UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA

ROGÉRIO BRITO RAMOS

Metodologias de análise integrada de segurança crítica e segurança cibernética em sistemas ciber físicos

São Paulo
2024

ROGÉRIO BRITO RAMOS

Metodologias de análise integrada de segurança crítica e segurança cibernética em sistemas ciber físicos

Versão Corrigida

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

Área de Concentração:
Engenharia de Computação

Orientador:
Prof. Dr. João Batista Camargo Júnior

São Paulo

2024

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

São Paulo, 5 de junho de 2024

Assinatura do autor: gov.br Documento assinado digitalmente ROGERIO BRITO RAMOS Data: 05/06/2024 19:22:40-0300 Verifique em https://validar.iti.gov.br

Assinatura do orientador: gov.br Documento assinado digitalmente JOAO BATISTA CAMARGO JUNIOR Data: 06/06/2024 14:01:29-0300 Verifique em https://validar.iti.gov.br

Catalogação-na-publicação

Ramos, Rogério Brito
Metodologias de análise integrada de segurança crítica e segurança cibernética em sistemas ciber físicos / R. B. Ramos -- versão corr. -- São Paulo, 2024.
156 p.

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

1. Engenharia de sistemas de computação 2.Gestão da segurança em sistemas computacionais 3.Computação aplicada (análise comparativa) 4.Pesquisa operacional I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Computação e Sistemas Digitais II.t.

Nome: RAMOS, Rogério Brito.

Título: Metodologias de análise integrada de segurança crítica e segurança cibernética em sistemas ciber físicos.

Dissertação apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

Aprovado em: 25/04/2024

Banca Examinadora

Membro: Prof. Dr. João Batista Camargo Júnior

Instituição: Escola Politécnica da Universidade de São Paulo (USP)

Julgamento: Aprovado

Membro: Prof. Dr. Valter Fernandes Avelino

Instituição: Faculdade de Engenharia Industrial (FEI)

Julgamento: Aprovado

Membro: Dr. Mário Aparecido Corrêa

Instituição: Avibras Indústria Aeroespacial S/A

Julgamento: Aprovado

À minha filha Rebecca, minha fonte de inspiração e felicidades.

AGRADECIMENTOS

Agradeço a Deus por todas as bençõe concedidas durante o curso, em especial, pelo nascimento da minha filha, Rebecca, que veio com muita saúde, trazendo ainda mais felicidades para a minha vida.

À minha esposa Natalia por todo incentivo dado e pela compreensão dos momentos em que tive que me abdicar do tempo com a família devido às demandas do curso.

À minha mãe Conceição, à minha irmã Jussara, e à Vanilde, mãe da minha esposa, expresso profunda gratidão por todo o carinho e apoio recebido.

Agradeço à Marinha do Brasil por ter me designado para esta missão profissional, na qual sou grato e me orgulho muito pelo voto de confiança.

À Escola Politécnica da USP que com sua excelência de ensino me proporcionou uma significativa evolução acadêmica.

Ao meu orientador, Prof. Dr. João Batista, por ter dedicado seu tempo na direção deste estudo provendo as instruções que me guiaram nesta campanha.

Ao Prof. Dr. Jorge Rady, ao Dr. Mário e ao Prof. Dr. Valter Avelino por terem aceitado a participação nas bancas em que os comentários foram fundamentais para conclusão e aprimoramento deste trabalho.

Aos meus antigos encarregados na Marinha, Capitão de Fragata Raquel Castilho e Capitão de Corveta Leandro Tolentino, pelo incentivo e apoio fornecido nas fases iniciais do curso.

Ao Capitão de Corveta Renato Cézar pelo acompanhamento desta fase junto a Marinha e pelo suporte das atividades da pesquisa.

Aos meus professores das disciplinas cursadas, Prof. Dra. Anarosa e Prof. Dra. Regina Silveira do Departamento de Engenharia da Computação e Sistemas Digitais (PCS), Prof. Dr. Max Luppe da Escola de Engenharia da São Carlos (EESC) e Prof. Dr. Newton Maruyama do Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos (PMR).

À Prof. Me. Priscila Hayama do Centro de Línguas da Faculdade de Filosofia, Letras e Ciências Humanas (FFLCH) pela minha preparação nos cursos de inglês que foram imprescindíveis para o êxito da apresentação do artigo em Congresso Internacional realizado no Reino Unido.

Aos meus professores de graduação da Universidade Federal do Amazonas (UFAM), Prof. Dr. Raimundo Barreto, Prof. Dr. José Luiz Pio e Prof. Dr. Cícero Ferreira Fernandes Costa Filho, pelos votos de confiança ao terem encaminhado as cartas de recomendação para ingresso neste curso de mestrado.

Aos meus colegas do Grupo de Análise de Segurança (GAS), Lucio Vismari, Daniel Baraldi, Ricardo Franco e Tiago Demay, pela oportunidade de convívio e pelos inúmeros esclarecimentos de questões relacionadas ao curso.

Aos meus colegas, Diego Miranda, Adriano Sousa e Wilyton Machado, pelas recomendações valiosas na preparação para o processo seletivo e para uma boa realização do curso.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro para participação de congresso científico internacional.

Por fim, quero estender os agradecimentos para todos os meus familiares e amigos que acompanharam e torceram para o sucesso desta missão. Muito obrigado!

“Podem ser encontrados aspectos positivos até nas situações negativas e utilizar isso como experiência para o futuro...”

Ayrton Senna

RESUMO

RAMOS, R. B. **Metodologias de análise integrada de segurança crítica e segurança cibernética em sistemas ciber físicos**. 2024. Dissertação de Mestrado. Escola Politécnica da Universidade de São Paulo. São Paulo, 2024.

Ativos de tecnologia da informação têm sido gradativamente incorporados em aplicações de domínio crítico. Na indústria marítima, na aviação e nos sistemas ferroviários, por exemplo, é possível encontrar sistemas ciber físicos, nos quais componentes digitais são integrados a dispositivos mecânicos. Como consequência, navios, aeronaves e trens tornam-se suscetíveis a ataques cibernéticos com interferência em suas ações físicas, podendo resultar em acidentes. Em projetos de alta criticidade, são geralmente obedecidas normas que preconizam a realização de análises de segurança crítica, conhecida como *safety,* voltadas para proteção contra falhas não intencionais, e também análises de segurança cibernética, conhecida como *security* ou *cybersecurity*, voltadas para proteção contra ataques maliciosos. No entanto, ambas as análises costumam ser realizadas de forma separada, utilizando diferentes ferramentas e por equipes distintas com pouca ou até nenhuma interação. Essa disjunção nas análises dificulta a detecção de conflitos entre as áreas de segurança crítica e de segurança cibernética. Visto que é possível que uma modificação para mitigar uma vulnerabilidade cause um aumento na probabilidade de uma falha acidental e, de forma recíproca, uma alteração com o intuito de reduzir a probabilidade de uma falha produza uma vulnerabilidade. Portanto, as análises integradas emergem como soluções para este problema. Esta pesquisa visa avaliar e comparar a eficiência das principais metodologias de análise integrada de segurança crítica e segurança cibernética, além disso, propor uma estratégia para condução de análise. Os resultados obtidos indicam que algumas metodologias foram úteis na identificação de conflitos que provavelmente passariam despercebidos em análises distintas. Como oportunidade para trabalhos futuros, são apontados o desenvolvimento de métodos para estimar probabilidades de ataques cibernéticos e de ferramentas de simulações para metodologias de análise.

Palavras-chave: Segurança Crítica. Segurança Cibernética. Sistemas Ciber Físicos (CPS). Análise Integrada. Sistemas Críticos.

ABSTRACT

RAMOS, R. B. **Methodologies of safety and security integrated analysis for cyber physical systems**. 2024. Master's Dissertation. Polytechnic School of the University of Sao Paulo. Sao Paulo, 2024.

Information technology assets have been gradually deployed into critical applications. In the maritime industry, aviation and railway systems, for example, cyber-physical systems are increasingly prevalent, with digital components integrated into mechanical devices. As a consequence, vessels, aircrafts and trains become susceptible to cyber-attacks, which can interfere their physical actions, potentially resulting in accidents. In high-criticality projects, according to most standards, both safety analysis, focused on safeguarding a system against unintentional failures, and cybersecurity analysis, aimed at protecting a system against malicious attacks, are mandatory. However, both analyses are typically performed separately, using different tools and conducted by different teams, often with limited or any interaction between them. These disjointed analyses make it difficult to detect conflicts between safety and security disciplines. Whereas, it is possible that a change deployed to fix a cyber vulnerability could increase the likelihood of an accidental failure, and conversely, a measure intended to increase the safety level could introduce a new vulnerability. Therefore, integrated analyses emerge as a solution to this problem. This study aims to assess and compare the effectiveness of integrated methodologies, which encompass safety and security analyses, as well as to propose a strategy to perform integrated analysis. The achieved results indicate that certain methodologies were effective in identifying conflicts that would have probably gone unnoticed in distinct analyses. As opportunities for future works, it is suggested to explore the development of methods for estimating the probabilities of cyber-attacks and to enhance the analysis tools for the methodologies assessed.

Keywords: Safety. Security. Cybersecurity. Cyber Physical Systems (CPS). Integrated Analysis. Critical Systems.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

AA	*Attack Action*
AIS	*Automated Information System*
ATCO	*Air Traffic Controller*
BDMP	*Boolean Logic Driven Markov Processes*
CCECOE	*NATO Cooperative Cyber Defense Centre of Excellence*
CCF	*Common Cause Failures*
CENELEC	*European committee for electrotechnical standardization*
CHASSIS	*Combined Harm Assessment of Safety and Security for Information Systems*
COMDCIBER	Comando de Defesa Cibernética
CPS	*Cyber-Physical System*
CVSS	*Common Vulnerability Scoring System*
D-MUC	*Misuse Case Diagram*
D-UC	*Use Case Diagram*
ESREL	*European Safety and Reliability Conference*
EUROCAE	*European Organisation for Civil Aviation Equipment*
FMECA	*Failures Mode Effects and Critical Analysis*
FMVEA	*Failure Mode, Vulnerabilities and Effects Analysis*
FSD	*Failure Sequence Diagram*
FTA	*Fault Tree Analysis*
GSN	*Goal Structure Notation*
GTST-MLD	*Goal Tree Success Tree Master Logic Diagram*
HAZOP	*Hazards and Operability study*
IAEA	*International Atomic Energy Agency*
IEC	*International Electrotechnical Commission*
IMO	*International Maritime Organization*
IOT	*Internet of things*
ISA	*International Society of Automation*
ISE	*Instantaneous Security Event*
ISID	*Industrial Security Incidents Database*
LS	*Loss Scenarios*
MA	*Markov Analysis*

MBSE	*Model-Based System Engineering*
MCS	*Minimal Cut Sets*
MISP	*Malware Information Sharing Platform*
MSC	*Maritime Safety Committee*
MUSD	*Misuse Sequence Diagram*
NATO	*North Atlantic Treaty Organization*
NSS	*Nuclear Security Series*
NIST	*National Institute of Standards and Technology*
RISI	*Repository of Industrial Security Incidents*
RTCA	*Radio Technical Committee for Aeronautics*
SBTA	*Brazilian Society of Air Transportaton Research*
SD	*Sequence Diagram*
SCADA	Sistemas de controle, supervisão e aquisição de dados
SL	*Security Level*
SIL	*Safety Integrity Level*
STAMP	*Systems-Theoretic Accident Model and Processes*
STPA	*System-Theoretic Process Analysis*
STPA-Sec	*System-Theoretic Process Analysis for Security*
T-MUC	*Textual MisUse Case*
T-UC	*Textual Use Case*
TSE	*Timed Security Event*
TWR	*Aerodrome Control Tower*
UCA	*Unsafe Control Actions*
UAV	*Unmanned Aerial Vehicles*
UML	*Unified Modeling Language*
VANT	Veículos Aéreos Não Tripulados
YAMS	*Yet Another Monte Carlo Simulation Tool*

SUMÁRIO

1 INTRODUÇÃO

Devido a diversas razões estratégicas de negócios, as aplicações empregadas em domínios críticos de segurança tendem a incorporar em seus projetos sistemas digitais cada vez mais integrados com ações mecânicas. O uso de tecnologias emergentes como dispositivos *Internet-Of-Thing* (IOT), arquitetura em nuvem, comunicações sem fio, equipamentos com novos tipos de *hardware*, já fazem parte do contexto dos sistemas críticos atuais e tem seu uso progressivamente ampliado como notadamente na aviação, na indústria marítima, em usinas nucleares e em sistemas ferroviários.

Essa migração de conceitos em projetos aplicados a infraestruturas críticas com incorporação de tais tecnologias os categorizam como Sistemas Ciber Físicos, ou *Cyber-Physical System* (CPS) onde dados e comandos digitais resultam em ações mecânicas (GILL apud GUZMAN; KOZINE; LUNDTEIGEN, 2021). Além de falhas acidentais, os CPS abrem a possibilidade para que ações cibernéticas maliciosas interfiram em operações do sistema, podendo resultar em catástrofes. Os métodos de análise de segurança crítica, para serem efetivos, devem levar também em consideração os riscos das vulnerabilidades de sistemas de tecnologia da informação. Esse estudo, portanto, consiste em abordar técnicas de análise integrada de aspectos de vulnerabilidades cibernéticas e de segurança crítica de sistemas.

1.1 CONTEXTUALIZAÇÃO

A principal preocupação em projetos de sistemas críticos é garantir que consequências catastróficas não ocorram. A proteção da vida e da integridade física de pessoas, a ausência de danos ao meio ambiente e a capacidade de evitar prejuízos financeiros significativos são requisitos que se tornam mais prioritários do que até mesmo a própria utilidade funcional do sistema. Nesses domínios de aplicação, as análises de segurança crítica devem ser realizadas de forma extremamente rigorosas para que sejam captadas todas as condições de perigo que possam levar a situações indesejáveis. Ainda mesmo que rígidos, as metodologias de análise de segurança sempre esbarram no desafio de acompanhar as evoluções tecnológicas dos sistemas, os quais apresentam uma evolução exponencial em comparação com sistemas puramente mecânicos ou eletrônicos. Ao mesmo tempo que essas evoluções trazem

ganhos de negócios como melhorias no desempenho, otimização de custo e mão de obra e aumenta o nível de competitividade das organizações, também trazem complicações para previsibilidade do comportamento do sistema que podem não ser cobertos por métodos de análises de segurança tradicionais (BUTTGEREIT et al. 2021). Além disso, traz a necessidade de análise em uma nova dimensão que se trata do contexto cibernético.

Tradicionalmente, análise de segurança crítica, voltada para proteção contra acidentes, e análise de segurança cibernética, voltada para proteção contra ataques intencionais, são realizadas de forma separada, por diferentes ferramentas e abordagens, por diferentes equipes de especialistas com pouca ou até nenhuma interação (KRIAA; BOUISSOU; LAAROUCHI, 2015). No entanto, com o aumento da dependência em que as plataformas críticas passam a ter de sistemas computacionais, torna-se necessário que as abordagens das análises adentrem nas especificidades em que um ativo de tecnologia da informação possa influenciar nos requisitos de segurança crítica do sistema. Portanto, analisar segurança crítica e segurança cibernética de forma conjunta torna-se fundamental para que se possa obter uma eficiente identificação de problemas correlacionados entre ambas as áreas. Além disso, existem situações de conflitos em que uma medida aplicada para reforçar uma área pode prejudicar a outra, e tais situações apenas são possíveis de serem identificadas se as análises forem feitas de forma conjunta.

Algumas ações que visam elevar o nível de proteção de segurança cibernética podem impactar negativamente no nível de segurança crítica de um sistema. Podemos citar como exemplo, a implementação de processos de autenticação e de criptografia na comunicação entre módulos de um sistema crítico. Ao mesmo tempo, em que tais implementações aumentam o nível segurança cibernética, também aumentam a latência no sistema, já que despejam mais dados no canal de comunicação e exigem mais processamento para validar uma troca de informação entre os módulos. Esses efeitos também reduzem o tempo de resposta e aumentam as chances de indisponibilidade temporária já que também podem invalidar uma troca de informação legítima. O impacto pode ser refletido no atendimento a requisitos críticos de um sistema de tempo real que preza pela garantia de execução de tarefas em prazos estritamente rigorosos. Em um sistema que tem como propósito evitar colisão de um trem com obstáculos no trilho, o

processamento da detecção do obstáculo até a ação de frear o trem deve ser feito no menor intervalo de tempo possível. Nesse cenário, qualquer implementação que aumente a latência impactará nas chances de êxito do trem não colidir. Por outro lado, manter a comunicação clara e direta, sem criptografia e sem autenticação, pode expor o sistema a indivíduos com habilidades cibernéticas e mal intencionados, os quais podem introduzir falsos obstáculos. Isso pode forçar o trem a frear indevidamente, resultando prejuízos à utilidade funcional e até mesmo provocando acidentes.

Similarmente, ações implementadas para elevar o nível de segurança crítica contra acidentes podem trazer e expor novas vulnerabilidades facilitando ações maliciosas por terceiros. Um exemplo mencionado por Sun apud Kriaa (2016) afirma que um modelo específico de carro de um fabricante europeu chegou a ser alvo de inúmeros furtos de forma extremamente desproporcional em relação aos demais modelos de veículos. Esse mistério apenas foi compreendido quando um suspeito detido revelou que as portas sempre destravavam ao pular em cima do teto do carro. Os projetistas haviam incluído uma funcionalidade de segurança crítica que permitia as portas destravarem se o carro se envolvesse em um acidente e capotasse. A intenção era facilitar que os ocupantes do veículo escapassem ou fossem retirados mais facilmente em um resgate. Tal medida de segurança crítica gerou uma vulnerabilidade que foi explorada por um agente malicioso. Situações equivalentes a esta também se aplicam no contexto cibernético.

