias mínimas entre Objetos de Interesse (OoI) em sistemas contendo objetos em movimento. Neste método, apenas as regiões de interseção (cruzamento) formadas pelas trajetórias de cada par OoI são monitoradas. Quando um dos OoI sai de uma região, é calculada a distância entre o outro OoI do par e a região de interseção. O valor obtido usando

este método é considerada a distância mínima entre um par OoI ($dmin(OoI', OoI) = min(dist(OoI', OoI_k)(t)), \forall k, 0 < t < t_i$).

A Figura 1 ilustra um cenário hipotético incluindo 4 objetos em movimento (numerados de '1' a '4'). Suas trajetórias também são ilustradas. Como pode ser observado, os objetos têm suas trajetórias se cruzando em alguns pontos da superfície. Assim, conforme os objetos se movem em suas respectivas trajetórias, eles terão algum risco de colisão nessas interseções.

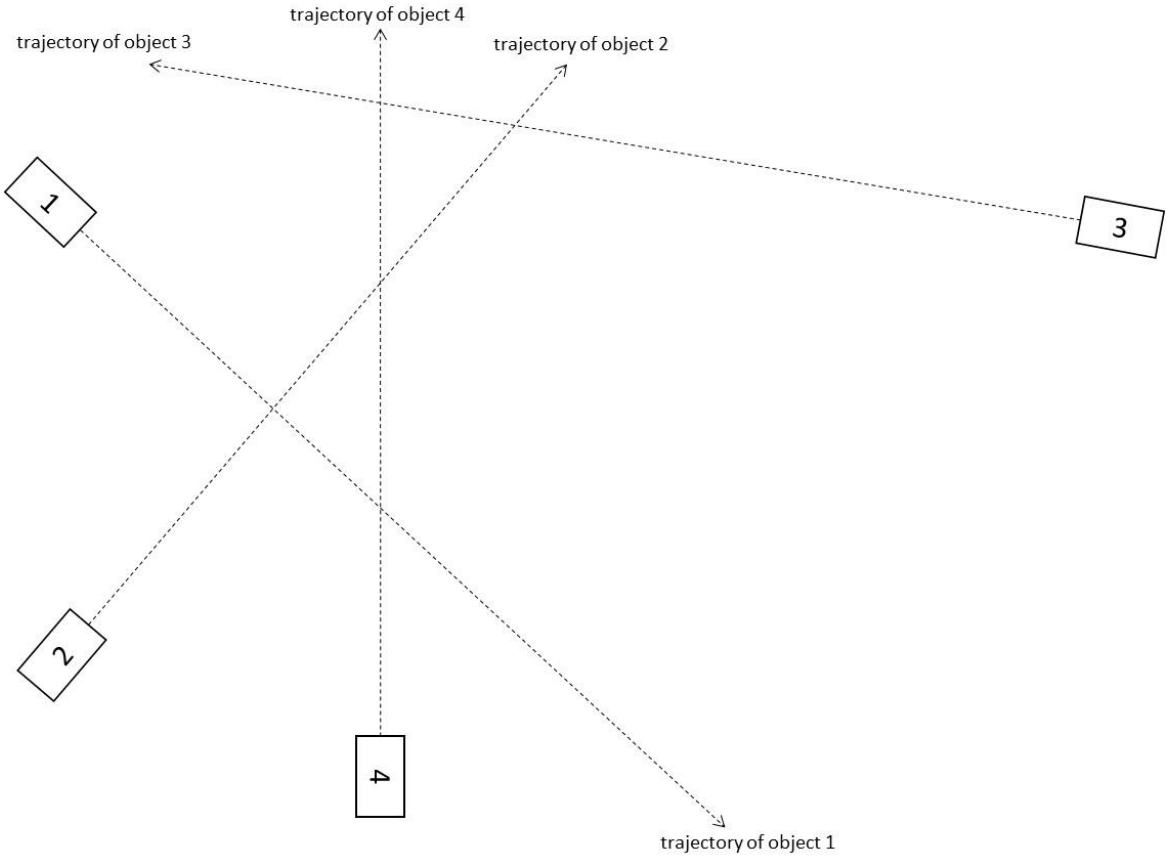


Figura 1. Exemplo de cenário de tráfego com 4 objetos em movimento e suas trajetórias

Considerando o cenário de tráfego apresentado na Figura 1, a Figura 2 ilustra as regiões nas quais pelo menos dois objetos podem colidir entre si. Denominadas como ‘regiões de conflito’ (ou ‘zonas de conflito’), são regiões físicas ao redor de cada interseção de trajetória onde há um potencial de interações entre objetos em movimento. Cada região é rotulada por $I_{i,j}$, onde, por convenção, 'i' é o número da trajetória do objeto (trajetória) que se espera chegar primeiro na região de conflito, e 'j' é o número da trajetória do objeto se chegará depois de ‘i’ na região do conflito. Assim, $I_{1,2}$ é o rótulo da região de conflito entre as trajetórias 1 e 2, onde se espera que o objeto 1 chegue primeiro e o objeto 2 chegue depois.

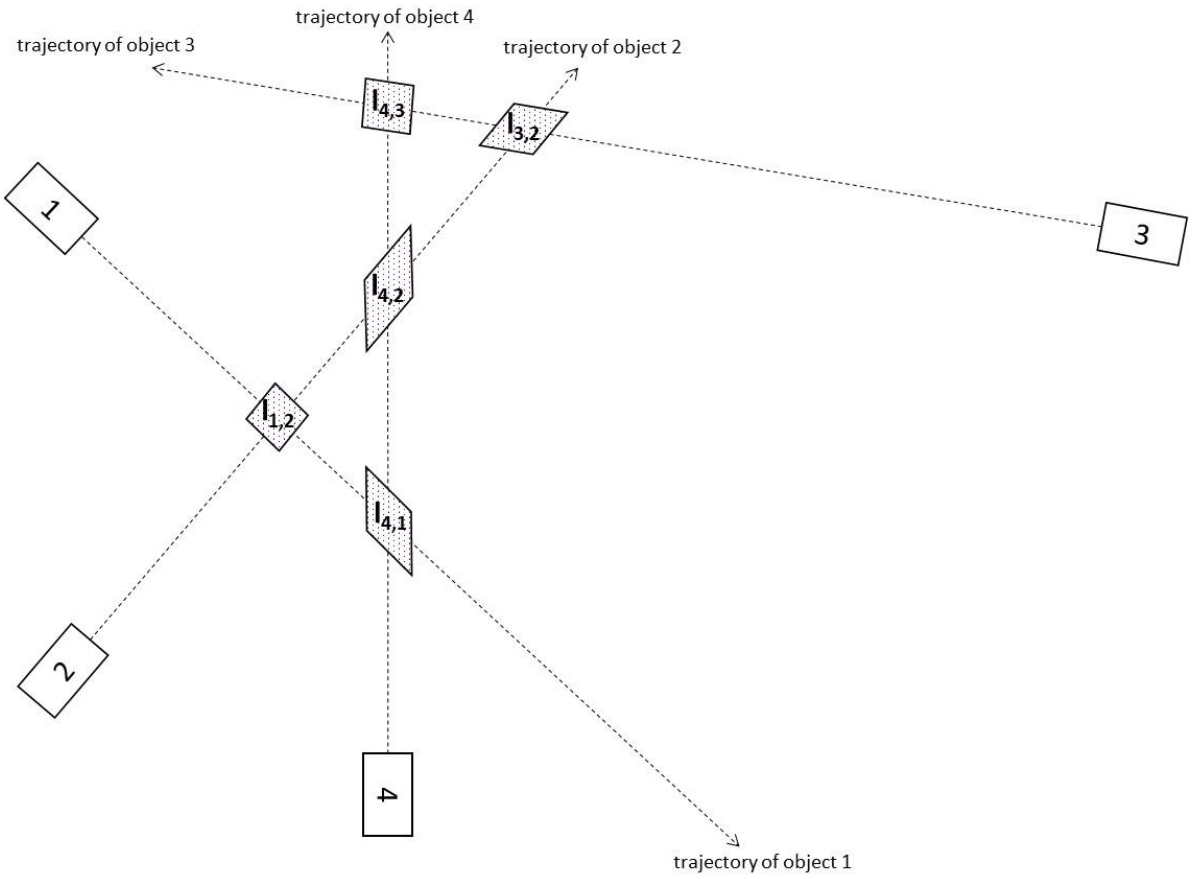


Figure 2. Regiões de colisão para o cenário de tráfego da Figura 1

A Figura 3 ilustra uma sequência de quatro instantes, representando a aproximação dos objetos 1 e 2 à sua região de conflito ($I_{1,2}$). Caso não seja possível determinar qual deles será o primeiro e o segundo a atingir a região de conflito, a distância de cada objeto à região de conflito precisa ser medida continuamente. Neste exemplo, o objeto 1 chegará primeiro, e o objeto 2 chegará na região depois do objeto 1. Assim, sua distância a $I_{1,2}$ será a única medida avaliada. A variável $d_{2,I12}$ indica a distância entre o objeto 2 e a interseção I1,2. Em cada amostra monitorada, essa distância é avaliada até que o objeto 1 saia da região de conflito (Figura 3. [D]). Esta última medida é a distância mínima da interação em potencial. Em outras palavras, é o quão perto o objeto 2 esteve de colidir com o objeto 1.