1.2 OBJETIVOS DA PESQUISA

Objetivo Geral

Apresentar e comparar os resultados obtidos ao aplicar as principais metodologias de análise integrada de aspectos de segurança crítica e de segurança cibernética em estudos de casos de sistemas ciber físicos.

Objetivos Específicos

Avaliar a eficiência das metodologias em identificar vulnerabilidades, situações de perigo e de conflitos entre medidas de segurança crítica e de segurança cibernética.

Correlacionar as metodologias mais adequadas aos requisitos de um determinado tipo de análise. As associações entre as características das metodologias e os objetivos dos resultados indicam que uma análise mais eficiente pode depender de uma seleção assertiva de uma ou mais metodologias aplicadas.

Propor uma estratégia de condução de análise integrada de segurança crítica e segurança cibernética, visando obter um levantamento mais consistente das informações de entradas e facilitar a seleção das metodologias.

1.3 MOTIVAÇÕES E JUSTIFICATIVAS

O ataque do vírus *Stuxnet* em usinas nucleares em meados de 2010 repercutiu mundialmente sendo considerado um marco na história da segurança de infraestruturas críticas. Pela primeira vez até então, um vírus de computador era empregado para causar impacto em um meio físico. O vírus havia sido projetado para se infiltrar em sistemas que controlam processos industriais, onde se espalhava principalmente através de dispositivos USB, aproveitando-se de vulnerabilidades do sistema operacional. Uma vez dentro de uma rede, ele tentava se propagar para outros sistemas. O ataque explorou uma vulnerabilidade de um Controlador Lógico Programável (PLC) onde conseguia alterar as velocidades das centrífugas de enriquecimento de urânio e ainda disponibilizava um *feedback* adulterado. O *Stuxnet* abriu caminho para uma maior conscientização sobre a necessidade de proteger sistemas de controle industrial contra ameaças cibernéticas. Falliere, Murchu e Chien (2011) elaboraram um relatório detalhado sobre o *Stuxnet* e Kriaa, Bouissiu e Pièrre-Cambacedes (2012) realizaram uma modelagem do ataque.

Shipunov et al. (2021) mencionaram que ataques cibernéticos afetaram empresas de transporte marítimo da empresa Maersk e da China *Ocean* e também da Guarda Costeira dos Estados Unidos. Tais ataques resultaram em perdas financeiras expressivas e provocaram exposição da tripulação. Os ataques tiveram como efeitos a alteração de velocidades das embarcações e falsificação de *feedbacks* dos sistemas de navegação. Ashraf et al. (2023) menciona uma ocorrência de falsificação de dados do Sistema de Identificação Automática (AIS) que é responsável por fornecer segurança de navegação no mar e prevenção de colisões. O invasor reproduzia transmissões AIS se passando por autoridade portuária e direcionava os navios para a rota desejada.

Em 2017, a Ucrânia sofreu com os ataques *BadRabbit* que deixou o aeroporto inoperante e o ataque *NotPetya* que paralisou as operações do metrô de Kiev (ALOTAIBI; VASSILAKIS, 2021). Ambos os ataques são do tipo *ransomware*, que criptografa arquivos tornando inacessíveis e há uma exigência de pagamentos para liberar o arquivo ao usuário (MOS; CHOWDHURY, 2020).

Os registros de ataques cibernéticos com efeito em meio físico são numerosos e devem ser levados em consideração em uma análise de segurança crítica e de forma integrada. As ocorrências mencionadas já são exemplos que estimulam iniciativas de elaboração e aperfeiçoamento de métodos que contemplam aspectos de ambas as áreas de segurança.

1.4 ESTRUTURA DO TRABALHO

No capítulo 2 está descrito a fundamentação teórica necessária para entendimento deste estudo onde são apresentados os conceitos relacionados à segurança cibernética e à segurança crítica, as relações entre as duas áreas, os tipos de classificação das abordagens de análise integrada. Além de apresentar fontes de referências e métodos para obtenção de índices quantitativos para ataques cibernéticos.

No capítulo 3 são apresentadas as iniciativas de integração e os principais métodos de análise conjunta de segurança crítica e segurança cibernética. São expostas as caraterísticas e objetivos dos métodos quando aplicados para análise de um sistema ciber físico em domínio crítico.

No capítulo 4 está relatada a metodologia de pesquisa adotada neste estudo, com a definição das questões de pesquisa, critérios de seleção das abordagens, elaboração e aplicação em estudos de caso e validação dos resultados.

Nos capítulos 5 e 6 constam os estudos de caso nos quais foram aplicadas as metodologias de análise integrada selecionadas. O primeiro estudo de caso trata-se de uma versão simplificada de um Sistema *Anti-Heeling* automático (AHS) de navio. O segundo estudo de caso trata-se de um sistema de torre de controle de aeródromo (TWR). Ambos os estudos de caso são analisados sob a ótica de cada metodologia.

No capítulo 7 constam os resultados obtidos pela pesquisa que se referem a comparação das características das metodologias de análise integradas que foram aplicados nos estudos de caso. Os resultados contribuem para um melhor entendimento das metodologias selecionadas e servem como subsídio para a definição de uma estratégia de condução de análise.

No capítulo 8 são apresentadas as considerações finais com base nos resultados alcançados. Nas conclusões está sumarizada a importância da aplicação de abordagens de análises integrada de segurança crítica e segurança cibernética em sistemas ciber físicos para a detecção de vulnerabilidades e situações de perigos que poderiam não ser detectadas por abordagens convencionais. Nas perspectivas de trabalhos futuro é exposto a necessidade de se estimular pesquisas, métodos e ferramentas para obtenção de dados quantitativos de ameaças cibernéticas.

2 FUNDAMENTAÇÃO TEÓRICA

O objetivo deste capítulo é esclarecer os principais conceitos e propriedades que definem e diferenciam a segurança crítica e a segurança cibernética. Tais conceitos básicos em ambas as áreas servem de subsídios para uma compreensão melhor deste trabalho.

2.1 CONCEITOS DE SEGURANÇA CRÍTICA E DE SEGURANÇA CIBERNÉTICA

Em Almeida, Camargo e Cugnasca (2013) a Segurança Crítica, referenciada na língua inglesa como *Safety*, é entendida como a forma de proteção praticada nas Aplicações Críticas visando a salvaguarda da integridade física de pessoas, do meio ambiente ou de propriedades. O foco de segurança crítica está ligado a técnicas de prevenção de acidentes. A Segurança Cibernética, referenciada como *Security*, significa a proteção da informação e dos sistemas contra acessos não autorizados de forma a reduzir intervenções maliciosas. Outro conceito relacionado é de Aplicações Críticas que Almeida, Camargo e Cugnasca (2013) define como uma classe de sistema cuja falha pode afetar a integridade física de pessoas e da infraestrutura agregada. São exemplificados nessa classe: usinas nucleares, plantas químicas e petroquímicas, sistemas de transporte ferroviário, sistemas marítimos, sistemas ligados à aeronáutica, sistemas de controle de transmissão e distribuição de energia elétrica, além de dispositivos médicos.

Tradicionalmente as plataformas críticas eram compostas somente de equipamentos puramente mecânicos e eletroeletrônicos. Então, as abordagens de análise de segurança inicialmente desenvolvidas se concentravam em falhas acidentais, ou seja, *safety*. No entanto, por necessidade de acompanhar evoluções de negócios, como otimização de custos para implantar, manter e operar, e acompanhar as tendências de inovação em um contexto industrial competitivo. Novas tecnologias passaram a ser cada vez mais incorporadas nas aplicações críticas. Esta migração para tecnologias de comunicação com protocolos abertos, com sistemas digitais e com diversos dispositivos conectados aumentou a complexidade e a imprevisibilidade das aplicações críticas e tornando-a expostas as vulnerabilidades de segurança da cibernética, *security*. Apesar de ambos tipos de análises terem objetivos convergidos de proteger a aplicação crítica, seja contra falha acidental ou intencional, ambas são

tratadas de formas separadas, por diferentes metodologias, diferentes equipes de especialistas e que muitas vezes não se comunicam. Neste estudo, foram abordadas ações de *safety* e *security*, que possuem interações e efeitos entre si, ou melhor, uma ação para melhorar o nível de em uma das áreas pode prejudicar a outra. Portanto, torna-se necessário abordagens de análise conjunta considerando aspectos de *safety* e *security* para que seja possível a realização de uma cobertura efetiva de riscos e conflitos das interações entre as áreas.

Sun (2018) menciona que análise de segurança crítica, *safety*, é direcionada para o interior do sistema, devido a erros aleatórios, erros de sistema, efeito ambiental, erros de operação e falhas oriunda de softwares. Os acidentes são resultados dos efeitos dessas falhas aleatórias e não previstas. A identificação dos perigos, *hazards*, é uma das primeiras etapas da análise de *safety*, onde Sun (2018) afirma que quando algumas combinações ocorrem, a condição de perigo fica ativada, podendo evoluir para a situação de acidente. As combinações podem ser: uma operação inesperada, um uso indevido, condições climáticas adversas, falha aleatória de *hardware*, falha de *software*. A Figura 1 ilustra a evolução das combinações até o resultado indesejável, o acidente.

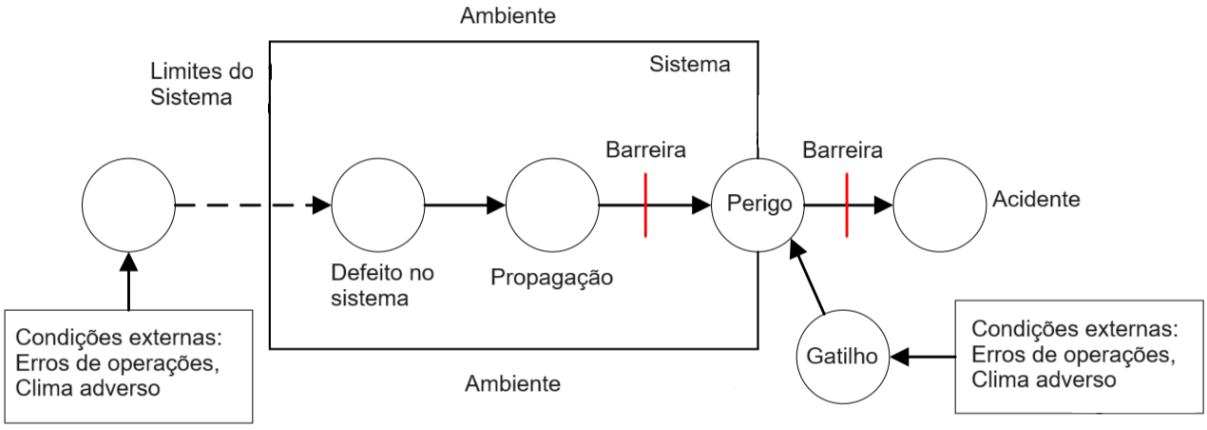

Figura 1 – Da falha e combinação de fatores ao acidente
Fonte: Adaptado de Sun (2018)

Da ocorrência de eventos adversos externos e falhas internas, caso as proteções do sistema sejam inexistentes ou insuficientes, a combinação ativa o gatilho do perigo, *hazard*. Se os perigos não forem identificados, provavelmente não haverá barreiras, e os perigos associados a eventos externos podem ser catastrófico.

Guzman, Kozine e Lundteigen (2021) mencionam a definição de segurança da cibernética, *security*, como a capacidade de manter a confidencialidade, integridade e autenticidade contra intervenções não autorizadas e maliciosas. Para *security*, O conceito de ameaça, *threat*, é equivalente ao de perigo, *hazard*, para *safety*. Na evolução de uma ameaça, os incidentes de segurança partem da exploração de vulnerabilidades por um agente malicioso até converter na condição de perigo que, associado a outras condições adversas podem resultar em um acidente. A Figura 2 ilustra os caminhos.

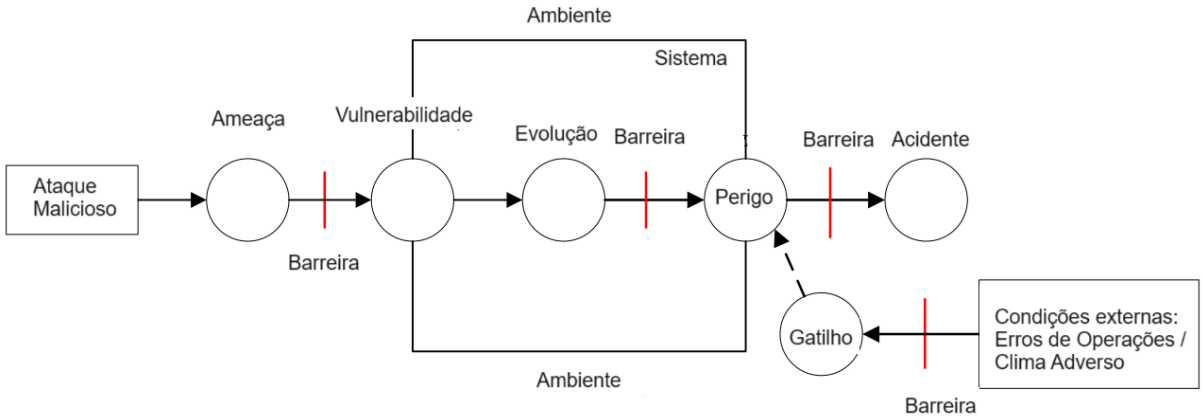

Figura 2 – Caminhos de um ataque até o acidente
Fonte: Adaptado de Sun (2018)

A porta de entrada é a existência da vulnerabilidade sem barreiras adequadas de proteção. Uma vez que um ataque acontece, o estado do sistema se altera, a ausência de meios de monitoramento e medidas de proteção agrava o efeito resultando na ativação do *Hazard*. Neste ponto, sob certas condições pode evoluir para acidentes.

Em uma visão integrada, podemos inserir em um mesmo diagrama ambos os caminhos oriundos dos pontos de partida de *safety* e *security* até resultar na situação catastrófica. Na Figura 3 é ilustrado essa trajetória integrada com as interações entre ambas as áreas. A diferença é apenas do início, mas o caminho de evolução após falha ou ataque é o mesmo.

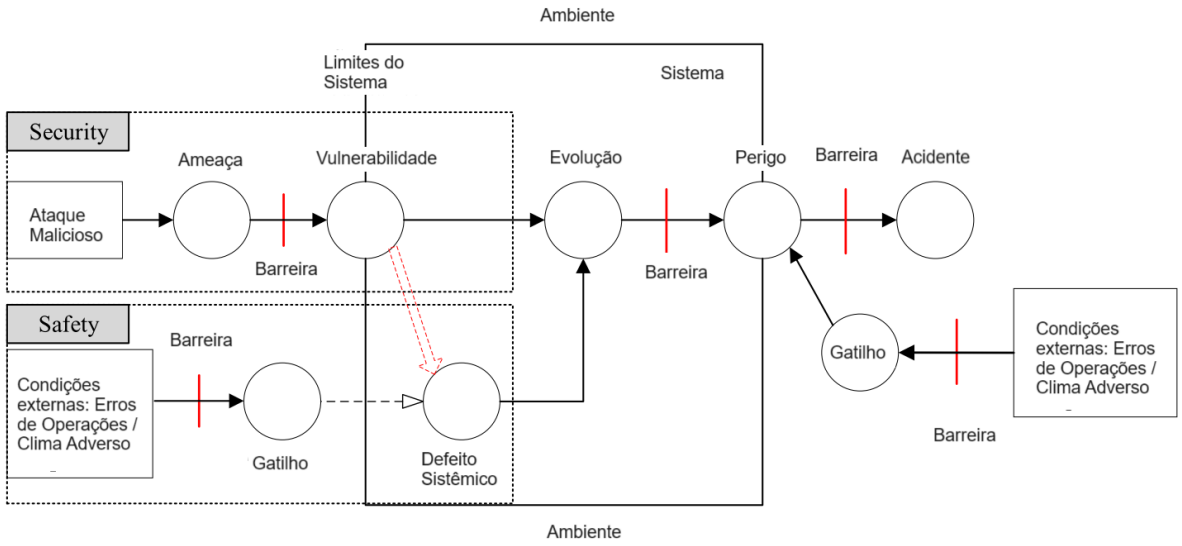

Figura 3 – Evoluções de ataque e de falha até o acidente
Fonte: Adaptado de Sun (2018)

Tanto por meio da ocorrência de uma falha ou de um ataque, ambos podem evoluir para ativação de um *Hazard* que por sua vez pode resultar no acidente. Para exemplificar melhor, podemos citar que um acidente ocasionado a partir da ineficiência de um sistema de proteção contra excesso de velocidade pode ter tido sua origem em uma falha aleatória interna ou por uma desativação maliciosa por invasor. Quando o sistema de proteção contra excesso de velocidade fica indisponível, independentemente se foi ataque ou falha, o resultado é o mesmo. Ainda através da Figura 3 é possível observar, pelas setas duplas vermelhas, que um ataque pode contribuir para ocorrência de falhas no sistema.

Em uma definição mais clássica, Avižienis et al. (2004) apresenta o conceito de *Dependability*, que seria um termo mais geral associado a sistemas críticos, com o significado de capacidade de oferecer um serviço que pode ser justificadamente confiável envolvendo atributos de Confiabilidade, Disponibilidade, Integridade, Manutenibilidade e Segurança (*safety*):

- Confiabilidade: continuidade do serviço;
- Disponibilidade: prontidão para entregar o serviço;
- Integridade: ausência de alteração indevida do sistema;
- Manutenibilidade: capacidade de sofrer modificações e reparos;

- Segurança (*safety*): ausência de consequências catastróficas para os usuários do sistema e para o meio ambiente.