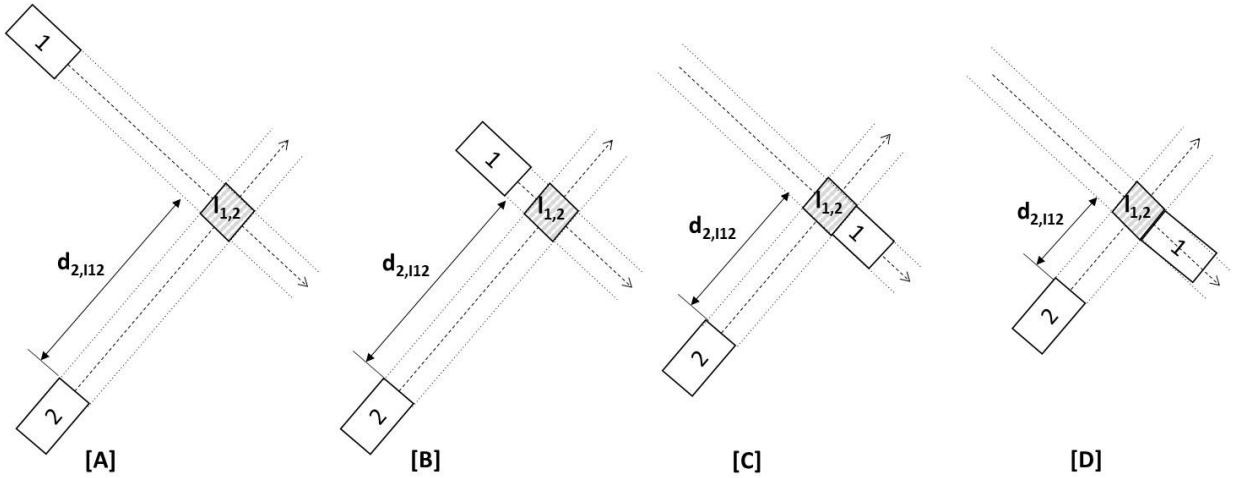


Figura 3. Exemplo do processo de medição de distância (4 instantes)

A Figura 4 ilustra o cenário [A] da Figura 3, no qual foram incluídos o sistema de coordenadas (0_{xy}) e as distâncias $d_{1,I12}$, $d_{2,I12}$ e $d_{1,2}$. Neste exemplo, o eixo 'x' foi convenientemente definido como paralelo a $d_{2,I12}$, e 'y' está paralelo a $d_{1,I12}$. Usando este sistema de coordenadas, a **Eq. 1** apresenta o cálculo da distância de $d_{1,I12}$ e $d_{2,I12}$.

$$d_{2,I12} = \left|X_{I1,2} - X_2\right| \quad and \quad d_{1,I12} = \left|Y_{I1,2} - Y_1\right| \qquad \text{(Eq. 1)}$$

O cálculo de $d_{1,2}$ pode ser executado por meio de algumas heurísticas existentes. A **distância euclidiana** é dada pela **Eq. 2**. A **distância de Manhattan** é dada pela **Eq. 3**. A **distância Octile** é dada pela **Eq. 4**. Finalmente, a **distância Chebyshev** é dada pela **Eq. 5**. Vale ressaltar, neste caso especial ilustrado pela **Figura 4** – onde as duas trajetórias são ortogonais – o método proposto para o cálculo da distância coincide com a distância de Chebyshev. Entretanto, se as trajetórias não forem ortogonais, o cálculo da distância proposto não é dado pela distância de Chebyshev (**Figura 5**), conforme demonstrado a seguir.

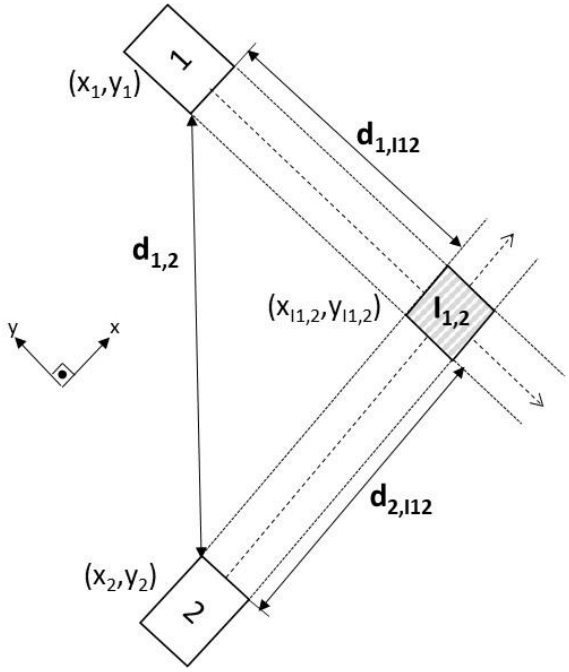


Figura 4. Sistemas de coordenadas e distâncias entre OoI

$$\sqrt[2]{{d_{2,I12}}^2 + {d_{1,I12}}^2} \quad \text{(Eq. 2)}$$

$$d_{2,I12} + d_{1,I12}. \quad \text{(Eq. 3)}$$

$$\max(d_{1,I12}, d_{2,I12}) + (\sqrt[2]{2} - 1) \cdot min(d_{1,I12}, d_{2,I12}). \quad \text{(Eq. 4)}$$

$$\max(d_{1,I12}, d_{2,I12}). \quad \text{(Eq. 5)}$$

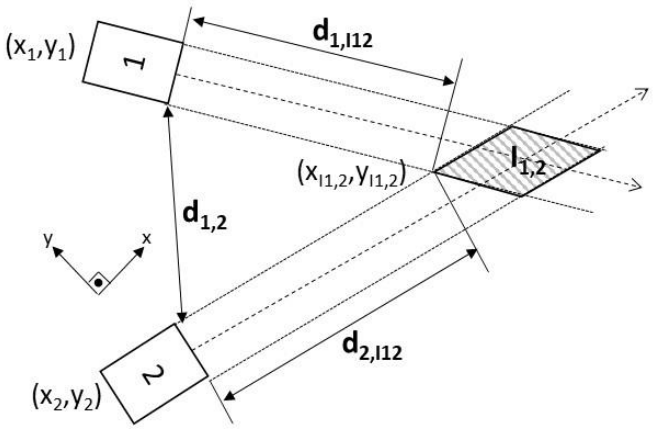


Figura 5. Exemplo de trajetórias não ortogonais

A Figura 5 representa duas trajetórias não ortogonais, e a coordenada $(X_{I1,2}, Y_{I1,2})$ refere-se à fronteira mais próxima da região de conflito (um paralelogramo) ao objeto 1 e ao objeto 2. Nesta situação, a formulação das distâncias $d_{1,I12}$ e $d_{2,I12}$ muda conforme mostrado na Eq. 6.

$$d_{2,I12} = \sqrt[2]{\left|X_{I1,2} - X_2\right|^2 + \left|Y_{I1,2} - Y_2\right|^2} \quad d_{1,I12} = \sqrt[2]{\left|X_{I1,2} - X_1\right|^2 + \left|Y_{I1,2} - Y_1\right|^2} \quad \text{Eq.6}$$

O cálculo de $d_{1,2}$ pode ser executado por meio das mesmas heurísticas aplicadas no cenário da Figura 4. Assim, a **distância euclidiana** é dada pela Eq.7. A **distância de Manhattan** é dada pela Eq.8 (mas, suas condições de contorno não são respeitado aqui). A **distância Octile** é dada pela Eq.9. Por fim, a **distância Chebyshev** é dada pela Eq.10. Neste caso, observamos que nenhuma dessas heurísticas coincide com a Eq.6, distância calculada neste método.

$$\sqrt[2]{|X_1 - X_2|^2 + |Y_1 - Y_2|^2}. \quad \text{(Eq.7)}$$

$$|X_1 - X_2| + |Y_1 - Y_2|. \quad \text{(Eq.8)}$$

$$\max(|X_1 - X_2|, |Y_1 - Y_2| + (\sqrt[2]{2} - 1) \cdot \min(|X_1 - X_2|, |Y_1 - Y_2|. \quad \text{(Eq.9)}$$

$$\max(|X_1 - X_2|, |Y_1 - Y_2|) \quad \text{(Eq.10)}$$

A Figura 6 apresenta equações para um caso geral de trajetórias não ortogonais considerando as duas possíveis confluências de trajetórias: ângulos de confluência de trajetórias entre 0º e 90º (esquerda) e entre 90º e 180º (direita). As distâncias Superior e Inferior de um objeto até a interseção ($d^U_{i,L12}$, $d^L_{i,L12}$, respectivamente) são calculadas. Observamos que a trajetória ortogonal é um caso especial de trajetória não ortogonal, em que $d^U_{i,L12} = d^L_{i,L12}$. Ao invés de medir apenas uma distância a cada objeto de passagem ($d_{1,L12}$, $d_{2,L12}$), mede-se $d^U_{i,L12}$ e $d^L_{i,L12}$ a cada objeto 'I'.

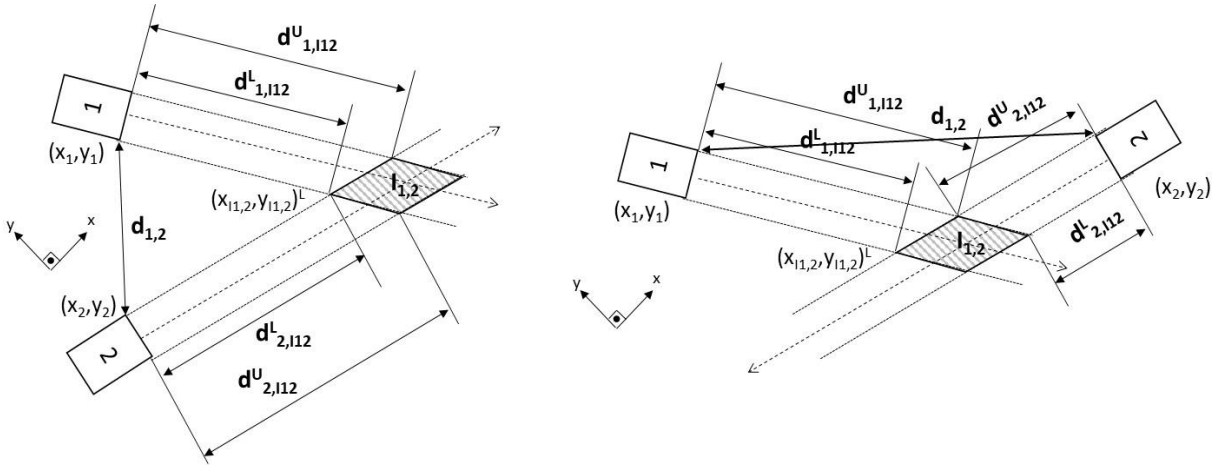


Figure 1. Equações gerais de distância para duas possíveis trajetórias de confluência: com ângulo entre]0º; 90º[(esquerda) e]90º; 180º[(direita)

A Figura 7 ilustra o caso geral apresentado na Figura 3.[D], representando todas as possíveis situações de cruzamento em relação ao objeto que sai primeiro da região de conflito: Objeto 1, [A] e [D]; Objeto 2, [B] e [C]. Podemos observar que $d_{1,2} = d_{U2,I12}$ [A], $d_{1,2} = d^U_{1,I12}$ [B], $d_{1,2} \approx d^L_{1,I12}$ [C] e $d_{1,2} \approx d^L_{2,I12}$ [D]. Consequentemente, podemos definir a seguinte regra geral para medir distâncias mínimas entre objetos que se cruzam:

- Quando "ângulo de confluência da trajetória" =]0º e 90_o]: mínimo $d_{1,2} = d^U_{i,I12}$ quando a parte de trás do primeiro objeto que sai começa a sair da região de conflito ($I_{1,2}$).

- Quando "ângulo de confluência da trajetória" =]90º e 180º[: mínimo $d_{1,2} = d^L_{i,I12}$ quando a parte de trás do primeiro objeto que sai termina de sair da região de conflito ($I_{1,2}$).

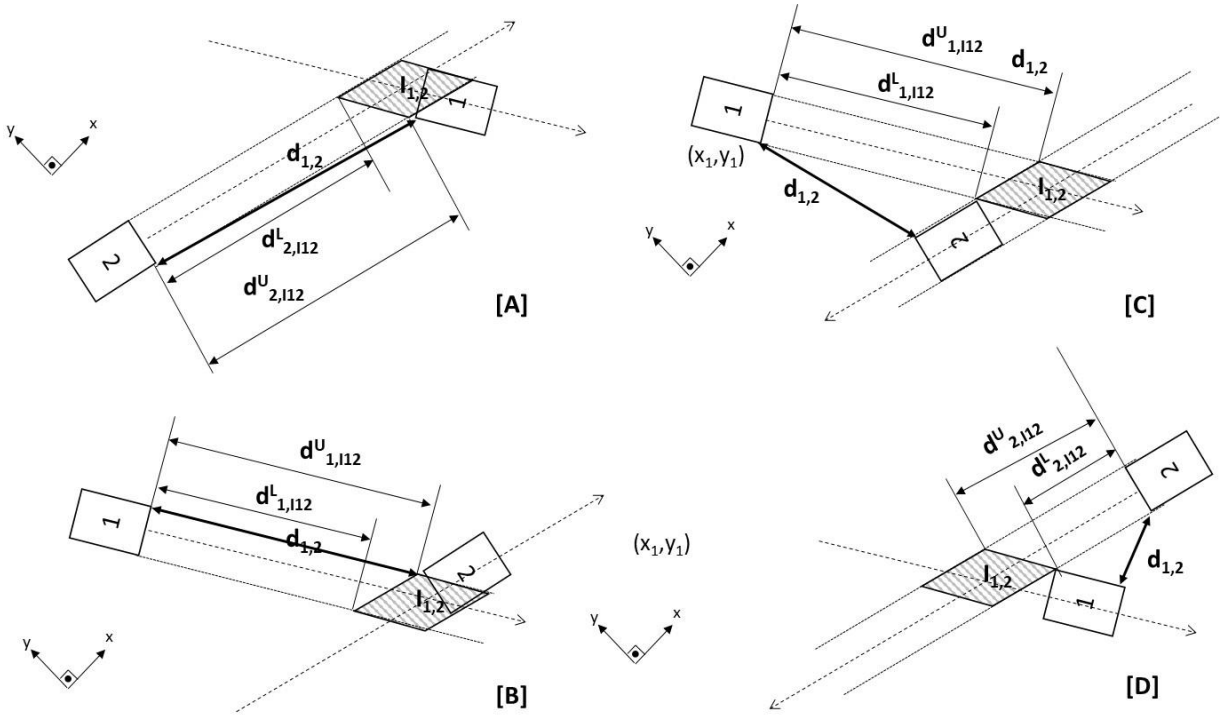


Figure 2. Processo de cálculo de distância mínima em trajetórias não ortogonais

Em muitos cenários de transporte, as características dos objetos (como suas dimensões) e suas trajetórias (incluindo coordenadas de cruzamento e ângulos) **são conhecidas desde o início da tarefa de monitoramento** (e, em aplicações de simulação, desde o momento da modelagem do cenário). Assim, as distâncias gerais $d^{U}_{i,Li,j}$ e $d^{L}_{i,Li,j}$ podem ser medidas indiretamente para cada objeto em função da **largura dos objetos (W_i, W_j)**, **ângulo de confluência da trajetória ($Ø_{i,j}$)** e **centro de confluência da trajetória** (centro geométrico de região de conflito $I_{i,j} – (x_{Ii,j}, y_{Ii,j})$), que são ilustrados na **Figura 8**.