Para Avižienis et al. (2004), o conceito de dependabilidade incorpora os requisitos de disponibilidade e integridade, mas exclui confidencialidade. Não associando portando o requisito de confidencialidade ao risco de consequência catastrófica. A Figura 4 ilustra essa relação de atributos.

Figura 4 – Atributos de Dependabilidade e de Segurança Cibernética
Fonte: Adaptado de Avižienis et al. (2004)

Em uma classificação mais ampla para *security*, os pesquisadores Praerit Garg e Loren Kohnfelder da Microsoft classificaram as ameaças e propuseram o modelo STRIDE. Abuemera et al. (2022) apresentam uma abordagem baseada em tal modelo. STRIDE leva esse nome devido ao mnemônico para a classificação das propriedades de *security* que seriam ***S**poofing*, ***T**ampering*, ***R**epudiation*, ***I**nformation disclosure, **D**enial of Service* e ***E**levation of privilegie*. A Tabela 1 mostra as ameaças, propriedades atingidas e a descrição das ações.

Tabela 1 – Classificação de Ameaças pelo Modelo STRIDE

Ameaça	Propriedade atingida	Descrição
***S**poofing*	Autenticação	Assumir a identidade de outro dispositivo
***T**ampering*	Integridade	Modificar dados de forma fraudulenta
***R**epudiation*	Não repúdio	Falsificar autoria de uma ação
***I**nformation disclosure*	Confidencialidade	Vazamento informações sigilosas
***D**enial of Service*	Disponibilidade	Tornar o serviço indisponível. Também denominado por *Jamming*
***E**levation of privilegie*	Autorização	Assumir um controle indevidamente

Fonte: Adaptado de Abuemera, Elzouka e Saad (2022)

O modelo STRIDE foi desenvolvido com foco somente em *security*, cabendo ao analista associar as ameaças com efeitos em *safety* no sistema.

2.2 RELAÇÕES ENTRE SEGURANÇA CRÍTICA E SEGURANÇA CIBERNÉTICA

Embora possuam algumas convergências, as diferenças entre *safety* e *security* também se refletem na distinção de uso de ferramentas de análise, normas e na forma como a gestão de riscos é conduzida em ambos domínios. Avaliar uma ameaça de *security* é radicalmente diferente de se avaliar um perigo de *safety*. No primeiro caso, as fontes das ameaças a serem avaliadas geralmente não são bem conhecidas pelo analista, e abrangem uma gama extremamente ampla de cenários possíveis.

No segundo caso, as características dos perigos são mais acessíveis, e o número de cenários a serem considerados também pode ser reduzido a um conjunto restrito. Nas abordagens quantitativas as métricas para *safety* são mais estáveis ao longo do tempo e o histórico sobre acidentes anteriores tende a ser plenamente confiável e suficiente para aplicação de uma medida de proteção. Enquanto os atributos de *security* são menos previsíveis e dependentes de muitos fatores como o perfil do invasor, habilidades e motivação. Isso torna mais difícil para um analista quantificar possíveis cenários. Portanto, o uso de probabilidades em abordagens quantitativas é amplamente adotado em *safety*, mas nem sempre são aceitas em *security* (YAQOOB; ABBAS; SHAFQAT, 2020).

Requisitos e contramedidas de uma área, *safety* ou *security*, podem impactar na outra. Tradicionalmente, ambas áreas foram analisadas de forma distinta, mas com a crescente incorporação das tecnologias de informação nas aplicações críticas elevou-se a preocupação para tratar ambas as disciplinas em conjunto. Pietre-Cambacedes e Bouissou (2010) relacionou as interdependências de *safety* e *security* em quatro tipos de relações:

- Dependência condicional: o atendimento a um requisito de *security* é uma condição para um requisito de *safety*, ou vice-versa.
- Reforço mútuo: o atendimento a um requisito de *security* contribui para um requisito de *safety*, ou vice-versa.

- Antagonismo: situação de conflito onde o atendimento a um requisito de *security* prejudica um requisito de *safety*, ou vice-versa.
- Independência: quando o atendimento a requisitos de qualquer área não influencia na outra.

2.3 CLASSIFICAÇÃO DAS ABORDAGENS INTEGRADAS

Kriaa, Bouissou e Laarouchi (2019) dividem as abordagens de análise conjunta de segurança crítica e proteção cibernética em dois grupos: orientadas a processos e baseadas em modelos. Di Maio et al. (2020) acrescenta uma terceira classificação, orientada a objetivos.

As abordagens orientadas a processos estabelecem novos ciclos de vida e sequencias metodológicas, e são empregadas nas fases iniciais da concepção do sistema. Normalmente, se baseiam em requisitos oriundos de normas e fornecem descrições genéricas de ciclos de vida, indicando que tipos de atividades devem ser executadas e em que ordem. Esse tipo de abordagem tem características muito macroscópica, proporcionando uma visão global que auxiliam os analistas a identificarem alguns problemas antes mesmo da implementação do sistema. Porém, devido ao foco nas fases iniciais, não costumam ser tão eficientes para representar e capturar características oriundas do baixo nível do sistema. Também não são adequadas para análises quantitativas pela ausência de representação matemática.

As abordagens baseadas em modelos dependem de uma representação formal dos aspectos funcionais e não funcionais do sistema e geralmente são suportadas por ferramentas. Ainda podem ser subdivididas se dependem de modelos gráficos ou não gráficos (KRIAA, 2016). Costumam ser muito mais práticas e objetivas para análises quantitativas, porém demandam esforços para adaptar mudanças nos sistemas e características de novos dispositivos, além de exigir alta experiência e domínio dos analistas nas ferramentas empregadas.

A terceira classificação, baseada em objetivos, propõe decompor o sistema em funções e subfunções de forma hierárquica, partindo-se do alto nível para o baixo nível, e representar as interações entre os componentes do sistema e seu vínculo com as funções. A hierarquia desse tipo de abordagem é construída respondendo à pergunta

“como” uma função de nível superior é alcançada, até que as funções de nível mais baixo sejam alcançadas, ou seja, quando as funções foram suficientemente decompostas (DI MAIO et al., 2020).

Para o contexto de integração *safety* e *security*, Kriaa, Bouissou e Laarouchi (2019) ainda mencionam abordagens unificadas, que visam unir técnicas de *safety* e *security* em uma única metodologia e ferramenta, e as abordagens harmonizadas, que tratam os aspectos de *safety* e *security* em metodologias e ferramentas diferentes, mas alinham de forma manual. Quanto à fase de aplicação de abordagens, podem ser agrupadas entre fase de projeto, que foca na identificação de requisitos e caso de uso e que podem levar a situações catastróficas, e a fase de operação que visa avaliar o sistema enquanto está em uso. Em seu estudo, Kriaa, Bouissou e Laarouchi (2019) sugerem quatro critérios para que uma abordagem possa ser considerada ideal:

- Critério 1: permite modelagem formal com base em conceitos matemáticos e métodos de raciocínio;
- Critério 2: produz análises qualitativas e quantitativas, uma abordagem com recursos qualitativos e quantitativos tem maior cobertura para capturar as interdependências de segurança crítica e de segurança cibernética;
- Critério 3: gera modelos de risco automaticamente, isso torna a abordagem fácil de usar e garante a objetividade dos modelos;
- Critério 4: robustez, ou seja, alterações no sistema não impactam na necessidade de reformular todo o modelo de análise.

O critério 3 é o menos atendido pela ampla maioria das abordagens. Há uma predominância de metodologias onde a construção de modelos de análise é manual e a identificação dos resultados partem mais da percepção dos analistas do que originada pela própria metodologia. Outra implicação dessa característica é que alterações no sistema levam novamente a construções manuais do modelo de análise, que pode tornar a análise muito onerosa e exaustiva. Por estas razões, Kriaa, Bouissou e Laarouchi (2019) reforçam a necessidade de desenvolvimento de modelos automatizados a fim de facilitar a construção de simulações, a reprodutibilidade dos resultados e redução de erros nas análises.

2.4 ANÁLISE QUALITATIVA E QUANTITATIVA

Uma análise de segurança crítica também pode ser classificada como qualitativa ou quantitativa de acordo com os tipos de resultados fornecidos (LYU; DING; YANG, 2019). As abordagens qualitativas fornecem as identificações de causas e consequências, onde os resultados podem ser apresentados como uma combinação de eventos que levam a uma situação indesejada, como por exemplo, uma lista de *Minimal Cut Sets* (MCS) que são obtidas através do método de Árvore de Falhas.

As análises quantitativas associam indicadores numéricos como a probabilidade de ocorrer um evento em um determinado tempo. Em análises por meio de cadeias de Markov (MA) utiliza-se o parâmetro taxa de falha para os eventos, que combinados matematicamente, podem indicar a probabilidade de um sistema atingir um estado inconsistente em um instante de tempo. Resultados quantitativos são exigidos para atendimento a parâmetros como o *Safety Integrity Level* (SIL) que é definido na norma IEC 61508, e assim, a aplicação ter aceite para ser comissionada (HOLLERER; KASTNER; SAUTER, 2021).

Para indicadores quantitativos em eventos de segurança crítica com falhas não intencionais, os valores de taxa de falha de um componente normalmente são os fornecidos pelos fabricantes ou são estimados. Já para indicadores em eventos maliciosos intencionais há uma grande lacuna sobre como estimá-los, e alguns métodos utilizados são questionáveis por muitos especialistas devido ao uso de bases de dados limitadas (ZHOU et al., 2020).

2.5 OBTENÇÃO DE PARÂMETROS QUANTITATIVOS

Existem abordagens que englobam tanto análises qualitativas quanto quantitativas. No entanto, ainda são enfrentados diversos desafios para obtenção de resultados quantitativos a partir de métodos de análises que integram aspectos de segurança crítica e de segurança cibernética. Fatores que incrementam essa dificuldade podem ser atribuídos a complexidade para se estimar probabilidades de ataques cibernéticos e a escassez de base de dados comum com registros de ocorrência de ataques.

Uma referência para obtenção de índices quantitativos é o *Common Vulnerability Score System* (CVSS) que está em sua quarta versão e é desenvolvida e mantida pela

Organização denominada FIRST (2023). O CVSS trata-se de um método para identificar as características fundamentais de uma vulnerabilidade, permitindo calcular uma pontuação numérica que represente o potencial de dano. Com tais valores obtidos, os analistas podem melhor gerenciar as vulnerabilidades aplicando correções para tratar elementos com riscos mais elevados. A pontuação resultante é dividida em três métricas: Base, Temporal e Ambiental.

A métrica Base refere-se às características intrínsecas da vulnerabilidade, considerando fatores como a complexidade de exploração, o impacto no sistema e a acessibilidade remota. A métrica Temporal leva em conta elementos que podem variar ao longo do tempo, como a disponibilidade de *patches* de segurança ou a evolução das técnicas de exploração. Já a métrica ambiental considera o contexto específico do ambiente no qual a vulnerabilidade está inserida, como a criticidade do sistema afetado. A pontuação final é apresentada em uma escala de 0 a 10, sendo 10 o nível máximo de gravidade. Além disso, o CVSS atribui também uma classificação de gravidade, dividindo as pontuações em categorias como Baixa, Média, Alta e Crítica (FIRST, 2023).

O *National Institute of Standard and Technology* (NIST) do Departamento de Comércio dos Estados Unidos disponibiliza uma calculadora CVSS que permite aos analistas de segurança obterem a pontuação referente ao ambiente computacional (NIST, 2023). As entradas de dados correspondem às características apontadas para um determinado ativo, por exemplo, como na métrica de vetor de ataque do grupo base. Neste caso, deve ser assinalado se o ativo é acessível através da internet, de uma rede adjacente, de uma rede local ou somente de contato direto. Para esta métrica quanto maior a exposição maior a pontuação. A pontuação final será combinada com as demais métricas. A Figura 5 ilustra a calculadora com as métricas que compõe o grupo Base.

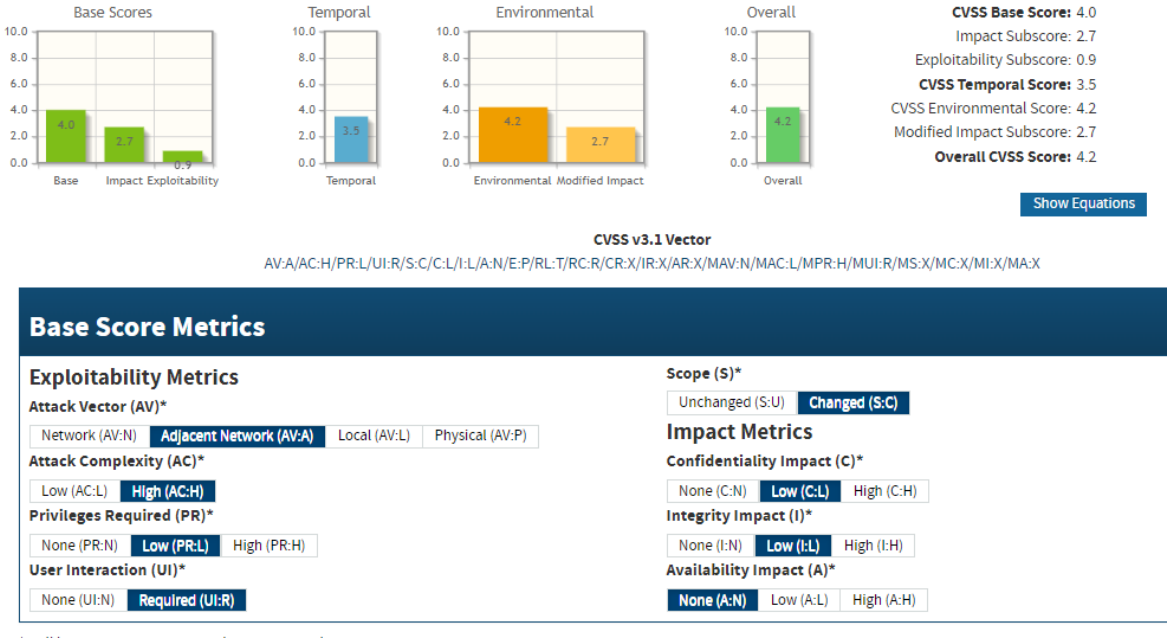

Figura 5 – Calculadora de Pontuação CVSS
Fonte: (NIST, 2023)

A implementação do CVSS é bastante útil para a priorização de correções de segurança cibernética permitindo que as organizações avaliem, classifiquem e priorizem o tratamento de vulnerabilidades com maiores riscos. Porém, o CVSS por si só não fornece um índice como o de probabilidade de ataque cibernético que poderia estar na mesma escala e ser comparado diretamente com um índice de probabilidade de falha acidental de um componente. A comparação direta seria fundamental para uma análise quantitativa de aspectos de segurança crítica e de segurança cibernética de forma integrada.

Hollerer, Sauter e Kastner (2022) apresentam um modelo de análise quantitativa para ambientes operacionais englobando aspectos de segurança crítica e segurança cibernética. O modelo converte a pontuação CVSS para as faixas do indicador *Security Level* (SL). Os indicadores SL são definidos na norma IEC 62443: *Security for industrial automation and control systems* (INTERNATIONAL ELECTROTECHNICAL COMMISSION, 2020) onde o nível mais de segurança mais alto, SL 4, refere-se à existência de proteções mais sofisticadas e o nível mais baixo, SL 1, às proteções mais simples. As proteções exigidas em um determinado nível de SL são inseridas na calculadora CVSS e são extraídas uma faixa de intervalo. Por exemplo, para o SL 4 as pontuações das proteções obtidas no CVSS variaram de 1.6 a 1.7 e para o SL 1 variaram de 4.5 a 5.8. A referência no CVSS é quanto mais alta a pontuação mais

vulnerável é o sistema. Essa conversão permite uma comparação do CVSS com o indicador *Safety Integrity Level* (SIL) aplicando as relações entre SL e SIL definidas nas normas IEC 62443 (INTERNATIONAL ELECTROTECHNICAL COMMISSION, 2020) e também apresentadas por Braband (2016).

Yaqoob, Abbas e Shafqat (2020) apresentam um modelo integrado de avaliação de riscos de segurança crítica, cibernética e de privacidade para dispositivos médicos. O modelo aplica a metodologia CVSS para estimar pontuação de elementos de segurança crítica, caminho oposto ao anterior. No entanto, a proposta estabelece a criação de três variáveis, que são denominadas eficiência de controle, e associadas a cada um dos três tipos de riscos: crítico, cibernético e de privacidade. Para cada ativo digital crítico, são levantadas essas três pontuações. Os resultados permitem aos analistas identificarem que uma correção de uma vulnerabilidade pode alterar o valor da variável de eficiência de controle de segurança crítica do ativo.

Apesar de ambos estudos apresentarem propostas de conversão de índices, que são bem úteis para direcionar uma análise integrada, ainda existem muitas lacunas para se obter índices quantitativos fidedignos de probabilidade de ataques cibernéticos devido à natureza não estocástica. Uma das formas de contornar esse problema poderia ser a implantação de uma base de dados onde fossem registrados todos ataques cibernéticos e catalogados. Em uma analogia com dados de segurança pública, as polícias calculam a taxa de criminalidade de uma cidade ou região pela quantidade de registros de ocorrências divididos pela população em um intervalo de tempo. Porém, se houver crimes em que a vítima não registra a ocorrência os dados começam a perder a fidelidade. A existência de uma base de múltiplo acesso, inclusive em âmbito mundial, surge como uma ideia interessante para pesquisas futuras. Como projeto aderente a ideia existe a ferramenta *Malware Information Sharing Platform* (MISP) desenvolvida com a proposta de permitir compartilhamento de dados sobre vulnerabilidades, ataques cibernéticos entre os membros (MISP Project, 2024).