Em aplicações virtuais, os parâmetros W_i, W_j, $Ø_{i,j}$, $(x_{Ii,j}, y_{Ii,j})$ são conhecidos desde o início da simulação. Assim, quase todos os parâmetros exigidos para calcular as distâncias (ou seja, arestas $I_{i,j}$, $(x_{i,Iij}, y_{i,Iij})$ e $(x_{j,Iij}, y_{j,Iij})$) podem ser definidos apenas uma vez, no momento em que o monitoramento tarefa começa. Durante o processo de cálculo da distância mínima, é necessário apenas obter (x_i, y_i) e (x_j, y_j) em cada amostra de monitoramento e calcular $d_{i,Lij}$ e $d_{j,Lij}$. Depois de calcular essas distâncias, todas as outras distâncias podem ser calculadas como consequência.

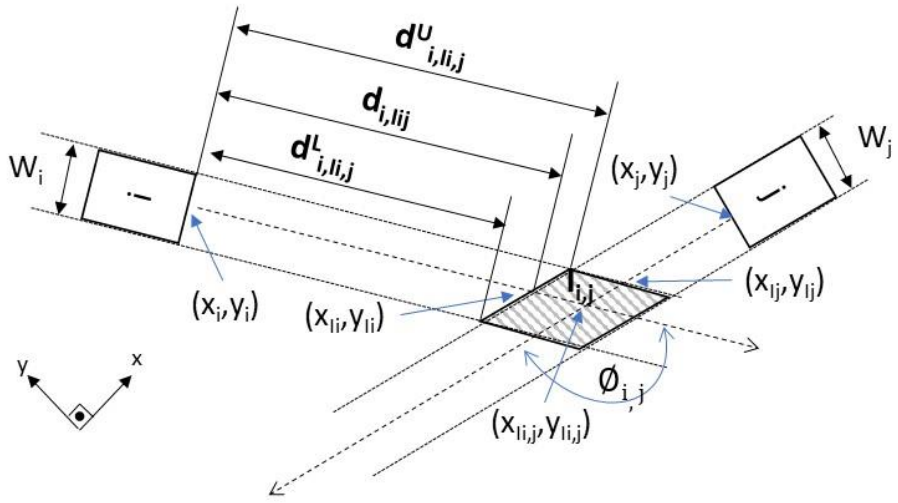


Figura 8. Equações gerais para cálculo de distância ($\mathbf{d^L_{i,Ii,j}}$, $\mathbf{d^U_{i,Ii,j}}$) em função das dimensões dos objetos (W_i, W_j), ângulo de confluência das trajetórias ($Ø_{i,j}$) e centro geométrico da região de conflito ($I_{i,j} – (x_{Ii,j}, y_{Ii,j})$)

A Eq.11 apresenta a relação entre parâmetros e as distâncias calculadas.

$$(\mathbf{d^L_{i,Ii,j}}, \mathbf{d^U_{i,Ii,j}}) = f(W_i, W_j, Ø_{i,j}, x_{Ii,j}, y_{Ii,j}) \quad \text{(Eq.11)}$$

Onde:

W_i, W_j = largura dos objetos;
$Ø_{i,j}$ = ângulo de confluência das trajetórias;
$(x_{Ii,j}, y_{Ii,j})$ = centro geométrico da região de conflito ($I_{i,j}$).

No entanto, em **cenários de transporte em que cada OoI é capaz de mudar sua própria trajetória durante o tempo de execução** (ou seja, **em sistemas de transporte não guiados**), as **regiões potenciais de conflitos precisam ser reavaliadas constantemente**, pois $(x_{Ii,j}, y_{Ii,j})$ e $Ø_{i,j}$ podem mudar ao longo do tempo. Assim, **em ambientes de movimento não guiado**, as bordas $I_{i,j}$, $(x_{i,Iij}, y_{i,Iij})$ e $(x_{j,Iij}, y_{j,Iij})$) devem ser calculadas em cada amostra de monitoramento. Portanto, espera-se que as vantagens deste novo método sejam potencializadas quando aplicada na estimativa de riscos de colisão em ambientes de movimentação guiada.

3. CONSIDERAÇÕES FINAIS A RESPEITO DO MÉTODO

Este trabalho desenvolveu um método que proporciona uma maneira mais eficiente, rápida e computacionalmente menos dispendiosa de calcular as distâncias mínimas entre objetos em ambientes de movimento guiado. Este método é aplicável a cenários reais e simulados, sendo aplicável em atividades de gerenciamento de riscos *runtime* (em tempo de operação do sistema) e *offline* (durante o desenvolvimento do sistema).

Este método apresenta a máxima eficiência computacional quando aplicado a ambientes móveis guiados. Nesses ambientes, os OoIs estão sempre se movendo ao longo de um conjunto conhecido de trajetórias estáticas (por exemplo, trilhos em sistemas ferroviários e estradas em sistemas de transporte rodoviário). Assim, a quantidade de possíveis regiões de interseção (#I) tende a ser muito menor do que a quantidade de OoI (#OoI) percorrendo essas trajetórias. Além disso, as interseções têm posições conhecidas, dado que as trajetórias são estáticas. Portanto, o método proposto pode promover uma redução significativa tanto na quantidade de elementos monitorados – de #OoI para #I – quanto nos esforços computacionais demandados, visto que as posições monitoradas são estáticas.

Dado que as posições das trajetórias e interseções são estáticas, a maioria dos parâmetros exigidos para o cálculo da distância mínima são sempre conhecidos em ambientes de movimento guiado. Desta forma, muitas tarefas computacionais não são exigidas quando este método é utilizado em ambientes móveis guiados. Portanto, o método proposto pode reduzir significativamente os esforços computacionais exigidos em cada tarefa de cálculo de distância. Além disso, este método pode promover uma redução significativa no número de distâncias calculadas, de $\binom{\#OoI}{2}$ para um valor muito menor que #I, visto que a tarefa de cálculo de distância é executada apenas quando um OoI sai de uma interseção.

ANEXO I. INFERÊNCIA DE MODELOS EXECUTIVOS UTILIZANDO APRENDIZADO DE MÁQUINA EM INTELIGÊNCIA ARTIFICIAL

Este anexo detalha o estudo realizado para avaliar a eficiência de diversas técnicas de classificação e de regressão em Aprendizado Supervisionado de Máquina no desenvolvimento do modelo executivo da Abordagem Proposta desenvolvida nesta tese. O objetivo deste estudo é identificar qual(is) técnica(s) de aprendizado supervisionado produzem o modelo mais eficiente tendo, como entrada do processo, uma base de dados com as características inerentes – sobretudo, pequena quantidade de dados – às produzidas pelo processo de simulação definido para o cenário-base (Figura 44).

Para cumprir este objetivo, o processo de desenvolvimento do modelo executivo por meio de ML/IA utilizou a **base de dados produzida na terceira campanha de simulação (Campanha III, Tabela 11)**. Cada uma das aplicações utilizou uma técnica distinta de aprendizado supervisionado, produzindo um modelo executivo. Ao final, os modelos produzidos foram avaliados e comparados, chegando-se à conclusão adotada por esta tese: as técnicas de **Árvore de Decisão** são a opção mais recomendada para o desenvolvimento do modelo executivo da Abordagem Proposta.

Dado que modelos de classificação (2 classes) foram utilizados nesse trabalho, a **íntegra** da avaliação realizada para as técnicas de classificação (2 classes) é apresentada. Para os modelos de classificação (3 classes) e de regressão, as tabelas detalhadas com as análise técnica-a-técnica foram suprimidas desta apresentação.

Este estudo é de autoria do pesquisador **Alexandre Moreira Nascimento**[60], ao qual agradecemos pela disponibilização do material.

[60] http://lattes.cnpq.br/6415796236777877

Avaliação da eficiência de técnicas de Aprendizado Supervisionado na inferência de modelos de AI/ML

Este estudo explora os resultados de diferentes tipos de técnicas de AI/ML com várias estratégias de indução e iterações. Um conjunto de dados de entrada foi usado para gerar 3 versões distintas de conjunto de dados: 2 classes, 3 classes e numérico. As duas primeiras versões (**2 classes** e **3 classes**) foram usadas para induzir **19 modelos** de classificação usando validação cruzada **10x10** (*10x10-fold cross-validation*). A **versão numérica** foi usada para induzir 29 modelos de regressão usando validação cruzada 10x10 (*10x10-fold cross-validation*). Isso resultou em 6.700 modelos induzidos.

A seguir, são apresentados os modelos de ML considerados neste estudo, bem como as estratégias usadas para treinar/validar os modelos de ML, as métricas de avaliação usadas para avaliar a qualidade (desempenho) dos modelos de ML induzidos, os conjuntos de dados aplicados para treinar/ testar os modelos e, finalmente, os resultados e discussão obtidos.