Di Maio, Mascherona e Zio (2020) utilizaram dados de probabilidade de ataques a um sistema SCADA a partir dos registros de um Banco de Dados de Incidentes de Segurança Industrial ou *Industrial Security Incident Database* (ISID). No banco constava o registro de 23 invasões em período de 1 ano para o total de 500 sistemas SCADA monitorados. Logo, foi considerado que a probabilidade de um sistema SCADA ser

invadido seria de 0,046 ataques por ano. Apesar do ISID conter registros históricos de incidentes de cibersegurança industrial, o valor estimado ainda pode ser bem divergente de um valor real e uma das razões é não considerar a influência de medidas de proteção cibernética e vulnerabilidades. Porém, já é um caminho para permitir uma compreensão realistas das ameaças cibernéticas. Byres, Lowe e Leversage (2015) desenvolveram o ISID como projeto de pesquisa acadêmico e posteriormente foi evoluído e rebatizado para *Repository of Industrial Security Incidents* (RISI).

3 METODOLOGIA DE PESQUISA

Este estudo adaptou alguns preceitos propostos por Peffers et al. (2014) para o desenvolvimento de pesquisas nas áreas de engenharias e tecnologia da informação, dividindo nas seguintes atividades: definição das questões da pesquisa, filtros de critérios para seleção, aplicação em estudos de casos e discussão dos resultados.

3.1 DEFINIÇÃO DAS QUESTÕES DE PESQUISA

A definição das questões de pesquisas resultou a partir de uma sequência de processos descritos como: fato observado, os problemas relacionados ao fato, as soluções correntes para os problemas, as lacunas das soluções correntes e, então, a proposta de solução para preenchimento dessas lacunas.

Processo 1 (Fato Observado): Plataformas críticas estão ampliando o uso ativos de tecnologia da informação integrados nas operações físicas. As ações mecânicas estão dependentes de processamento digital.

Processo 2 (Problema do Fato Observado): um ataque cibernético pode resultar em consequências físicas indesejáveis, ou seja, acidentes.

Processo 3 (Soluções Atuais): diversas propostas de abordagens para análises de segurança cibernética em plataformas críticas foram lançadas. Porém, usualmente as análises de segurança cibernética ainda são aplicadas separadamente das análises de segurança crítica.

Processo 4 (Lacunas das Soluções Atuais): análises em processos distintos podem deixar de capturar relações entre segurança crítica e segurança cibernética. Como por exemplo, os casos de antagonismo onde uma medida aplicada para beneficiar uma área pode prejudicar a outra.

Processo 5 (Propostas): avaliar as principais abordagens de análise integrada de segurança crítica e segurança cibernética apresentando as comparações entre os métodos por meio de aplicação em estudos de caso. Com os resultados, apresentar uma estratégia de condução de análise integrada.

Processo 6 (Questões de Pesquisa): outros estudos com abordagens de análise conjunta têm sido propostos, mas ainda possuem poucas aderências de uso. Então, a primeira questão de pesquisa refere-se ao levantamento dessas metodologias de análise integrada de segurança crítica e segurança cibernética considerando os últimos 5 anos. A segunda questão de pesquisa envolve uma busca pelas características dos novos sistemas críticos, quais tipos de dispositivos digitais e quais tipos de ameaças cibernéticas vinculadas. A terceira questão de pesquisa refere-se ao ambiente de aplicação e requisitos de aceitação de um sistema para ser incorporado a plataforma.

QP1: Quais métodos de análise integrada de *safety* e *security* são mais utilizados atualmente (de 2019 a 2024)?

QP2: Quais tecnologias de ponta estão sendo incorporadas em sistemas críticos e quais ameaças cibernéticas envolvidas?

QP3: Quais critérios de avaliação de sistemas críticos são utilizados para permitir seu emprego em ambientes cuja falha pode resultar em óbitos?

3.2 CRITÉRIOS DE SELEÇÃO DAS ABORDAGENS

Com as questões de pesquisa definidas, é elaborado um protocolo simplificado (*string* de busca) para revisão da literatura. As *strings* são aplicadas nas principais bases científicas IEEE Xplore, ACM Digital Library, Springer e Web of Science. Além de conferências relevantes para área como SAFECOMP e ESREL.

String de busca da QP1: (SAFETY II SECURITY II CYBERSECURITY) && (JOINT II INTEGRATED II COANALISYS II COMBINED II APPROACH)

String de busca da QP2: (CRITICAL SYSTEM) && (FPGA II FOG II CLOUD II WIRELESS II IOT) && (CYBER THREATS)

String de busca da QP3: (SECURITY REQUIREMENT II SAFETY REQUIREMENT) && (MARITIME II RAILWAY II AVIATION II NUCLEAR II CRITICAL FACILITIES).

As buscas foram restringidas para intervalo de 2019 à 2024, porém, alguns dos trabalhos foram de aplicações de métodos criados anteriormente. Como por exemplo, o

BDMP criado em 2008, adaptado em 2010, mas com aplicações recentes como de Czekster e Morisset (2021).

Nas primeiras execuções de busca, foi arbitrado como filtro a quantidade de ao menos dez trabalhos publicados. Em seguida, foram selecionadas somente as metodologias com propostas de integração *safety* e *security*. Avançando, foram realizadas comparações das características das metodologias como em que se baseiam como árvores, modelagem em UML e engenharia de requisitos. Entre as metodologias com muita similaridade foi escolhida aquela considerada mais relevante como, por exemplo, as metodologias AADL, SysML-Sec baseadas em modelagem UML foram excluídas por terem suas características já absorvidas pelo CHASSIS. Outro tipo de exclusão foi baseado considerando estudos que comparam metodologias como é o caso do AT-BT que Di Maio, Macherona e Zio (2020) apontam ser menos eficiente que a GTST-MLD. Similarmente a Leveson e Thomas (2018) afirmam característica do STPA e STPA-Sec que são capazes de identificar problemas que passam despercebidos em análises FMEA e FMVEA. A Tabela 2 mostra a relação das abordagens verificadas e o resultado da seleção.

Tabela 2 – Resultados da seleção das abordagens

N.	Abordagem verificada	Resultado seleção
1	*Architecture Analysis and Design Language* (AADL)	Substituída por CHASSIS
2	*Attack Tree Bow Tie* (AT-BT)	Substituída por GTST-MLD e BDMP
3	*Boolean logic Driven Markov Processes* (BDMP)	**Selecionada**.
4	*Combined Harm Assessment of Safety and Security for Information System* (CHASSIS)	**Selecionada**
5	Extended Fault Tree	Substituída por BDMP
6	*Failure Mode, Vulnerabilities and Effects Analysis* (FMVEA)	Substituída por STPA / STPA-Sec
7	*Goal Tree Success Tree Master Logic Diagram* (GTST-MLD)	**Selecionada**
8	S-Cube	**Selecionada**
9	*System-Theoretic Processes Analysis for Security* (STPA-Sec)	**Selecionada**
10	*System Modeling Language for Security* (SysML-Sec)	Substituída por CHASSIS
11	*Threat Analysis and Risk Assessment* (TARA)	Substituída por STPA / STPA-Sec

3.3 APLICAÇÃO EM ESTUDOS DE CASOS

Após a seleção das metodologias, a etapa seguinte consiste em submeter estudos de casos para serem analisados sob a ótica de cada metodologia. Em seguida comparar os resultados obtidos com as análises. Foram adaptados dois estudos de casos, sendo o primeiro, uma representação simplificada de um sistema *Anti-Heeling* (AH) de navio, cujo propósito é manter o navio balanceado pelos ajustes de níveis de água em tanques de lastro. O segundo estudo de caso trata-se de uma representação de sistema de torre de controle de aeródromo (TWR) que tem por objetivo promover o pouso e a decolagem de aeronaves de forma livre de acidentes. Ambos os sistemas são analisados pelos métodos BDMP, S-Cube, CHASSIS, GTST-MLD e STPA-Sec.

Para o método BDMP, são construídas as árvores com os elementos que representam *safety* e *security*, e ainda o que representa as transições, *triggers,* que diferencia de uma árvore comum. Os modelos BDMP são simulados pela ferramenta *Model Builder* da organização Risk Spectrum (RISK SPECTRUM AB, 2023). Foi obtida uma versão para fins acadêmicos. Além, da árvore BDMP, pode-se entrar com valores quantitativos para cada folha. O resultado fornecido pela simulação são os *Cut Sets,* ou Conjuntos de Cortes, e a probabilidade do evento topo ocorrer em um determinado tempo.

Para o método S-Cube, são construídas as arquiteturas do sistema e a simulação fornece os resultados de cenários de ataques e falhas. O S-Cube é simulado também no *Model Builder* (RISK SPECTRUM AB, 2023) e pode ter suas classes modificadas alterando a programação do sistema desenvolvida na linguagem Figaro (KHAN et al., 2021).

No método CHASSIS, qualquer ferramenta que represente diagramas UML pode ser utilizada. A condução da análise é realizada de forma conceitual seguindo os processos da abordagem onde são construídos casos de uso, diagramas de sequência e os demais processos adaptados de UML. A saída será a relação de requisitos de segurança crítica e segurança cibernética.

No método GTST-MLD são construídas as árvores cruzadas de funções e elementos necessários para cada função. Qualquer ferramenta visual que representa diagramas MLD pode ser utilizada.

Por fim para o método STPA-Sec, foram utilizadas as ferramentas Visual Pro, STAMP Workbench e SafetyHAT disponíveis pelo grupo desenvolvedor (MIT PSASS GROUP, 2023). As entradas para o STPA-Sec correspondem a uma sequência de definições hierárquicas entre perdas, perigos e restrições, estruturas de controle e ao fim são refinados os requisitos e contramedidas para os problemas identificados.

3.4 VALIDAÇÃO DOS RESULTADOS OBTIDOS

As metodologias são comparadas com base em suas **características**, com base nos **resultados** obtidos com as aplicações nos estudos de caso, e também com bases em **critérios** levantados como desejáveis para uma metodologia de análise ser considerada ideal.

Características: onde considera-se os atributos apontados pelo autor de cada metodologia. Como, o tipo de entrada de dados (árvore, diagramas UML ou arquitetura), o tipo de resultado (qualitativo ou quantitativo) e as ferramentas necessárias para condução da análise. Portanto, trata-se de uma comparação das descrições da metodologia extraídas da revisão bibliográfica.

Resultados: corresponde a comparação das saídas fornecidas pelas metodologias aplicadas nos estudos de caso. O objetivo é verificar como que as metodologias são capazes de identificar problemas de segurança cibernética que afetam a segurança crítica e os conflitos relacionados a ambas as áreas. Compara-se também como as metodologias apresentam esses resultados e qual foi o ponto de partida para as identificações.

Critérios desejáveis: corresponde as capacidades em atender requisitos para uma metodologia ser considerada ideal. Os critérios são mostrados na Tabela 3. Foram compiladas a partir de referências do modelo de engenharia de requisitos *Model Based*

System Engineering (MBSE) (BHOLE; KASTNER; SAUTER, 2022), (JAPS, 2020) e dos estudos de Kriaa (2016), El-Kady et al. (2023), Khan, Katoen e Bouissou (2023).

Tabela 3 – Critérios desejáveis em uma abordagem integrada

	Critérios Desejáveis	**Descrição**
1	Integração segurança crítica e segurança cibernética	Capacidade de representar elementos de segurança crítica, segurança cibernética e as relações entre ambas as áreas.
2	Qualitativa e Quantitativa	Produção de resultados qualitativos e quantitativos
3	Resultados automáticos	Capacidade de identificar problemas de segurança crítica e segurança cibernética mesmo sem ter sido previamente definido nas entradas.
4	Gerenciável e Legível	Capacidade de permitir uma organização facilitada, de uma construção intuitiva e produção de documentos de fácil compreensão.
5	Ampla aplicação	Capacidade de ser utilizada nas fases de concepção, desenvolvimento e operação em qualquer tipo de sistema ciber físico.
6	Flexível, Escalável e Modular	Capacidade da metodologia atender alterações, crescimento e divisões do sistema sem invalidar análises anteriores.
7	Rastreável	Capacidade de conectar efeito de baixo nível com requisito de alto nível do sistema.
8	Ferramentas dedicadas	Disponibilidade de ferramentas para realização de análises.

Fonte: obtidos de Bhole, Kastner e Sauter (2022), Japs (2020), Kriaa (2016), El-Kady et al. (2023) e Khan, Katoen e Bouissou (2023)

Com os resultados das comparações das metodologias, o estudo ainda apresenta uma proposta de estratégia de condução de análise integrada de segurança crítica e segurança cibernética. O intuito da proposta é subsidiar melhor na escolha do modelo de análise e na definição correta dos valores de entrada.

4 REVISÃO DA LITERATURA

Desde que as plataformas críticas passaram a ter suas operações dependentes de sistemas de informação, as preocupações com a proteção cibernética aumentaram e as análises de *security* passaram a fazer parte do contexto de *safety*, inicialmente de forma separada. No entanto, com os avanços das análises foi observado a necessidade de integração de métodos entre ambas as áreas. Algumas iniciativas de pesquisas, novas metodologias e ferramentas surgiram, porém, devido à alta velocidade de avanços em ativos digitais, as metodologias logo precisam ser aprimoradas para cobrir as características evolutivas desses ativos, considerando as novas vulnerabilidades que podem trazer. Yang et al. (2022) destaca o incremento de dispositivos com recursos de Inteligência Artificial (IA), sensores e dispositivos de comunicação sem fio cada vez mais presentes em equipamentos tradicionais em aplicações críticas. Também estima que o número de dispositivos Internet das Coisas (IoT) aumentará para 3,5 bilhões em 2024. Por esses dados apresentados, Yang et al. (2022) destaca a importância de pesquisas que avaliem os efeitos que tais dispositivos, suscetíveis a ataques maliciosos, podem trazer ao campo de *safety*.

4.1 INICIATIVAS DE INTEGRAÇÃO

Alguns grupos de pesquisa foram estabelecidos com o propósito de desenvolver metodologias, ferramentas, normas e guias nos aspectos de proteção cibernética em conjunto com análises de segurança crítica.

ISA-99 WG7: é um grupo de trabalho estabelecido dentro do comitê ISA-99 para tratar de questões relacionadas a *safety* e *security* na automação industrial e sistemas de controle. O objetivo deste grupo de trabalho é estender o ciclo de vida de segurança crítica existente para considerar aspectos de segurança cibernética em diferentes fases do ciclo de vida do processo industrial (ou seja, projeto, implementação, comissionamento e manutenção), a fim de garantir uma plataforma confiável, eficiente e segura (INTERNATIONAL SOCIETY OF AUTOMATION, 2023).

IEC TC65: é um grupo de trabalho formado pela *International Electrotechnical Commission* (IEC), com os propósitos de: elaborar normas internacionais para sistemas e elementos utilizados na medição, controle e automação de processos industriais

considerando aspectos de coordenar as atividades de padronização que afetam a integração de componentes e funções em tais sistemas, incluindo aspectos de *safety* e *security*. O TC 65 atua na Cibersegurança para Tecnologias Operacionais que inclui envolvimento em todo o ciclo de vida desde o projeto até o descarte (incluindo cadeia de suprimentos, etc.), nos requisitos técnico, organizacional e processual e nos componentes, subsistemas e sistemas. (INTERNATIONAL ELECTROTECHNICAL COMMISSION, 2023).

DO-326/ED-202: conjuntos de normas, desenvolvido por grupos de trabalho da *Radio Technical Committee for Aeronautics* (RTCA) dos Estados Unidos e da *European Organisation for Civil Aviation Equipment* (EUROCAE) que tem como objetivo aumentar a orientação atual para certificação de aeronaves no tratamento de ameaças cibernéticas, *security*, para a segurança crítica da aeronave, *safety*. O foco é impedir que os sistemas de aviação sejam violados por *hackers* ou infectados por *malware*, por exemplo, qualquer falha resultante pode ameaçar drasticamente a segurança de passageiros e operadores. As normas foram projetadas para abordar todo o ciclo de vida de desenvolvimento da segurança cibernética do sistema de aviação, desde o conceito até a implantação e comissionamento. (EUROCAE, 2023; RTCA, 2023). Outra iniciativa é a *Brazilian Society of Air Transportaton Research* (SBTA) que visa estimular a pesquisa aplicada ao setor de transporte aéreo brasileiro (SBTA, 2024).

CEN/CLC/JTC 13: corresponde a iniciativa do *European Committee for Electrotechnical Standardization* (CENELEC) e tem como propósito o desenvolvimento de normas de segurança cibernética e proteção de dados utilizadas como suporte para desenvolvimento, programação e monitoramento de sistemas ferroviários. O grupo é responsável por descrever as ações que devem ser tomadas para demonstrar a segurança do sistema considerando o contexto cibernético (CEN, 2021).

Nuclear Security Series (NSS): corresponde a um conjunto de guias da *International Atomic Energy Agency* (IAEA) que visa incorporar aspectos de cibersegurança para apoiar especialistas em todo o mundo na implementação de medidas de proteção cibernética para fortalecer suas implementações nacionais de segurança nuclear. A IAEA fornece orientação em todas as fases de avaliação de segurança cibernética de sistema voltado para atividades nucleares. Ainda provê treinamentos para que os Estados Membros possam desenvolver seus próprios

programas de segurança cibernética na aplicação nuclear. A Agência realiza missões de consultoria, treina inspetores e fornece experiência em planejamento na condução de exercícios de segurança de computadores como parte do programa de segurança nuclear (IAEA, 2023).

Maritime Safety Committee (MSC): é um grupo técnico da *International Maritime Organization* (IMO) responsável por temas relacionados à segurança crítica no ambiente marítimo. Possui um subcomitê que trata especificamente dos riscos cibernéticos que podem resultar em falhas operacionais e de segurança crítica relacionadas ao transporte marítimo. Como iniciativas, foram lançados os guias MSC-FAL.1-Circ.3-Rev.2, *Guidelines on maritime cyber risk management*, o anexo 10 da resolução MSC.428(98), *Maritime Cyber Risk Management in Safety Management Systems*, e o guia *The Guidelines On Cyber Security Onboard Ships* (IMO, 2023).

***NATO Cooperative Cyber Defence Centre of Excellence* (CCDCOE)**: grupo formado por países membros da Organização do Tratado do Atlântico Norte (NATO) que visa o desenvolvimento e cooperação no campo de defesa cibernética aplicada às infraestruturas críticas militares e civis. O grupo promove pesquisas, treinamentos e exercícios que contemplam efeitos de ameaças aos sistemas ciber físicos associados (CCDCOE, 2021). No Brasil, com estruturas e objetivos equivalentes, foi estabelecido o Comando de Defesa Cibernética (COMDCIBER, 2017).

4.2 FORMALISMO BDMP

O formalismo *Boolean logic Driven Markov Processes* (BDMP) foi proposto por Bouissou (2008) inicialmente apenas para análise de segurança crítica, *safety*, como uma alternativa ao método tradicional de Árvore de Falhas, mas posteriormente adaptado por Pietre-Cambacedes e Bouissou (2010) para incluir aspectos de segurança cibernética. Recentemente, Czekster e Morisset (2021) apresentaram a ferramenta *BDMPathfinder* que visa explorar possíveis ataques cibernéticos considerando a progressão ao longo do tempo.

O BDMP basicamente se diferencia de Árvore de Falhas por incorporar a funcionalidade denominada *trigger*, que proporciona um caráter dinâmico a ferramenta, e por permitir aplicações de Cadeias de Markov para análises quantitativas. Na Figura 6

temos um exemplo de uma representação gráfica em BDMP. A seta em vermelho pontilhado é o *trigger*, e significa que incialmente apenas a ocorrência de P1 e P2 são relevantes, e P3 e P4 irrelevantes. Porém, quando o evento G1 ocorre resultante de P1 e P2 é como se agora apenas o lado G2 existisse e o evento topo se torna dependente da ocorrência de P3 e P4. Convertendo para estados na cadeira de Markov, temos 3 estados: um primeiro considerando P1 e P2, um segundo estado considerando P3 e P4, e último estado considerando o de falha, onde o evento r é obtido.

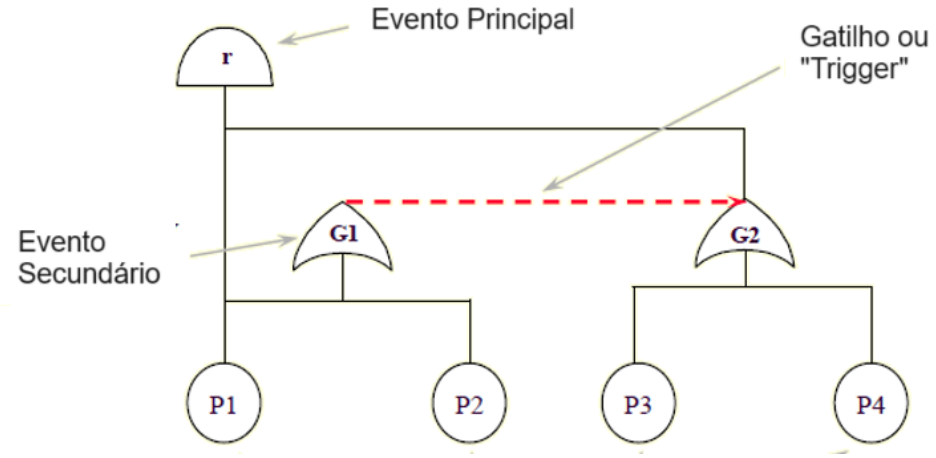

Figura 6 – Representação gráfica de um modelo BDMP
Fonte: Adaptado de Bouissou (2008)

Nas propriedades do BDMP, cada folha pode estar em dois modos: *required* ou *standby*, como temos um trigger de G1 para G2, antes da ocorrência G1, temos P1 e P2 em modo *required* e P3 e P4 em modo *standby*. Após G1 ocorrer, P3 e P4 vão para o modo *required*. Cada estado em cadeia de Markov pode ser associado a combinação da variação de modos entre os eventos básicos. No BDMP os reparos também podem ser considerados na análise, por exemplo, supondo que a partir da condição inicial o evento P2 ocorra, então G1 ocorre, o *trigger* é ativado, e o evento P3 e P4 ficam em *required*. Se P2 retornar ao funcionamento, então o *trigger* é desativado, retornando o sistema à condição inicial com somente P1 e P2 em *required*. Bouissou (2008) fornece mais detalhes com definições matemáticas do BDMP.

As adaptações de Pietre-Cambacedes e Bouissou (2010) adicionaram folhas com propriedades de proteção cibernética, *security*, onde um tipo representa uma tentativa de ataque, *Attack Action* (AA), e outras duas representam medidas de proteção, sendo uma temporizada *Timed Security Event* (TSE) e outra instantânea, *Instantaneous*

Security Event (ISE). Os modos das folhas são referenciados como *Idle* e *Active.* Para os eventos de segurança crítica, *safety*, foram divididos entre *Failure in Operation* e *Failure on demand*. Na Tabela 4 estão descritos os cinco tipos de eventos básicos.

Tabela 4 – Tipos das Folhas BDMP

Elemento	Tipo	Descrição
 Attack Action (AA)	Security	Representa um passo do atacante em direção ao objetivo. No modo *Idle* significa que o invasor ainda nem tentou executar o ataque. O modo *Active* corresponde que já houve tentativa.
 Timed Security Event (TSE)	Security	Representa um evento de proteção cibernética que retarda o progresso dos ataques de um agente malicioso.
 Instantaneous Security Event (ISE)	Security	Representa um evento proteção cibernética que pode ocorrer instantaneamente.
 Failure in Operation	Safety	Representa uma falha na operação quando está no modo *required*. O componente possui os parâmetros de probabilidade de falha e de reparo.
 Failure on demand	Safety	Representa uma falha quando a folha muda imediatamente do modo de *standby* para *required*. O componente possui os parâmetros de probabilidade de falha e de reparo.

Fonte: Adaptado de Pietre-Cambacedes e Bouissou (2010)

Kriaa et al. (2012) modelou o ataque cibernético STUXNET em BDMP considerando as adaptações de Pietre-Cambacedes e Bouissou (2010). Pela análise seria possível identificar os caminhos do ataque permitindo uma visualização sobre os pontos de adição barreiras cibernéticas. Czekster e Morisset (2021) exploram as características do BDMP para propor a ferramenta BDMPathfinder que executa múltiplas iterações para computar a quantidade de caminhos de ataques em modelos BDMP.

Uma simulação BDMP é realizada através da aplicação proprietária Risk Spectrum pertencente a empresa Risk Spectrum AB (RISK SPECTRUM AB, 2023). Na ferramenta, carrega-se as bases de conhecimento BDMP que foram escritas na Linguagem Figaro (KHAN et al., 2021). A partir desse procedimento, é possível montar os modelos BDMP, configurar os parâmetros dos módulos e a simulação irá gerar os estados de Markov automaticamente. As combinações que levam ao evento topo, *Minimal Cut Sets* (MCS), e as curvas de probabilidade falha são fornecidas como resultado da simulação.

4.3 ARQUITETURA S-CUBE

Kriaa, Bouissou e Laarouchi (2015) elaboraram a ferramenta S-CUBE voltada para sistemas de controle, supervisão e aquisição de dados (SCADA). A ferramenta fornece possíveis cenários de riscos oriundos dos domínios de *safety* e *security*. O S-CUBE tem como entrada a descrição da arquitetura e gera como saída os possíveis cenários com ataques cibernéticos e falhas que podem resultar em uma condição indesejável.

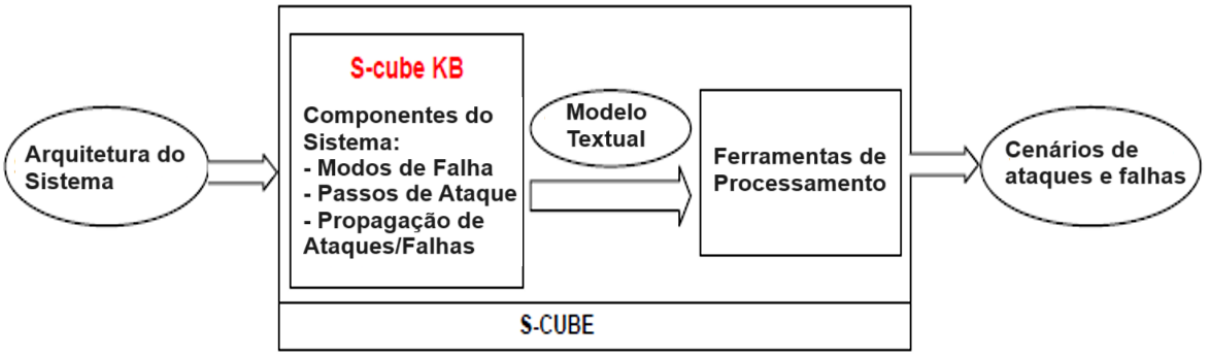

Figura 7 – Diagrama S-Cube
Fonte: Adaptado de Kriaa, Bouissou e Laarouchi (2015)

Na Figura 7, o componente S-CUBE KB representa o núcleo do sistema escrito na linguagem Figaro (KHAN et al., 2021), onde é possível representar os módulos típicos de infraestruturas industriais digitais com aspectos de segurança crítica e proteção cibernética (autenticação, controle de acesso, redundância) e as interações e dependência entre os módulos. Também são associados os tipos e probabilidades de ataques e falhas que cada módulo pode sofrer. Os resultados de saída do S-cube KB podem ser exportados para ferramentas de análise quantitativa como o *Yet Another*

Monte Carlo (YAMS) e o Figseq, uma ferramenta para cálculo de confiabilidade e disponibilidade de sistemas. Por fim, os cenários de ataques cibernéticos e de falhas são gerados para servirem de subsídios aos analistas de segurança crítica na tomada de decisões sobre o sistema.

O S-cube tem a vantagem em relação ao BDMP sobre o fato de poder gerar cenários automaticamente a partir da descrição da arquitetura do sistema. Enquanto, no BDMP temos que gerar as árvores de forma manual para obtermos as probabilidades dos estados e os *cut sets*, no S-Cube foca-se em descrever a arquitetura do sistema.

A ferramenta S-CUBE oferece uma boa oportunidade de expansão adaptando o módulo S-CUBE que está escrito na linguagem FIGARO para as características de novos módulos. Oueidat et al. (2022) apresentou um caso de uso do S-CUBE onde gera automaticamente cenários de ataques cibernéticos e de falhas para uma arquitetura industrial crítica.

4.4 PROCESSOS CHASSIS

Raspotnig et al. (2012) propôs o método *Combined Harm Assessment of Safety and Security for Information Systems* (CHASSIS) que é baseado em *Unified Modeling Language* (UML) e visa avaliar aspectos de segurança crítica (*safety*) e de proteção cibernética (*security*) de forma conjunta e com aplicabilidade na fase inicial do projeto de um sistema e tem como saída a definição de requisitos. Basicamente, este método é dividido em três etapas principais: Levantamento de Requisitos Funcionais, Levantamento de Requisitos de *Safety/Security* e Especificação de Requisitos de *Safety/Security*. Estas etapas e subatividades são ilustradas na Figura 8.

Na primeira etapa, denominada Levantamento de Requisitos Funcionais, são definidas as funções e serviços globais e elaborados os Casos de Uso em UML e os Diagramas de Sequência. Esta etapa ainda se subdivide em três atividades sequenciais: atividade (1) que correspondem aos Diagramas de Caso de Uso (D-UC), atividade (2) onde são detalhados os Casos de Uso Textual (T-UC) e a atividade (3) que são os Diagramas de Sequência (SD).

Na segunda etapa, Levantamento de Requisitos de *Safety/Security*, temos mais três atividades envolvidas: atividade (4) onde são definidos os diagramas casos de erros

de uso (D-MUC), atividade (5) onde os D-MUC são convertidos para escrita textual de erros de uso (T-MUC) e atividade (6) onde são gerados os diagramas sequenciais dos cenários de falhas (FSD) e de erros de uso (MUSD). Com os FSD e MUSD é possível identificar a necessidade de incluir contramedidas para *safety* e *security*, e reestabelecer tais medidas na atividade anterior e reeditar o T-MUC ou retornar para a atividade (1) editando D-UC.

A última etapa, Especificação de Requisitos de *Safety/Security*, possui duas atividades: atividade (7) responsável por coletar informações do T-MUC e criar uma tabela HAZOP estendida contemplando aspectos de *safety* e *security*. E por fim a atividade (8) que corresponde aos requisitos gerados englobando aspectos de ambas as áreas.

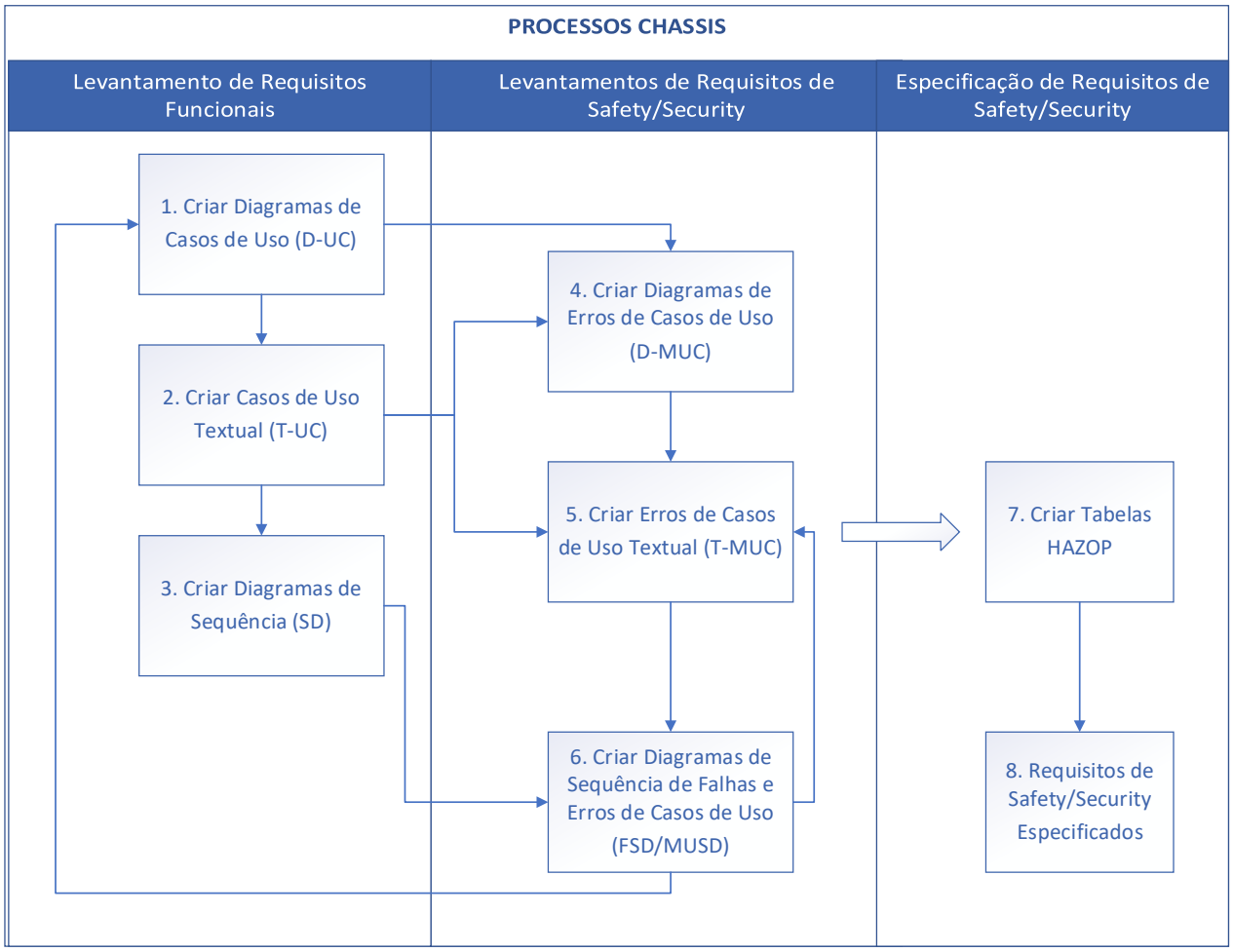

Figura 8 – Visão Geral dos Processos CHASSIS
Fonte: Adaptado de Raspotnig, Karpati e Katta (2012)

O resultado gerado pela abordagem CHASSIS corresponde a um conjunto de requisitos, que se forem plenamente atendidos, o sistema alcançará o mais alto nível de segurança crítica e proteção cibernética. Em outro estudo mais recente, Raspotnig, Karpati e Opdahl (2018) apresentaram uma aplicação em que o método CHASSIS foi bem-sucedido em capturar questões *Safety/Security* em um sistema de gerenciamento de tráfego aéreo.

Diferente do BDMP e do S-Cube, o CHASSIS permite apenas análises qualitativas. Se for necessário executar uma análise quantitativa, deve-se empregar outra ferramenta, que poderá até ser usada em complemento ao CHASSIS.