1. Modelos de Aprendizado de Máquina (ML) avaliados

Dois tipos de modelos de ML foram usados neste estudo: **modelos de classificação** e **modelos de regressão**. Os modelos de classificação foram usados para classificar configurações distintas de redes de comunicação (<F, L>) em classes de risco de segurança (alto risco, médio risco ou baixo risco). Inicialmente, esses modelos foram testados com 2 classes de risco (Alto e Baixo). No entanto, para ajudar a equilibrar o conjunto de dados, foram adotadas 3 classes de risco de segurança. Um total de **19 técnicas de ML de modelos de classificação** foram consideradas neste estudo. Eles estão listados com seus principais parâmetros na **Tabela 1**.

Tabela 1. Técnicas de Classificação avaliadas

Id	Tipo	Nome
[1]	bayes	NaiveBayes
[2]	functions	Logistic
[3]	functions	MultilayerPerceptron
[4]	lazy	LWL
[5]	meta	Bagging
[6]	meta	RandomCommittee
[7]	meta	RotationForest
[8]	meta	RandomSubSpace
[9]	rules	ConjunctiveRule
[10]	rules	DecisionTable
[11]	rules	PART
[12]	rules	Ridor
[13]	trees	DecisionStump
[14]	trees	HoeffdingTree
[15]	trees	J48
[16]	trees	RandomForest
[17]	trees	RandomTree

[18]	trees	REPTree
[19]	functions	MultilayerPerceptron

Os modelos de regressão foram utilizados para retornar a probabilidade de colisão quando uma AV está na região de cruzamento. Os modelos retornam uma probabilidade para cada configuração de rede de comunicação distinta (<F, L>) de acordo com a capacidade de abstração do modelo para aprender com o conjunto de dados. Um total de **29 técnicas** de regressão ML foram usadas neste estudo. Eles estão listados com seus principais parâmetros na **Tabela 2**.

Tabela 2. Técnicas de Regressão avaliadas

Id	Tipo	Nome
[1]	meta	AdditiveRegression
[2]	meta	AttributeSelectedClassifier
[3]	meta	Bagging
[4]	meta	CVParameterSelection
[5]	rules	ConjunctiveRule
[6]	rules	DecisionTable
[7]	trees	DecisionStump
[8]	functions	GaussianProcesses
[9]	misc	InputMappedClassifier
[10]	lazy	IBk
[11]	lazy	KStar
[12]	lazy	LWL
[13]	functions	MultilayerPerceptron
[14]	functions	MultilayerPerceptron
[15]	meta	MultiScheme
[16]	rules	M5Rules
[17]	trees	M5P
[18]	meta	RandomCommittee
[19]	meta	RandomSubSpace
[20]	trees	RandomForest
[21]	trees	RandomTree
[22]	trees	REPTree
[23]	meta	RandomizableFilteredClassif
[24]	meta	RegressionByDiscretization
[25]	functions	SimpleLinearRegression
[26]	meta	Stacking
[27]	meta	Vote
[28]	meta	WeightedInstancesHandlerWrapper
[29]	rules	ZeroR

A ferramenta Weka[61] 3.7 e 3.8 foi usada para as induções e validação de ML. As explorações foram realizadas com o módulo *Weka Explorer* e os experimentos com o módulo *Weka Experimenter*. Durante a fase de experimentação, os parâmetros do modelo e algumas métricas de desempenho foram validados qualitativamente. Assim, com exceção de um MLP, as configurações padrão foram mantidas.

2. Estratégias de Dataset Splitting para treino/teste de modelos de ML

Estratégias distintas para dividir o conjunto de dados em subconjuntos de treinamento e validação foram empregadas com o objetivo de testar o desempenho dos modelos de ML quando induzidos para cada configuração. A divisão do conjunto de dados em 50/50 significa que metade dos elementos do conjunto de dados foi selecionada aleatoriamente e usada durante a fase de treinamento do modelo, enquanto os elementos restantes foram usados na fase de validação do modelo, quando o desempenho do modelo é avaliado de acordo com as métricas descritas a seguir. Nenhuma partição *hold-out* (partição de teste) foi usada neste estudo **devido ao tamanho limitado do conjunto de dados** e devido ao objetivo de avaliar e selecionar o modelo de ML adequado.

A **Figura 1** ilustra 5 das 6 estratégias utilizadas. Para melhor avaliação do desempenho dos modelos, foram utilizadas 10 iterações para cada estratégia. Ou seja, por exemplo, em uma divisão 50/50, o processo de divisão, treinamento e avaliação do modelo foi executado 10 vezes e foi calculada a média dos valores das métricas de desempenho. Essa abordagem ajuda a aumentar a confiabilidade dos resultados ao suavizar os extremos. Evita a avaliação pessimista do pior cenário onde foi realizada pior divisão possível que resultou em um modelo de baixo desempenho e evita a avaliação otimista do melhor cenário onde foi realizada a melhor divisão possível que resultou em um modelo de alto desempenho. Consequentemente, para cada estratégia de divisão de conjunto de dados e cada técnica de ML, 10 modelos de ML foram induzidos.

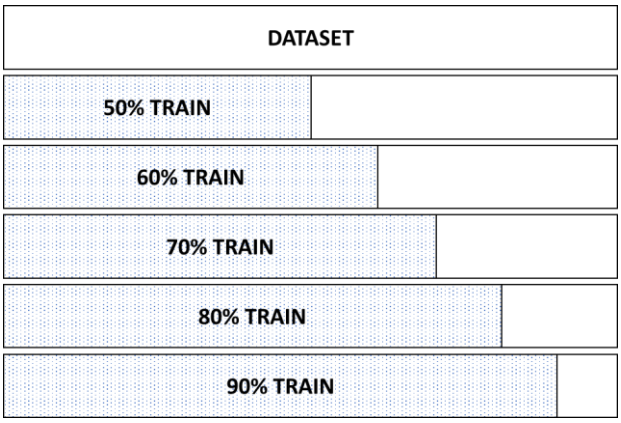


Figura 1. Estratégias de *Dataset Splitting*

É importante satisfazer ambas as condições no processo de divisão: (1) os conjuntos de dados devem ser grandes o suficiente para produzir resultados estatisticamente significativos; e (2) os conjuntos de dados são representativos do conjunto de dados como um todo. Em outras palavras, o conjunto de validação não pode ter características diferentes do conjunto de

[61] https://www.cs.waikato.ac.nz/ml/weka/

treinamento. **O tamanho do conjunto de dados utilizado como entrada deste estudo é menor do que o necessário para garantir a condição (1)**.

A outra estratégia utilizada é a validação cruzada k-fold (**Figura 2**). A validação cruzada é uma técnica para avaliar a capacidade de generalização (habilidades para prever dados não vistos) de um modelo de ML induzido a partir de um conjunto de dados específico. Essa técnica é amplamente empregada em problemas onde o objetivo da modelagem é a predição. Baseia-se em um procedimento de reamostragem para avaliar modelos de ML em uma amostra limitada de dados, que é o **caso da base de dados adotada neste estudo**. Ele ajuda a avaliar, comparar e selecionar um modelo de ML de um conjunto de modelos de ML induzidos para um determinado problema preditivo e resulta em estimativas de habilidade em dados não vistos que geralmente têm um viés menor do que outros métodos. A técnica possui um parâmetro conhecido por k que se refere ao número de grupos em que uma determinada amostra de dados deve ser dividida. O valor de k usado foi 10, ou seja, a validação cruzada de 10 vezes foi usada, resultando em 10 divisões distintas dos dados de treinamento, conforme ilustrado na **Figura 3**.

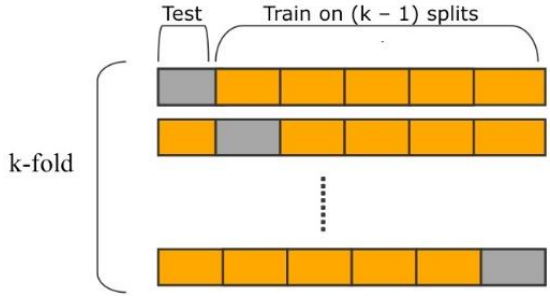


Figura 2. Estratégia de validação cruzada 'k-fold'
[Fonte: https://images.app.goo.gl/ntzPCGRcB4gDjGCJ6]

Figura 3. Estratégia "10-fold" (k = 10)
[Fonte: https://images.app.goo.gl/7LQLinpSAkypyMQr7]

3. Métricas de desempenho para avaliação de eficiência dos modelos

Conjuntos consideráveis de métricas foram utilizados para avaliar o desempenho dos modelos +de ML induzidos, listados na **Tabela 3**. À esquerda, são apresentadas as métricas de desempenho consideradas para os **modelos de classificação**. Na coluna da direita, são apresentadas as métricas consideradas para os modelos de regressão. O objetivo desta seção é apenas listar as métricas aplicadas, ao invés de explicá-las.