4.5 DIAGRAMAS GTST-MLD

Di Maio, Mascherona e Zio (2020) apresentaram o modelo *Goal Tree Success Tree Master Logic Diagram* (GTST-MLD) para analisar Sistemas Ciber Físicos (CPS) considerando possibilidade de falhas de componentes e ameaças cibernéticas. No estudo, também afirmaram que o GTST-MLD é útil para a análise quantitativa mesmo com lacunas de informações sobre probabilidade de ataques cibernéticos. Neste caso foi considerado o histórico de ataques cibernéticos ocorridos em sistemas SCADA e não em apenas ativos cibernéticos.

O GTST-MLD é uma abordagem orientada a objetivos devido ao fato de enfatizar a finalidade da função de segurança crítica. Seguindo um caminho *Top-Down,* a abordagem decompõe a função objetivo em subfunções e as relaciona com os componentes dos sistemas envolvidos no atendimento. A parte denominada de *Goal Tree* (GT) representa o lado da função principal decomposta em subfunções hierarquicamente relacionadas. As funções são decompostas guiada pelo questionamento de “como” a função superior pode ser alcançada. A função do nível mais abaixo é guiada pelo questionamento do “porquê’ é necessária para atender o nível acima. O lado denominado *Success Tree* (ST) representa a relação dos componentes do sistema com as funções do GT em todos os níveis de especificação. No lado ST temos como topo o Sistema, que é decomposto em uma combinação lógica de componentes. Também são representados os componentes auxiliares, que são necessários para operação dos componentes principais, que não possuem relacionamento direto com o GT.

Na Figura 9, a partir do elemento principal “System”, a elaboração da ST é guiada pelo questionamento de “quais são as partes” para decomposição até a representação dos componentes básicos do sistema. As conectividades entre as funções do GT e os componentes principais do ST são representadas pelo *Master Logic Diagram* (MLD). O MLD facilita essa representação que pode ser vista como uma matriz de conectividade entre elementos GT e ST (DI MAIO; MASCHERONA; ZIO, 2020).

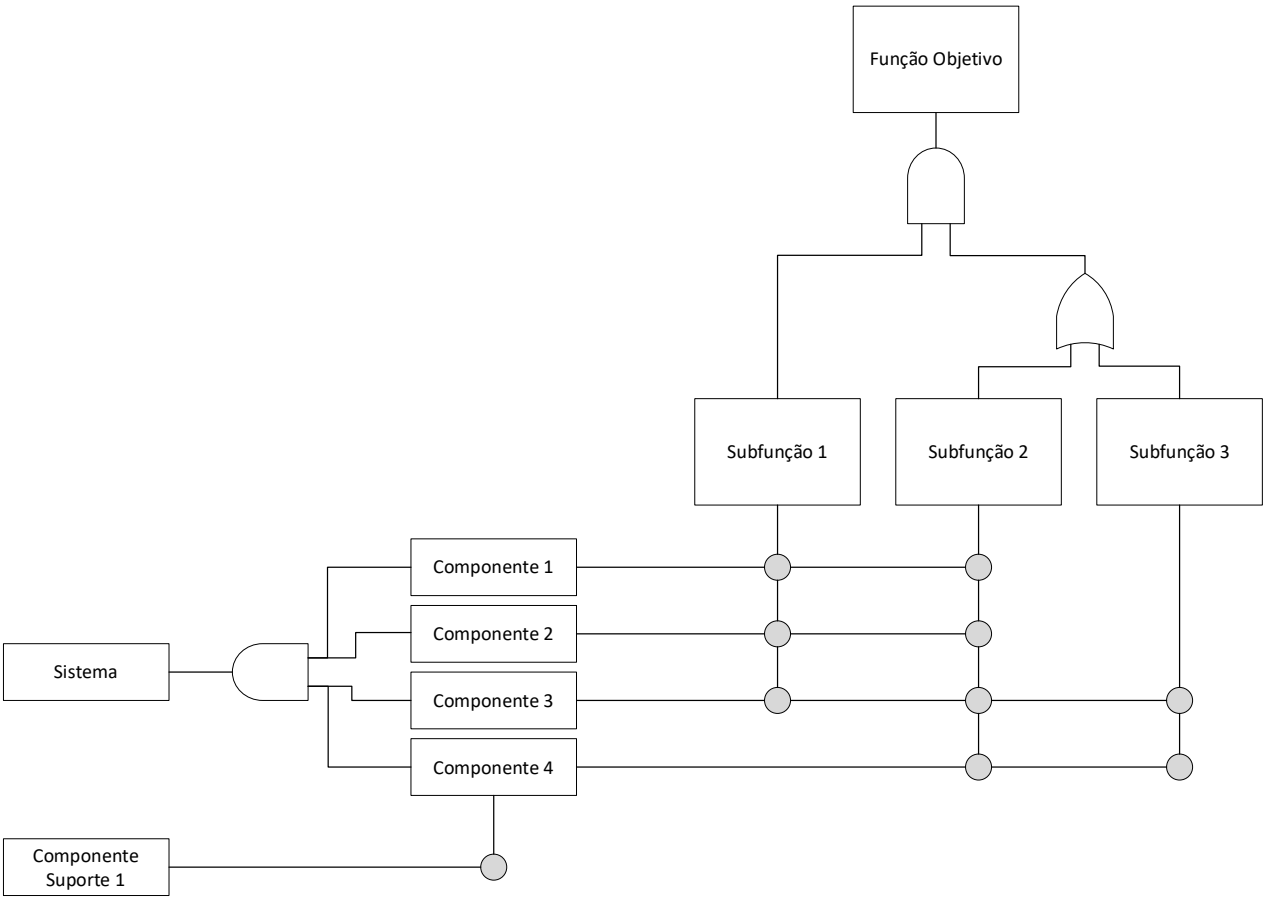

Figura 9 – Exemplo de uma estrutura GTST-MLD
Fonte: Adaptado Di Maio, Mascherona e Zio (2020)

Em ambos os lados GT e ST pode-se representar elementos de segurança cibernética. Essa abordagem também tem a vantagem de permitir uma construção mais intuitiva do diagrama já que são tratadas primeiro as funções de segurança mais óbvias de um sistema. Embora GTST-MLD permita análises quantitativas, também sofre o mesmo problema dos demais modelos BDMP e S-Cube a respeito da ausência de informação para probabilidades de ataques cibernéticos. O autor da proposta usa alguns parâmetros como fatores de influência em *security*, e algumas bases de dados que

mostram algumas estatísticas de ataques. Para análises qualitativas permite uma visualização facilitada das interações entre os módulos.

4.6 MÉTODOS STPA E STPA-SEC

O *System-Theoretic Process Analysis* (STPA) é um método de análise de perigos proposto como forma ampliada do *System-Theoretic Accident Model and Processes* (STAMP) que se trata de um modelo de causalidade de acidentes baseado na teoria de sistemas, ambos desenvolvidos por Nancy Leveson (LEVESON, 2014). No STAMP um sistema é tratado como uma estrutura de controle hierárquica onde as interações entre cada camada impõem as restrições necessárias ao comportamento de componentes na camada inferior imediata. Diferente de outras abordagens, no STAMP uma falha de segurança crítica na aplicação é associada a uma falha na estrutura de controle (LYU; DING; YANG, 2019).

O STPA corresponde então a técnica de análise de perigos na visão de modelos STAMP. Além das falhas de componentes, o STPA assume que os acidentes também podem ser causados por interações inseguras dos componentes do sistema mesmo que nenhum tenha falhado individualmente (LEVESON; THOMAS, 2018).

O STPA pode ser aplicado nas fases iniciais do projeto para auxiliar na identificação de requisitos e restrições de segurança crítica. A identificação na fase de concepção de um sistema contribui para eliminar retrabalhos dispendiosos quando são identificadas ou ocorrem falhas no final do desenvolvimento ou durante as operações. O STPA também inclui software e operadores humanos como elementos a serem analisados, garantindo que a análise de segurança crítica inclua todos os possíveis fatores que podem ter relação com uma falha sistêmica. O modelo também proporciona ao projetista documentar as funcionalidades dos sistemas de maneira mais clara e rastreável, que para sistemas grandes e complexos são extremamente úteis (LEVESON; THOMAS, 2018).

Leveson e Thomas (2018) ainda afirmam que muitas avaliações e comparações do STPA com métodos mais tradicionais de análise de perigos, como análise de árvore de falhas (FTA), *Failure Modes and Effects Criticality Analysis* (FMECA) e *Hazard and Operability Analysis* (HAZOP) foram realizadas. Em todas as análises, o STPA encontrou todos os cenários causais encontrados pelas análises tradicionais. Também

identificou mais cenários não capturados pelos métodos tradicionais, a maior parte relacionada a software e sem falhas individuais de componentes. Em alguns casos, onde houve um acidente do qual os analistas não foram informados, apenas o STPA encontrou a causa do acidente. Além disso, o STPA revelou-se muito menos dispendioso em termos de tempo e recursos do que os métodos tradicionais.

Em análise STPA deve-se considerar o sistema como um todo, analisando não apenas componentes individuais, mas também as interações entre eles e os processos que ocorrem no sistema. O objetivo é identificar cenários de falha que podem surgir devido a interações complexas entre elementos do sistema e os processos. Uma análise STPA é dividida em 4 etapas conforme ilustrado na Figura 10.

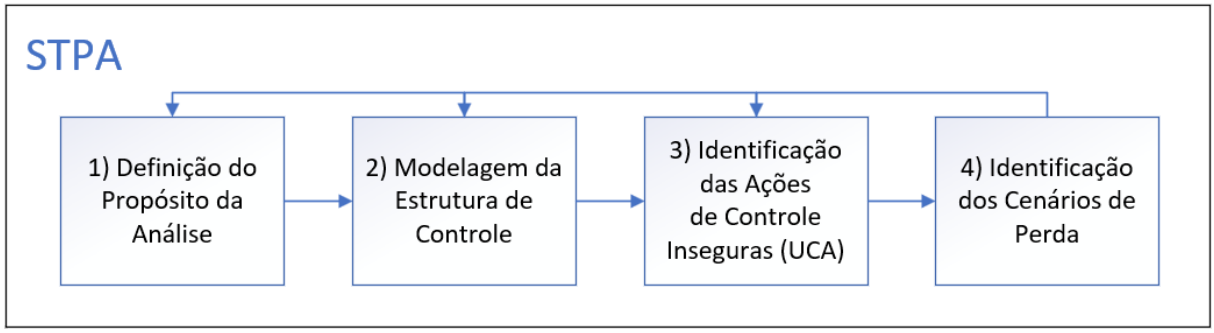

Figura 10 – Processos STPA
Fonte: Adaptado de Leveson e Thomas (2018)

A primeira etapa, Definição do Propósito da Análise, consiste em estabelecer quais perdas se pretende prevenir referente a aplicação. As perdas podem estar relacionadas a objetivos de segurança crítica e também a objetivos de negócios como desempenho ou outras propriedades. A partir da definição das perdas (*Losses*), define-se os perigos (*Hazard*s) a nível de sistema. Cada perda pode estar relacionada a vários perigos e cada perigo correlacionado a várias perdas. Em seguida, define-se as restrições relacionadas a cada perigo. Como saída dessa primeira etapa, temos a lista de perigos a nível de sistema, a lista de restrições e até opcionalmente uma lista de restrições e perigos secundários. A figura 11 a seguir ilustra os passos internos referente a primeira etapa de uma análise STPA.

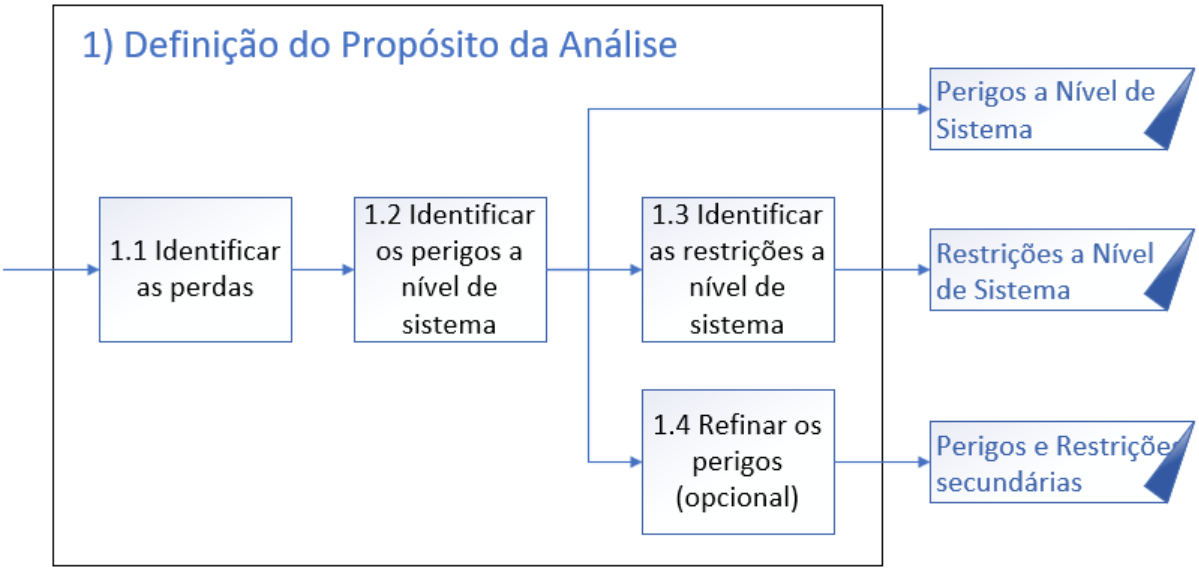

Figura 11 – Passos da Definição do Propósito da Análise
Fonte: Adaptado de Leveson e Thomas (2018)

A segunda etapa do STPA, Modelagem da Estrutura de Controle, corresponde a elaboração de malhas de estruturas de controles que serão responsáveis pela imposição das restrições a nível de sistema e das restrições secundárias nos elementos a serem controlados.

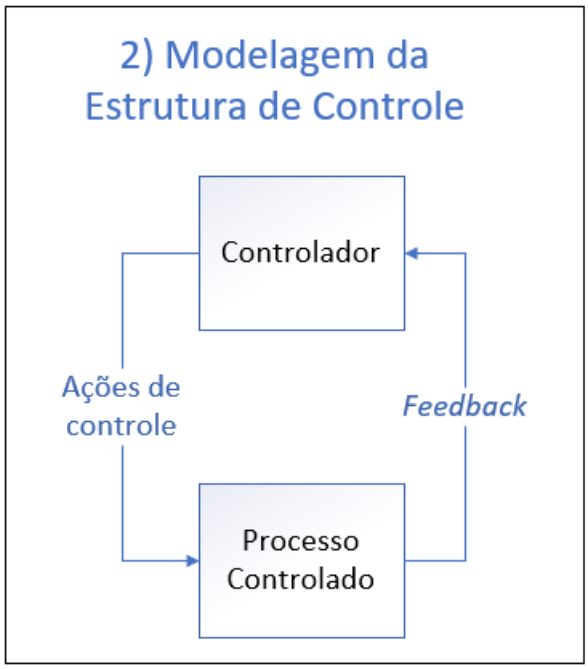

Figura 12 – Componentes básicos de uma estrutura de controle
Fonte: Adaptado de Leveson e Thomas (2018)

Um modelo mais básico de uma estrutura de controle, conforme ilustrado na Figura 12, é composto pelo controlador, pelas ações de controle, pelo processo controlado e pelos sinais de retorno (*feedback*). Porém, as malhas podem ser implementadas em cascata ou com trocas de informações em paralelo. Nesta etapa é assinalada uma lista de Responsabilidades para cada controlador. Essas responsabilidades são associadas ao atendimento das restrições identificadas na primeira etapa do STPA e auxiliam na identificação de qual elemento está associado a um perigo e por consequência a uma perda.

A terceira etapa do STPA refere-se ao levantamento de ações de controle insegura ou *Unsafe Control Actions* (UCA) que resultariam em um não atendimento a uma restrição imposta acionando um gatilho de um perigo. Cada ação de controle é analisada em quatro modos para enquadramento como UCA:

1. Não executada: verificação se a ação de controle não disparar quando solicitada pode ativar um perigo.

2. Executada: verificação se a ação de controle disparar quando não solicitada pode ativar um perigo.

3. Início de execução indevido: verificação se a ação de controle disparar muito cedo, muito tarde ou fora de ordem pode ativar um perigo.

4. Duração da execução: verificação se a ação de controle tiver duração prolongada ou for interrompida antes do tempo.

Quando uma ação de controle se enquadra em pelo menos um dos modos, então a ação deve ser incluída da lista de UCA. Cada registro de UCA deve conter cinco partes: a fonte (elemento), o tipo (executado, não executado, fora de ordem), a ação de controle propriamente dita, o contexto (ou estado da aplicação) e qual perigo ativado. Esta etapa ainda permite a definição de novas restrições para o controlador que pode ser implementado no sistema ou levada ao conhecimento do operador da plataforma crítica. A Figura 13 ilustra as entradas e as saídas desta etapa.

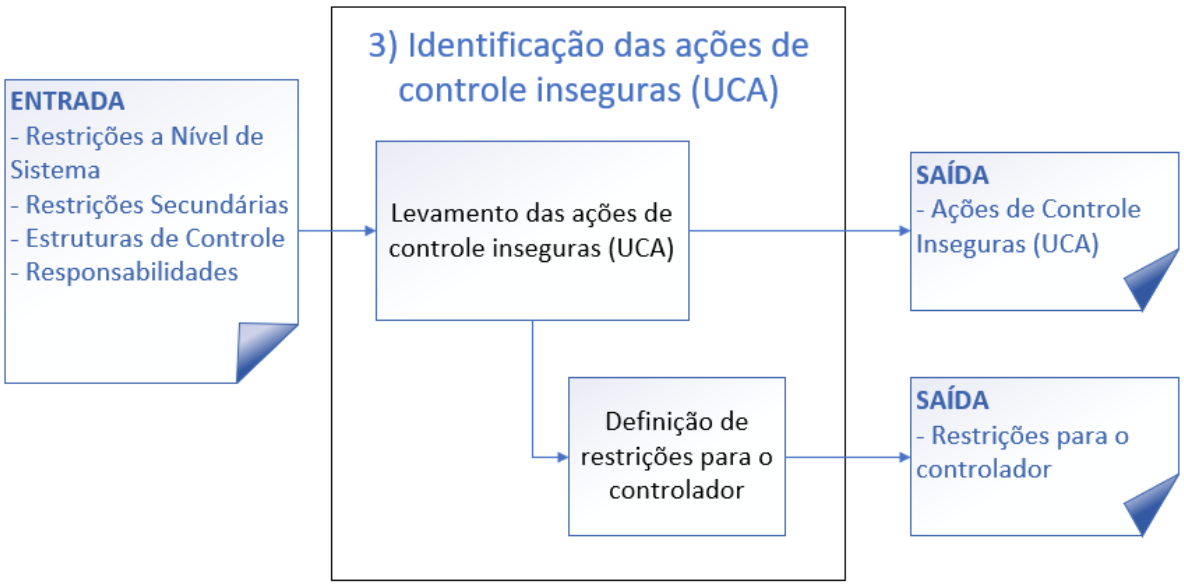

Figura 13 – Identificação das Ações de Controle Inseguras (UCA)
Fonte: Adaptado de Leveson e Thomas (2018)

A quarta e última etapa do STPA, Identificação dos Cenários de Perda, corresponde ao levantamento dos fatores que podem disparar as ações de controle inseguras (UCA) capturadas na etapa anterior. Os fatores considerados devem ser abrangentes ao ponto de analisar, por exemplo, possibilidade de falhas físicas em equipamentos, interrupção de energia, exposição a altas e baixa temperaturas, exposição a umidade, falhas em implementação de algoritmos, falhas de comunicação entre dispositivos, recebimento de *feedback* inconsistentes e também possibilidade de ataques cibernéticos. Além desses fatores, também devem ser analisados e documentados ações indevidas pelo operador ou tripulação da plataforma crítica. A Figura 14 ilustra a entrada e saída da etapa 4. Com os resultados dos cenários de perdas, o analista pode tomar a decisão para o tratamento ou aceitação do risco, podendo refazer todo ciclo do STPA novamente se julgar pertinente.

O STPA permite apenas uma análise qualitativa. Se for necessária a obtenção de resultados quantitativos então deve-se adotar outro método que pode ser complementar ao STPA, como o BDMP, S-Cube ou GTST-MLD.

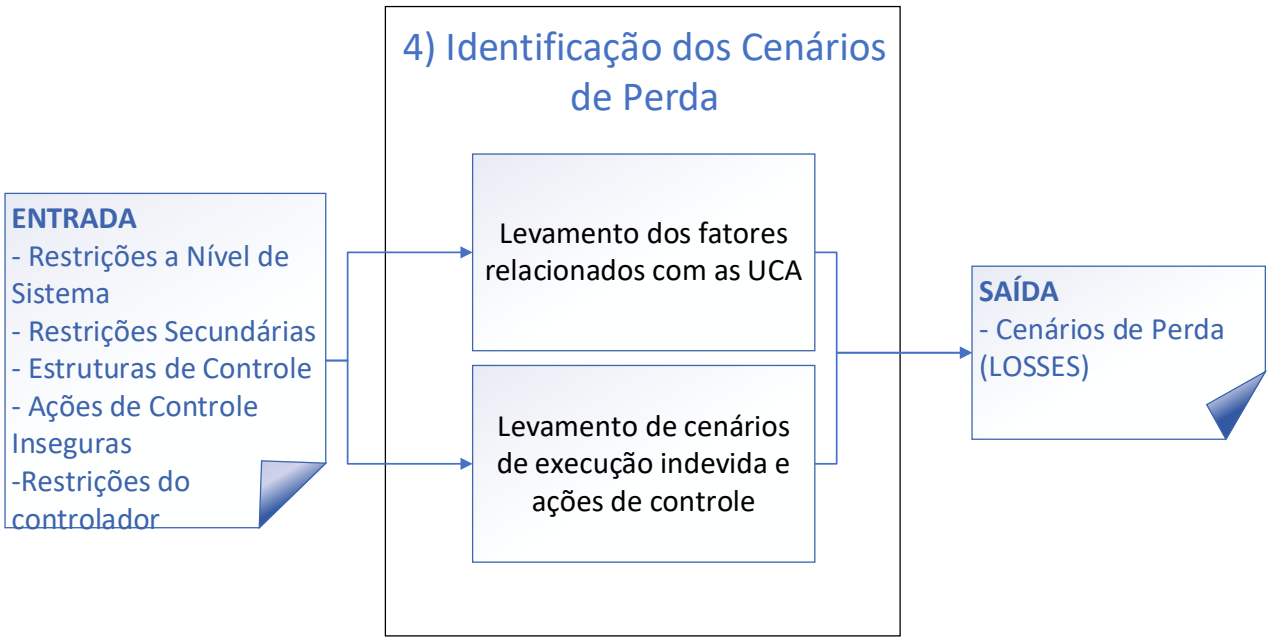

Figura 14 – Identificação dos Cenários de Perda
Fonte: Adaptado de Leveson e Thomas (2018)

O STPA, em seu modelo oficial vigente, já é capaz de considerar aspectos de segurança cibernética, mas ainda assim, há várias propostas com ajustes em seus processos com o objetivo de uma melhor adaptação as questões de vulnerabilidades cibernéticas. Oficialmente, Young e Leveson (2014) desenvolveram a versão adaptada para aspectos de segurança cibernética denominada de STPA-Sec. Essa nova versão adiciona os passos de:

1. Levamento de Problemas de Cibersegurança a nível de sistema: esse passo é realizado em conjunto com o Levantamento de Perigos no Processo inicial do STPA;

2. Identificação das ações inseguras no contexto cibernético: realizada em conjunto com o levantamento das UCA;

3. Realização de *Wargaming* ou jogo de guerra: trata-se de uma atividade prática onde os analistas são divididos em duas equipes: *Red Team* responsável por lançar ataques e descobrir vulnerabilidades no sistema e *Blue Team* responsável pela defesa.

Além das etapas mencionadas são considerados outros aspectos como a Gestão do Ciclo de Vida de Informação que é distribuída em sub atividades por todo o processo.

Pereira, Hirata e Nadjm-Tehrano (2019) apresentam um estudo onde aplicam uma versão ampliada do STPA capaz de capturar mais vulnerabilidades cibernéticas adaptando as quatros etapas do STPA. Uma das adaptações consiste em relacionar todos os fatores casuais com as propriedades de segurança da informação e assinalando qual se aplica e se há contramedidas. Fan e Yang (2022) expõe a necessidade de aplicar método como o STPA-Sec para análises voltadas a sistemas de transporte devido à alta digitalização, que já não são plenamente cobertas pelo método tradicional. Aktouche et al. (2021) aplicam o STPA-Sec para análise de um sistema de condução remota para veículos em trilhos considerando aspectos de riscos cibernéticos junto com a caraterística do STPA de analisar o sistema como um todo onde falhas podem ocorrer a partir de erros de interações.

4.7 OUTROS ESTUDOS CORRELATOS

As preocupações com ameaças cibernéticas em plataformas críticas não iniciaram recentemente. Existem diversas publicações que tratam o tema com variadas técnicas sugeridas ao longo do tempo. Lisova et al. (2019) apresentaram um estudo que trata de uma revisão sistemática da literatura para análise integrada de *safety e security* e o objetivo era identificar o estado atual do desenvolvimento das abordagens integradas. Como resultado, foram destacadas um total de trinta e três abordagens datadas entre 2012 e 2017, entre elas, as analisadas neste estudo como BDMP, S-Cube, CHASSIS e STPA-Sec. LISOVA et al. (2019) apresentam um quadro comparativo das abordagens mostrando se está associada com alguma norma, se tem origem acadêmica ou na indústria e se a área é específica como na aviação ou genérica.

Kriaa (2016) ao desenvolver o S-Cube também faz um levantamento de diversas abordagens anteriores que tratam *safety* e *security* de forma integrada. Em seu estudo a autora também menciona STPA-Sec, CHASSIS e BDMP. Kriaa (2016) compara as metodologias em critérios propostos para uma metodologia ser considerada ideal como permitir análise integrada, ter fundamentação matemática, produzir resultados qualitativos e quantitativos, fornecer identificação automática a partir de uma entrada de dados e prover flexibilidade em suas análises quando a aplicação sofrer alterações.

Fan e Yang (2022) fizeram o levantamento de metodologias de análise integrada de segurança crítica e segurança cibernética e compara suas aplicações no contexto de sistemas de transportes. No estudo também são apontados os métodos CHASSIS e STPA-Sec. Na sua comparação Fan e Yang (2022) exploram a classificação das abordagens integradas orientadas para *safety* e as orientadas para *security*.

Pekaric et al. (2023) apresentam um estudo de revisão sistemática do estado da arte de análise de segurança crítica e segurança cibernética para sistemas auto adaptativos (SAS). Os SAS por terem a capacidade de adaptar seu comportamento, modificando sua arquitetura e sua estrutura lógica de acordo com estímulos oriundos do ambiente, introduzem significativa complexidade na análise de segurança. Dos estudos analisados Pekaric et al. (2023) afirmaram que embora ataques cibernéticos já estejam endereçados na análise de SAS, ainda são necessários aprimoramentos nas abordagens para cobrir o processo de adaptação.

Fovino, Masera, De Cian (2009) propuseram uma adaptação à Análise de Árvores de Falhas (FTA) para incluir Árvores de Ataques (AT). O modelo foi batizado de Árvore de Falhas Estendidas e considera falhas acidentais e ameaças cibernéticas. A proposta visou possibilitar análises quantitativas também para eventos de segurança cibernética.

George e Renjith (2021) apresentam uma revisão de aplicações com redes Bayesianas na avaliação de riscos de segurança crítica e segurança cibernética em processos industriais. O estudo investiga os métodos gerais para o desenvolvimento de redes Bayesianas, apoiando plataformas de software, técnicas de validação, metodologias de análise de sensibilidade. Além disso, destaca os pontos fortes e as limitações do uso de redes Bayesianas comparados com as metodologias.

Yarza et al. (2022) apresentam uma estrutura para análise integrada de segurança crítica e segurança cibernética voltada para dispositivos computacionais embarcados de alto desempenho. A metodologia proposta leva em consideração as propriedades físicas e lógicas de dispositivos de lógica programável (FPGA) mapeando fontes de falhas e vulnerabilidades intrínsecas do dispositivo. O objeto é contribuir para o desenvolvimento de um algoritmo embarcado mais robusto ao indicar para o projetista os pontos de mitigação.

Busquim e Silva et al. (2021) apresentam um estudo que visa avaliar aspectos de segurança cibernética para usinas nucleares. A análise é feita por meio de um simulador que identifica impactos funcionais causados por ataques cibernéticos direcionados a ativos digitais específicos em uma instalação nuclear. O estudo conclui que os resultados da simulação facilitam o entendimento de como um ataque cibernético propaga, quais as consequências e quais contramedidas podem ser implantadas para proteger uma instalação nuclear.

Vismari (2023) apresenta uma proposta de abordagem orientada a riscos para o gerenciamento de recursos de comunicação em Sistemas de Transporte. A abordagem permite tratar durante a operação, situações inseguras não previstas na fase de projeto. Os requisitos de segurança podem ser alterados pela operação do sistema. Para trabalhos futuros, é sugerida a investigação de aspectos de segurança cibernéticas nos impactos sobre a eficiência da abordagem proposta.

Ramos e Camargo Júnior (2023) destacam que plataformas críticas, ênfase em navios e sistemas marítimos, possuem suas operações cada vez mais dependentes de sistemas de informação com novas tecnologias, tornando-as suscetíveis a ataques cibernéticos com efeitos físicos. Essa atualização em tais projetos requer que análises de segurança cibernética e de segurança crítica sejam feitas de forma conjunta. Como demonstração, o artigo apresenta um estudo de caso analisado por cadeias de Markov onde mostra que uma implementação feita para corrigir uma vulnerabilidade cibernética pode apresentar um impacto negativo no nível de segurança crítica do sistema. O resultado, portanto, sugere que análises integradas são importantes para identificar conflitos entre propriedades de segurança crítica e segurança cibernética, que não seriam identificadas se analisadas de forma separada.

Nicoletti et al. (2023) apresentam um estudo de revisão da literatura dos formalismos baseados em modelos que contemplam análises conjuntas de segurança crítica e segurança cibernética. Os formalismos selecionados são comparados em critérios como a capacidade de modelagem, onde são analisados quais tipos de representações podem ser expressos e até que ponto são capazes de capturar as relações de ambas as áreas. Outro critério de comparação é a capacidade analítica que visa analisar quais tipos de análises o formalismo suporta. O último critério de comparação trata-se da aplicabilidade prática, onde visa identificar até que ponto os

formalismos podem ser utilizados para analisar sistemas de diferentes tamanhos e complexidades. O estudo apresenta como conclusões alguns pontos como a maioria dos formalismos combinam técnicas já existentes sendo que muitas são semelhantes as árvores de falhas, ainda não há um foco em novas construções para modelar as interações entre segurança crítica e segurança cibernética e os formalismos apresentados podem capturar diferentes resultados dessas interações. Os autores destacam, portanto, que há ausência de estudos que visam identificar as relações entre ambas as áreas.

5 ESTUDO DE CASO 1 – SISTEMA *ANTI-HEELING (AH)*

O estudo de caso apresentado neste capítulo corresponde a um exemplo de sistema automático *Anti-Heeling* que tem como função manter a estabilidade de um navio controlando os níveis de água dos tanques de lastros. Incialmente, o sistema é analisado sob a ótica dos métodos tradicionais amplamente aceitos, FTA e MA, com o objetivo de comparar resultados qualitativos e quantitativos com o método BDMP. Continuando, aplica-se análises sob a ótica do S-Cube, CHASSIS, GTST-MLD e STPA-Sec. Em todas as análises são consideradas as possibilidades de falhas acidentais e de ataques cibernéticos.

5.1 DESCRIÇÃO DO SISTEMA AH

O sistema *Anti-Heeling* é responsável por detectar o ângulo de inclinação de um navio e ajustá-lo calculando o nível de água nos tanques de lastro, a finalidade é evitar que o navio tombe e afunde. A Figura *15* ilustra um diagrama exemplo de Sistema *Anti-Heeling* com um tanque de lastro (*Ballast Tank*) em cada lateral do navio.

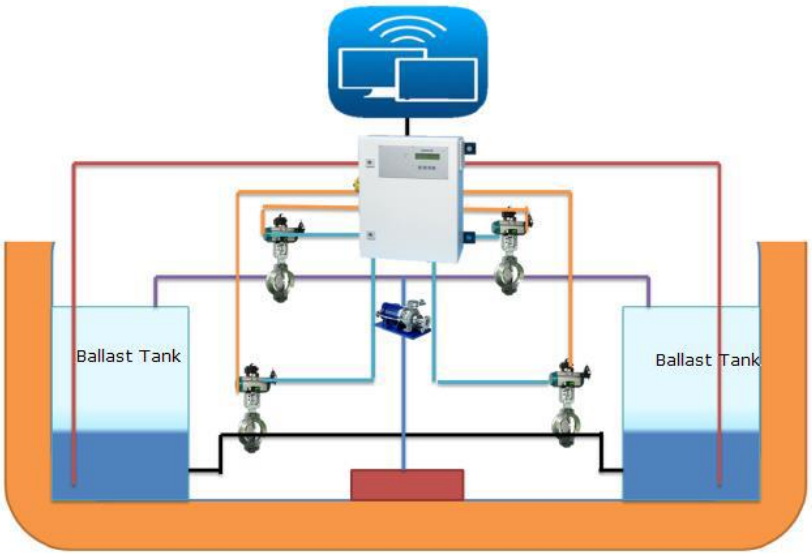

Figura 15 – Ilustração de um sistema *Anti-Heeling*
Fonte: Atlas Marine (2023)

Um sistema *Anti-Heeling* pode ser composto por:

1) Sensores: identificam o ângulo do navio e o envia para a unidade de processamento;
2) Tanques de lastro: são compartimentos localizados nas extremidades do navio, laterais esquerda (bombordo) e direita (boreste), ou frontal (proa) e traseira (popa) e que podem ser preenchidos com água e ar para proporcionar estabilidade ao navio;
3) Atuadores: responsável pela operação nos tanques de lastro, recebem o comando da unidade de processamento para aumentar ou reduzir o nível de água nos tanques;
4) Unidade de Processamento: recebe dados de sensores, realiza o cálculo do nível de água de cada tanque e os envia para os atuadores; e
5) Diversos módulos mecânicos e eletromecânicos como válvulas controladoras de vazão, válvulas e bombas de pressurização, tubulações, medidores de nível água. Além ainda de visores de monitoramento e alarmes e módulo para operação manual.

Para efeito mais didático neste estudo de caso, a modelagem foi simplificada considerando apenas os elementos correspondentes aos sensores, ao computador de bordo e aos atuadores. Consideremos também subsistemas como o de operação do tanque de lastro com sensor de nível já incorporados no módulo atuador. O diagrama da planta de controle onde o processo controlado corresponde ao próprio navio é ilustrado na Figura 16.