Tabela 3. Métricas de desempenho utilizadas neste estudo

Métricas de Desempenho	
Classificação	**Regressão**
Accuracy	Root Relative Squared Error
False Negative Rate	Mean Absolute Error
False Positive Rate	Root Mean Squared Error
Precision	Relative Absolute Error
Recall	Correlation Coefficient
F Measure	
Area Under ROC	
Area Under PRC	

4. Bases de dados de treinamento/validação aplicados

A característica "Taxa de Colisão" (CR) – uma variável numérica – do conjunto de dados de entrada do processo de desenvolvimento do modelo foi convertida em uma classe chamada "Risco de Acidente" (AR) para poder usar técnicas de classificação de ML. Experimentos também foram realizados com o uso de 2 classes (Alta/Baixa) e 3 classes (Alta/Média/Baixa). Portanto, tecnicamente, foram utilizadas 12 versões de conjuntos de dados, nas quais "Taxa de Atualização de Mensagens" (F), "Latência fim-a-fim" (L) e "Taxa de Colisão" (CR) - <F, L, CR> são os dados contidos nas bases de dados. Vale ressaltar que avaliações empíricas foram realizadas para estabelecer os critérios para converter o conjunto de dados numéricos em classes visando **minimizar a falta de balanceamento naturalmente imposta pela natureza dos experimentos** e pelo **pequeno tamanho do conjunto de dados**. Os limiares para discretização do conjunto de dados em 2 classes e 3 classes são apresentados na **Tabela 4** e na **Tabela 5**, respectivamente.

Tabela 4. Limiares para classificação de CR em AR (2 classes)

Classe de Risco	**Limiar**	**Quantidade de amostras**
Baixo	0.0093	31
Alto	1	29
Total		60

Tabela 5. Limiares para classificação de CR em AR (3 classes)

Classe de Risco	Limiar	Quantidade de amostras
Baixo	0.007	21
Médio	0.037	20
Alto	1	19
Total		60

A **Tabela 6**, mostra as características da base de dados de entrada (numérico). As estatísticas descritivas dos conjuntos de dados foram geradas pelo software estatístico R[62].

Tabela 6. Estatísticas descritivas da Base de Dados de entrada

Descriptive Statistic	Numérico			2 classes			3 classes		
	F	**L**	**AR**	**F**	**L**	**AR**	**F**	**L**	**AR**
nbr.val	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00	60.00
nbr.null	-	-	9.00	-	-	31.00	-	-	21.00
nbr.na	-	-	-	-	-	-	-	-	-
min	0.10	1.00	-	0.10	1.00	-	0.10	1.00	-
max	40.00	5,000.00	1.00	40.00	5,000.00	1.00	40.00	5,000.00	1.00
range	39.90	4,999.00	1.00	39.90	4,999.00	1.00	39.90	4,999.00	1.00
sum	716.00	38,706.00	9.61	716.00	38,706.00	29.00	716.00	38,706.00	29.00
median	5.50	45.00	0.01	5.50	45.00	-	5.50	45.00	0.50
mean	11.93	645.10	0.16	11.93	645.10	0.48	11.93	645.10	0.48
SE.mean	1.88	192.64	0.04	1.88	192.64	0.07	1.88	192.64	0.05
CI.mean.0.95	3.76	385.46	0.08	3.76	385.46	0.13	3.76	385.46	0.11
var	211.33	2,226,504.00	0.10	211.33	2,226,504.00	0.25	211.33	2,226,504.00	0.17
std.dev	14.54	1,492.15	0.32	14.54	1,492.15	0.50	14.54	1,492.15	0.41
coef.var	1.22	2.31	1.99	1.22	2.31	1.04	1.22	2.31	0.85

5. Resultados e Discussão

As métricas de desempenho foram calculadas para todos os modelos inferidos. Por se tratar de uma aplicação crítica de segurança, os resultados apresentados e discutidos estão relacionados apenas à indução de validação cruzada 10 x 10 vezes (ou seja, 10 avaliações completas de 10 vezes), onde as médias das métricas de desempenho são usadas no avaliação (t-Test). Essas são as avaliações mais rigorosas considerando todas as estratégias utilizadas e podem indicar melhor a capacidade de generalização dos modelos diante de dados novos.

[62] https://www.r-project.org/

5.1. Modelos de classificação (2 classes)

A **Tabela 7** apresenta uma visão geral dos resultados dos modelos de classificação ML induzidos a partir da base de dados de entrada e considerando **2 classes de risco** (baixo e alto riscos de segurança). A tabela apresenta os valores **mínimo** (Min), **máximo** (Max), **médio** e **desvio padrão** (DP) encontrados para cada métrica de desempenho avaliada calculada sobre os valores alcançados pelos modelos de ML em 10 iterações usando validação cruzada de 10 vezes. Ou seja, foram realizadas **100 induções por modelo**. Além disso, apresentam o **modelo que alcançou os valores mínimo e máximo para cada métrica de desempenho**. Assim, a tabela resume os principais resultados. Além disso, ajuda a avaliar qual modelo de ML obteve o melhor e o pior desempenho para cada métrica avaliada.

Tabela 7. Visão geral do desempenho dos modelos de classificação (2 classes)

Métrica de Desempenho (Tabela 3)	Min	Min Model ([ID], Tabela 1)	Max	Max Model ([ID], Tabela 1)	Média	DP
Accuracy	56	[1], [14]	78.3	[5]	68.93	5.707
False Negative Rate (FNR)	0	[4], [9], [13]	0.38	[2], [19]	0.1711	0.129
False Positive Rate (FPR)	0.27	[3]	0.76	[1], [14]	0.4605	0.164
Precision	0.55	[1], [14]	0.79	[5]	0.6953	0.067
Recall	0.62	[2],[19]	1	[4],[9],[13]	0.8289	0.129
F Measure	0.62	[1],[18]	0.82	[3]	0.7311	0.059
Area Under ROC	0.64	[9]	0.84	[5],[8]	0.7295	0.06
Area Under PRC	0.62	[9]	0.87	[5]	0.7426	0.074

É desejável que a maioria das métricas de desempenho seja maximizada. Já as taxas de Falso Positivo (FPR) e Falso Negativo (FNR), deseja-se que atinjam os valores mínimos. Precisão e Recall também são desejados para serem maximizados, onde a avaliação de sua combinação (F-Measure) oferece muitas vezes uma avaliação mais equilibrada. No entanto, neste domínio do problema, **falsos negativos resultarão em mais acidentes**, enquanto **falsos positivos aumentarão o uso da rede**. Embora o equilíbrio certo seja desejável, aumentar o consumo da rede para reduzir a taxa de acidentes é algo óbvio dentro de uma faixa razoável, que precisa ser avaliada pelas partes envolvidas e interessadas (provedor de serviços de comunicação, usuário do serviço de comunicação, autoridades reguladoras, etc).

É possível observar que não há predominância de um tipo de modelo sobre as métricas de desempenho. O modelo [5] tende a ser aquele com mais consistência, pois obteve a melhor Acurácia, Precisão, Áreas sob ROC e PRC. Sua precisão foi de cerca de 78%. Este modelo não atingiu o valor mínimo de TFN (0%), o que foi de fato alcançado pelos modelos [4], [9] e [13]. Apesar dos valores mínimos e máximos, vale a pena um olhar mais aprofundado sobre os resultados. As tabelas de **8** a **15** mostram os resultados do 'Teste-t' de todos os modelos usando o modo Modelo [1] como referência de avaliação para cada métrica de desempenho. Assim, cada modelo recebeu uma análise 'Test-t' para verificar a hipótese nula de cada métrica de desempenho.

A **Tabela 8** mostra os resultados da acurácia para todos os modelos. Os modelos [4], [5], [6], [7], [8], [11], [15], [17] e [18] alcançaram valores de precisão estatisticamente melhores do que a referência ([1]). De fato, os modelos [4] e [8] tiveram desempenhos bastante comparáveis na média quando comparados ao melhor ([5]).

Tabela 8. Modelos de classificação induzidos, 2 Classes, Acurácia

Modelos [1] a [10]									
[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
56	63.83	68.17	75.33v	78.33v	70.67v	71.33v	74.83v	65.67	67.67
(referência)	(=)	(=)	(>)	(>)	(>)	(>)	(>)	(=)	(=)
Modelos [11] a [19]									
[11]	**[12]**	**[13]**	**[14]**	**[15]**	**[16]**	**[17]**	**[18]**	**[19]**	
72.83v	71.17	68.33	56	73.00v	69.67	70.67v	72.00v	64.17	
(>)	(=)	(=)	(=)	(>)	(=)	(>)	(>)	(=)	

A **Tabela 9** apresenta os resultados da Área Sob RPC para todos os modelos. Os modelos [3], [4], [5], [8], [14] e [16] alcançaram valores estatisticamente melhores do que a referência. Os modelos [4], [5], [8] e [16] tiveram desempenhos bastante comparáveis em média.

Tabela 9. Modelos de classificação induzidos, 2 Classes, Área sob RPC

Models [1] to [10]									
[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
0.72	0.76	0.80v	0.84v	0.87v	0.7	0.71	0.84v	0.62	0.64
(referência)	(=)	(>)	(>)	(>)	(=)	(=)	(>)	(=)	(=)
Models [11] to [19]									
[11]	**[12]**	**[13]**	**[14]**	**[15]**	**[16]**	**[17]**	**[18]**	**[19]**	
0.71	0.69	0.64	0.80v	0.73	0.86v	0.7	0.71	0.77	
(=)	(=)	(=)	(>)	(=)	(>)	(=)	(=)	(=)	

A **Tabela 10** mostra os resultados da área sob ROC para todos os modelos. Os modelos [4], [5], [8] e [16] alcançaram valores estatisticamente melhores do que a referência. Esses modelos tiveram desempenho bastante comparável considerando os valores médios de ROC.

Tabela 10. Modelos de classificação induzidos, 2 Classes, ROC

Models [1] to [10]									
[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
0.67	0.66	0.74	0.80v	0.84v	0.71	0.73	0.84v	0.64	0.67
(referência)	(=)	(=)	(>)	(>)	(=)	(=)	(>)	(=)	(=)
Models [11] to [19]									
[11]	**[12]**	**[13]**	**[14]**	**[15]**	**[16]**	**[17]**	**[18]**	**[19]**	
0.74	0.71	0.67	0.72	0.77	0.82v	0.71	0.75	0.67	
(=)	(=)	(=)	(=)	(=)	(>)	(=)	(=)	(=)	

A **Tabela 11** mostra os resultados da Medida F (F-Measure) para todos os modelos. Os modelos [4], [5], [8] e [13] alcançaram valores estatisticamente melhores do que a referência. Esses modelos tiveram desempenho bastante comparável considerando a Medida F, em especial, os Modelos [4], [5] e [8].

Tabela 11. Modelos de classificação induzidos, 2 Classes, F-Measure

Models [1] to [10]									
[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
0.66	0.62	0.66	0.82v	0.81v	0.7	0.76	0.79v	0.76	0.77
(referência)	(=)	(=)	(>)	(>)	(=)	(=)	(>)	(=)	(=)
Models [11] to [19]									
[11]	**[12]**	**[13]**	**[14]**	**[15]**	**[16]**	**[17]**	**[18]**	**[19]**	
0.76	0.73	0.77v	0.66	0.76	0.7	0.7	0.77	0.62	
(=)	(=)	(>)	(=)	(=)	(=)	(=)	(=)	(=)	

A **Tabela 12** mostra os resultados da Taxa de Falsos Negativos (TFN) para todos os modelos. Os modelos [4], [9] e [13] alcançaram taxas estatisticamente melhores do que a referência. De fato, eles alcançaram valores de taxa de falso negativo de 0, o que significa que nenhum falso negativo na média foi encontrado, o que é um resultado bastante bom para o propósito do domínio de aplicação desejado.

Tabela 12. Modelos de classificação induzidos, 2 Classes, FNR

Models [1] to [10]									
[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
0.14	0.38v	0.37v	0.00*	0.11	0.31	0.1	0.1	0.00*	0.01
(referência)	(>)	(>)	(<)	(=)	(=)	(=)	(=)	(<)	(=)
Models [11] to [19]									
[11]	**[12]**	**[13]**	**[14]**	**[15]**	**[16]**	**[17]**	**[18]**	**[19]**	
0.16	0.19	0.00*	0.14	0.15	0.3	0.31	0.1	0.38v	
(=)	(=)	(<)	(=)	(=)	(=)	(=)	(=)	(>)	

A **Tabela 13** mostra os resultados da Taxa de Falsos Positivos (TFP) para todos os modelos. Com exceção dos modelos [9], [10], [13] e [14], todos os demais alcançaram índices estatisticamente melhores do que a referência. Os modelos [3], [6] e [17] obtiveram resultados bastante semelhantes, em torno de 28%. **Devido à natureza do problema atual sob pesquisa, os falsos positivos são preferidos aos falsos negativos, uma vez que o último resultaria em maiores taxas de acidentes**. Assim, parece que os modelos sem falsos negativos na média são modelos promissores.

Tabela 13. Modelos de classificação induzidos, 2 Classes, FPR

Models [1] to [10]									
[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
0.76	0.34*	0.27*	0.51*	0.33*	0.28*	0.48*	0.41*	0.71	0.66
(referência)	(<)	(<)	(<)	(<)	(<)	(<)	(<)	(=)	(=)
Models [11] to [19]									
[11]	**[12]**	**[13]**	**[14]**	**[15]**	**[16]**	**[17]**	**[18]**	**[19]**	
0.40*	0.39*	0.66	0.76	0.40*	0.31*	0.28*	0.47*	0.33*	

(<)	(<)	(=)	(=)	(<)	(<)	(<)	(<)	(<)

A **Tabela 14.** mostra os resultados da Precisão para todos os modelos. Os modelos [3], [4], [5], [6], [7], [8], [11], [12], [15], [16], [17] e [18] alcançaram valores estatisticamente melhor do que a referência. Os valores superiores foram alcançados por [5], [6] e [17], que tiveram valores bastante semelhantes.

Tabela 14. Modelos de classificação induzidos, 2 Classes, Precisão

Models [1] to [10]									
[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
0.55	0.70v	0.75v	0.70v	0.79v	0.77v	0.68v	0.74v	0.62	0.63
(referência)	(>)	(>)	(>)	(>)	(>)	(>)	(>)	(=)	(=)
Models [11] to [19]									
[11]	**[12]**	**[13]**	**[14]**	**[15]**	**[16]**	**[17]**	**[18]**	**[19]**	
0.73v	0.71v	0.64	0.55	0.73v	0.75v	0.77v	0.70v	0.7	
(>)	(>)	(=)	(=)	(>)	(>)	(>)	(>)	(=)	

A **Tabela 15** mostra os resultados do Recall para todos os modelos. Os modelos [4], [9] e [13] alcançaram 100% de Recall, o que é estatisticamente melhor do que a referência.

Tabela 15. Modelos de classificação induzidos, 2 Classes, Recall

Models [1] to [10]									
[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**	**[10]**
0.86	0.62*	0.63*	1.00v	0.89	0.69	0.9	0.9	1.00v	0.99
(referência)	(<)	(<)	(>)	(=)	(=)	(=)	(=)	(>)	(=)
Models [11] to [19]									
[11]	**[12]**	**[13]**	**[14]**	**[15]**	**[16]**	**[17]**	**[18]**	**[19]**	
0.84	0.81	1.00v	0.86	0.85	0.7	0.69	0.9	0.62*	
(=)	(=)	(>)	(=)	(=)	(=)	(=)	(=)	(<)	

Em relação aos resultados obtidos, as seguintes conclusões podem ser destacadas:

- **O tamanho do conjunto de dados importa**, e melhores resultados podem ser alcançados com conjuntos de dados maiores, gerados a partir da simulação de <F, L> cuidadosamente selecionados para tentar encontrar o comportamento das taxas de acidentes em alguns intervalos. Espera-se que, quanto maior o conjunto de dados, maior a resolução dos fenômenos e melhores resultados dos modelos inferidos por ML.

- Não houve resultados expressivos em termos de Precisão, o que reforça a primeira conclusão.
- Considerando que não há grande diferença nos resultados (desempenho dos modelos) em face a base de dados utilizada, **técnicas de árvore de decisão** são a melhor escolha para geração dos **modelos, sobretudo** por sua **explicabilidade**, o que ajuda a entender os fenômenos e a verificar os resultados dos modelos inferidos em relação aos gráficos

do conjunto de dados, podendo entender visualmente as regiões de tomada de decisão que o modelo executivo decidiu operar quando está operando.

5.2. Modelos de classificação (3 classes)

A **Tabela 16** apresenta uma visão geral dos resultados dos modelos de classificação ML induzidos a partir da base de dados de entrada e considerando **3 classes de risco** (baixo, médio e alto riscos de segurança). A tabela apresenta os valores **mínimo** (Min), **máximo** (Max), **médio** e **desvio padrão** (DP) encontrados para cada métrica de desempenho avaliada calculada sobre os valores alcançados pelos modelos de ML em 10 iterações usando validação cruzada de 10 vezes. Ou seja, foram realizadas **100 induções por modelo**. Além disso, apresentam o **modelo que alcançou os valores mínimo e máximo para cada métrica de desempenho**. Assim, a tabela resume os principais resultados. Além disso, ajuda a avaliar qual modelo de ML obteve o melhor e o pior desempenho para cada métrica avaliada.

Tabela 16. Visão geral do desempenho dos modelos de classificação (3 classes)

Performance Metric (Tabela 3)	Min	Min Model ([ID], Tabela 1)	Max	Max Model ([ID], Tabela 1)	Average	Std
Accuracy	48	[1]	71.83	[15]	60.351	8.4263
False Negative Rate	0.01	[1]	0.75	[13]	0.2395	0.2122
False Positive Rate	0	[13]	0.99	[15]	0.4684	0.2765
Precision	0.06	[3]	0.96	[1]	0.6405	0.2317
Recall	0.01	[19]	1	[13]	0.5316	0.2765
F Measure	0.07	[3]	0.82	[15]	0.6137	0.1797
Area Under ROC	0.4	[3]	0.88	[15]	0.7047	0.127
Area Under PRC	0.41	[9]	0.82	[16]	0.6311	0.1316

Nesta tabela, é possível observar que há predominância de um tipo de modelo sobre as métricas de desempenho. Em primeiro lugar, os **modelos baseados em árvore** tiveram um desempenho superior perceptível e o modelo [15] demonstrou alguma consistência em algumas métricas de desempenho consideradas. No entanto, o modelo [15] teve desempenho ruim para a Taxa de Falsos Positivos. O modelo [1] demonstrou um desempenho muito bom para taxa de falsos positivos (próximo de zero), enquanto o modelo [13] teve um desempenho ruim para essa métrica.

[N.A.] foram aplicadas sobre as métricas dos modelos de classificação (3 classes) as mesmas avaliações métrica-a-métrica aplicadas nos modelos de classificação (2 classes). Dado que modelos de classificação (2 classes) foram utilizados neste trabalho, as tabelas comparativas para os modelos de classificação (3 classes) foram suprimidas.

Em relação aos resultados obtidos, as seguintes conclusões podem ser destacadas:

- O tamanho do conjunto de dados: assim como nos modelos de classificação (2 classes), o tamanho da base de dados é importante, e melhores resultados podem ser alcançados com conjuntos de dados maiores, gerados a partir da simulação de <F,L>

cuidadosamente selecionados para tentar encontrar o comportamento das taxas de acidentes em alguns intervalos. Espera-se que, quanto maior o conjunto de dados, maior a resolução dos fenômenos e melhores resultados dos modelos inferidos por ML.

- Não houve resultados expressivos em termos de Precisão, o que reforça a primeira conclusão.

- Considerando que não há grande diferença nos resultados (desempenho dos modelos) em face a base de dados utilizada, **técnicas de árvore de decisão** são a melhor escolha para geração dos **modelos, sobretudo** por sua **explicabilidade**, o que ajuda a entender os fenômenos e a verificar os resultados dos modelos inferidos em relação aos gráficos do conjunto de dados, podendo entender visualmente as regiões de tomada de decisão que o modelo executivo decidiu operar quando está operando.

5.3. Modelos de regressão

A **Tabela 17** apresenta uma visão geral dos resultados dos modelos de regressão ML induzidos a partir do conjunto de dados de entrada. A tabela apresenta os valores **mínimo** (Min), **máximo** (Max), **médio** e **desvio padrão** (DP) encontrados para cada métrica de desempenho avaliada calculada sobre os valores alcançados pelos modelos de ML em 10 iterações usando validação cruzada de 10 vezes. Ou seja, foram realizadas **100 induções por modelo**. Além disso, apresentam o **modelo que alcançou os valores mínimo e máximo para cada métrica de desempenho**. Assim, a tabela resume os principais resultados. Além disso, ajuda a avaliar qual modelo de ML obteve o melhor e o pior desempenho para cada métrica avaliada.

Tabela 17. Visão geral do desempenho dos modelos de regressão

Performance Metric (Tabela 3)	Min	Min Model ([ID], Tabela 2)	Max	Max Model ([ID], Tabela 2)	Average	Std
Root Relative Squared Error	20.07	[18], [21]	100.00	[4], [9], [15], [26], [27], [28], [29]	58.95	26.63
Mean Absolute Error	0.03	[18], [20], [21]	0.22	[4], [9], [15], [26], [27], [28], [29]	0.11	0.07
Root Mean Squared Error	0.06	[2], [18], [20], [21]	0.29	[4], [9], [15], [26], [27], [28], [29]	0.15	0.08
Relative Absolute Error	13.45	[18], [21]	100.00	[4], [9], [15], [26], [27], [28], [29]	53.53	29.72
Correlation Coefficient	-	[4], [9], [15], [26], [27], [28], [29]	0.91	[18], [21]	0.50	0.33

Diferentemente dos resultados dos modelos induzidos com por classificação de 2 classes e 3 classes, muitos modelos atingiram os valores mínimos e máximos encontrados para cada métrica de desempenho. Além disso, diferentemente das métricas de desempenho consideradas anteriormente para modelos de classificação, deseja-se que os modelos atinjam os valores mínimos possíveis para a maioria das métricas de desempenho aqui consideradas. Uma exceção é para a métrica de **correlação**, que se deseja que seja a mais alta possível.

Os valores máximos dos erros, indesejados, foram obtidos pelos modelos [4], [9], [15], [26], [27], [28], [29]. Os valores mínimos dos erros foram consistentemente alcançados pelos

modelos [18] e [21], que são respectivamente uma técnica de meta-aprendizagem usando algoritmo baseado em árvore e uma técnica baseada em árvore. A mesma consistência também pode ser notada para a métrica do coeficiente de correlação, onde se deseja o valor máximo. Para esta métrica, os modelos [4],[9],[15],[26],[27],[28],[29] obtiveram o pior desempenho (0) e o melhor desempenho foi alcançado também pelos modelos [18] e [21]. **O modelo [20] também apresentou o melhor desempenho para erros**, com exceção do Root Relative Squared Error. Vale ressaltar que **o modelo [20] é baseado em árvore, o que agrega o valor da explicabilidade**.

[N.A.] foram aplicadas sobre as métricas dos modelos de regressão as mesmas avaliações métrica-a-métrica aplicadas nos modelos de classificação (2 classes). Dado que modelos de classificação (2 classes) foram utilizados neste trabalho, as tabelas comparativas detalhadas para os modelos de regressão foram suprimidas.

Em relação aos resultados obtidos, pode-se concluir que os **modelos baseados em árvore de decisão** também **demonstraram resultados bons e consistentes quando usados como técnicas de regressão**. Este é um resultado conveniente, pois eles são fáceis de entender e explicam suas saídas com base nas entradas para apoiar a tomada de decisão. Em termos de aplicações críticas de segurança, os **modelos white-box são obrigatórios** para a sua validação e certificação de segurança. Além disso, os modelos baseados em regras também são úteis para esses tipos de aplicativos. Por esta razão, há pesquisas consideráveis na extração de regras para modelos de caixa-preta.

6. Considerações Finais

Os modelos baseados em árvore de decisão demonstraram resultado bom e consistente ao longo das avaliações apresentadas. Este é um resultado conveniente, pois eles são fáceis de entender e explicam suas saídas com base nas entradas para apoiar a tomada de decisão. Em termos de aplicações críticas de segurança, os modelos caixa-branca (*white-box*) são obrigatórios para a sua validação e certificação de segurança. Além disso, os **modelos baseados em regras** também são úteis para esses tipos de aplicativos. Por esta razão, existem pesquisas consideráveis na extração de regras para modelos caixa-preta.