A embarcação sofre os efeitos físicos de ondas marítimas e variações de distribuição de carga interna e necessita estar equilibrada. O ângulo de inclinação do Navio não pode ultrapassar um limite tolerado, caso contrário, o navio emborcará. Falhando, portanto, no atendimento ao requisito de segurança crítica. Esse dado de inclinação é identificado pelos sensores e enviado para o computador de bordo. Em seguida, o algoritmo de controle do computador realiza o processamento tendo como saída a instrução para ajuste nos níveis de água em cada tanque. A instrução é enviada para os atuadores que por sua vez executam a ação mecânica nos tanques. O ciclo é feito novamente com uma nova identificação do ângulo de inclinação do navio. Também são considerados apenas dois tanques de lastro sendo um localizado a bombordo (lado esquerdo) e outro a boreste (lado direito).

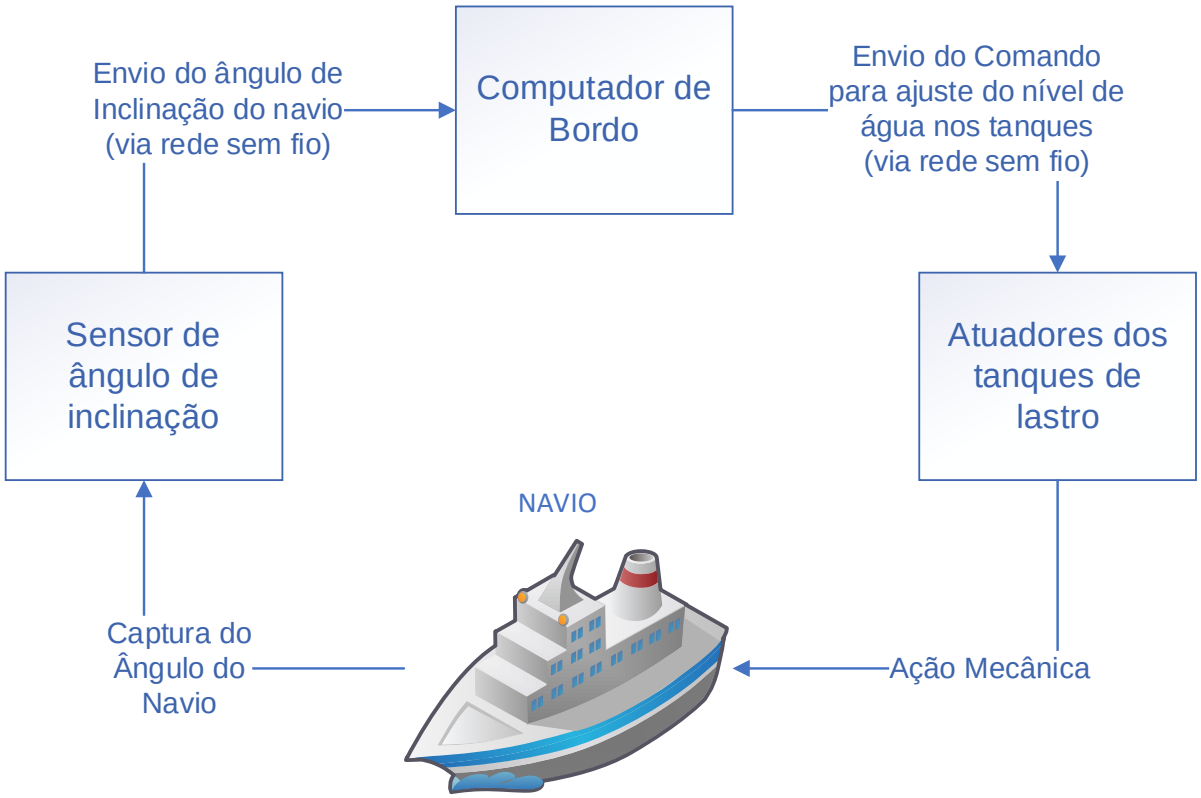

Figura 16 – Planta de Controle do *Anti-Heeling*

As comunicações entre os sensores, o computador e os atuadores são por meio de uma rede sem fio, suscetível a intervenções maliciosas como os ataques cibernéticos, *jamming* e *spoofing (*explicados na Tabela 1). Um ataque de *jamming* deixaria o Sistema Anti-Heeling indisponível e menos tolerante a uma onda marítima e variações de cargas internas. Um ataque de *spoofing* bem sucedido permitiria uma manipulação indevida nos níveis de água no tanque e poderia fazer o navio tombar até em uma situação de repouso, sem precisar que ocorra uma onda marítima ou variação.

5.2 APLICAÇÃO FTA AO ESTUDO DE CASO 1

O intuito da aplicação do método de Árvore de falhas é obter os resultados qualitativo que correspondem aos *Minimal Cut Sets* (MCS) ou grupos de cortes mínimos que são tradicionalmente utilizados em uma análise de segurança crítica. Tais dados servirão de subsídios para comparação com os demais outros métodos. A Árvore de Falhas construída para o Sistema *Anti-Heeling* é representada na Figura 17 e considera os dois tipos de ataques cibernéticos mencionados, *jamming* e *spoofing.* Nesta árvore é

observado que é possível atingir o Evento Topo “Navio Tombado” a partir de um “Ataque Spoofing” bem-sucedido e por uma operação maliciosa nos tanques de lastro representados por “Controle dos Tanques”.

A folha “Ataque Jamming” impacta a indisponibilidade do Sistema Anti-Heeling, representada por “AH Indisponível”. Porém, o evento topo ainda precisa que aconteça o evento “Variação de Carga”, que representa uma variação do ponto de equilíbrio do navio, que pode ser causado simplesmente por movimentação de cargas internas ou por ondas marítimas intensas.

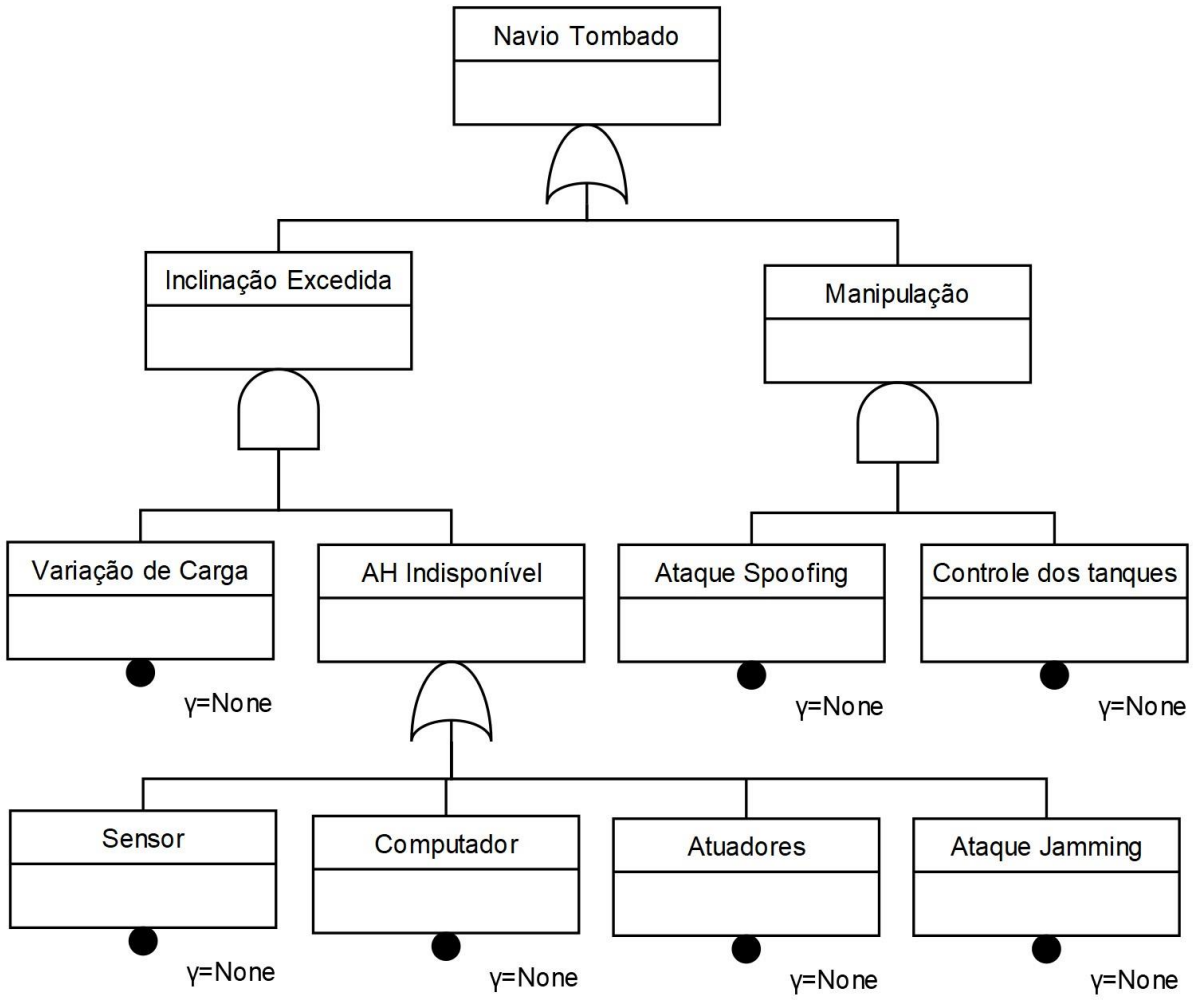

Figura 17 – Árvores de falhas para *o Anti-Heeling* System

Os MCS obtidos para a árvore de falhas construída mostrada na Figura 17 estão representados na Tabela *5*:

Tabela 5 – FTA: MCS do Sistema AH

Tipo	Nº de Eventos	Eventos
Security	2	“Ataque Spoofing” e “Controle dos Tanques”
Safety	2	“Sensor” e “Variação de Carga”
Safety	2	“Atuadores” e “Variação de Carga”
Safety	2	“Computador” e “Variação de Carga”
Security e Safety	2	“Ataque Jamming” e “Variação de Carga”

Foram identificados cinco *cut sets* ou grupo de cortes sendo todos com combinações de dois eventos. Os eventos “Ataque *Spoofing*” e “Controle dos Tanques” estão intimamente relacionados pois podem ser interpretados como uma progressão do ataque cibernético. Um ataque spoofing é um requisito para que se haja a manipulação do nível dos tanques. Neste cenário todos os eventos são de *security* permitindo a um agente malicioso o alcance do evento topo.

O evento “Ataque *Jamming*” deixaria o sistema AH indisponível, mas para atingir o evento topo “Navio Tombado” ainda seria necessário que o evento “Variação de Carga” ocorresse. Este evento representa a probabilidade de ocorrer uma onda marítima acima de um limite tolerável ou uma distribuição de carga interna do navio também acima de uma certa capacidade.

Para este cenário, qualitativamente podemos afirmar que um ataque *spoofing* tem uma gravidade superior a um ataque *jamming* e também superior aos demais *cut sets* de *safety* que representam falhas acidentais. Assim, uma contribuição desse resultado é direcionar os projetistas para priorizar a implantação de medidas de proteção contra ataques *spoofing*.

5.3 APLICAÇÃO MA AO ESTUDO DE CASO 1

Nesta etapa são aplicadas análises por cadeias de Markov (MA) para realização de uma análise quantitativa. Inicia-se com um diagrama que representa a continuidade operacional do sistema mostrado na Figura 18 para auxiliar na montagem dos estados e equações. Neste diagrama assumimos que se o bloco M1 falhar, isso representaria um ataque de *spoofing* completo, e então o requisito de segurança crítica não é atendido.

Com M1 funcionando, o requisito de segurança crítica não é atendido somente se M4 e M2, ou M4 e M3 falharem. A equação de continuidade operacional resulta na Eq. (1):

$$\mathrm{Opr} = \mathrm{M1}\,(\mathrm{M2M3} + \mathrm{M4}) = \mathrm{M1M2M3} + \mathrm{M1M4} \quad (1)$$

Pela equação (1) podemos visualizar 3 estados: P1 (que significa que os módulos M1, M2 e M3 estão funcionando), P2 (que significa que somente os módulos M1 e M4 estão funcionando) e PF (que significa o estado de falha, Navio Tombado). Para o requisito de segurança crítica, apenas os estados P1 e P2 que indicam que o navio não tombou.

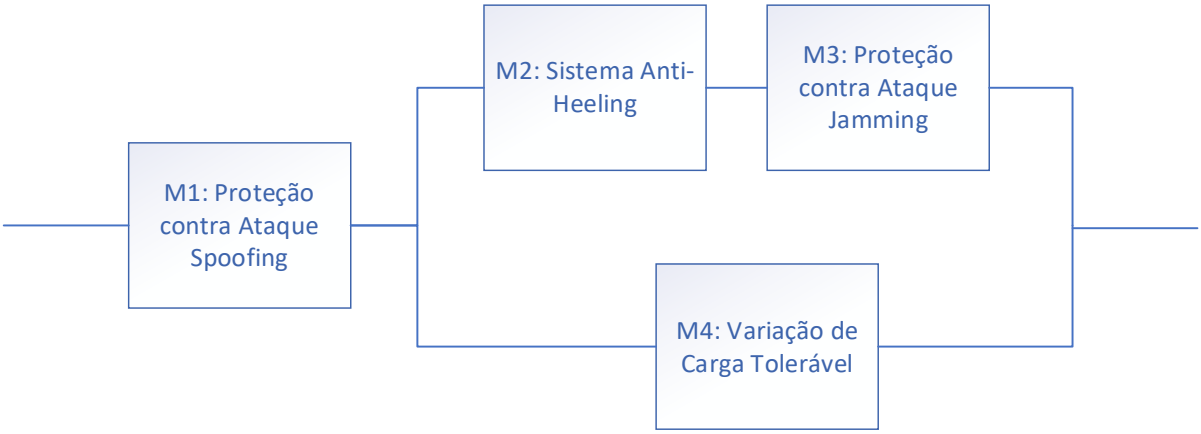

Figura 18 – Representação da continuidade operacional do sistema

Não são considerados os atributos fator de cobertura e taxa de reparo, e para os demais atributos, são especificados a seguir:

- λ1 é a probabilidade de M1 falhar, o que significa que ocorreu um ataque de *spoofing* completo e bem-sucedido.
- λ2 é a probabilidade de M2 falhar, o que significa que um dos ativos de tecnologia da informação: sensor, computador de bordo ou o atuador não estão funcionando, tornado o sistema *Anti-Heeling* indisponível.
- λ3 é a probabilidade de M3 falhar, o que significa que ocorreu um ataque de *jamming* bem-sucedido.
- λ4 é a probabilidade de M4 falhar: significa que ocorreu uma variação de carga acima dos limites tolerados pela física do navio.

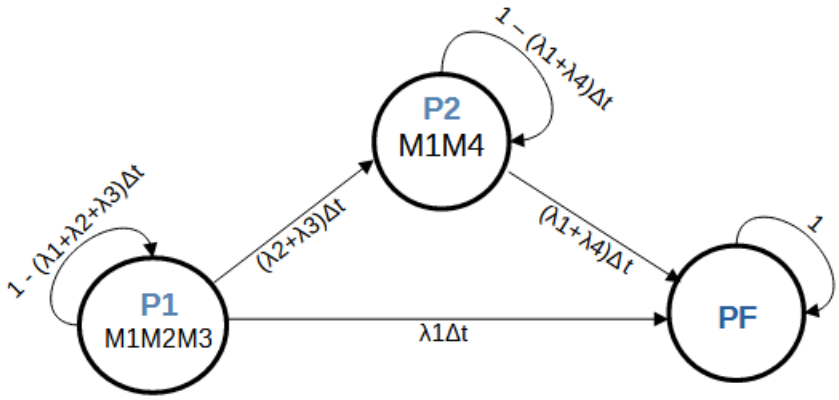

Figura 19 – Sistema AH em Cadeias de Markov

A Figura 19 representa o sistema *Anti-Heeling* em Cadeias de Markov, e Δt representa a variação no tempo. O próximo passo é elaborar a equação de probabilidade de cada estado:

$$P1(t + \Delta t) = (1 - (\lambda 1 + \lambda 2 + \lambda 3)\Delta t)\, P1(t) \quad (2)$$

$$P2(t + \Delta t) = (\lambda 2 + \lambda 3)\Delta t P1(t) + (1 - (\lambda 1 + \lambda 4)\Delta t) P2(t) \quad (3)$$

$$PF(t + \Delta t) = \lambda 1 \Delta t P1(t) + (\lambda 1 + \lambda 4)\Delta t P2(t) + PF(t) \quad (4)$$

Agora são transformadas as Eq. (2), Eq. (3) e Eq. (4), em equações diferenciais Eq. (5), Eq. (6) e Eq. (7) respectivamente:

$$dP1(t)/dt = (\lambda 1 + \lambda 2 + \lambda 3) P1(t) \quad (5)$$

$$dP2(t)/dt = (\lambda 2 + \lambda 3) P1(t) + (\lambda 1 + \lambda 4) P2(t) \quad (6)$$

$$dPF(t)/dt = \lambda 1 P1(t) + (\lambda 1 + \lambda 4) P2(t) \quad (7)$$

Foi elaborado um script no MATLAB, código fonte no Apêndice A, onde as equações diferenciais são usadas como entrada, os valores de λ1, λ2, λ3 e λ4 podem ser ajustados de acordo com a estimativa levantada pelo analista. As etapas do *script* são descritas na Tabela 6, as taxas de falhas são alteradas em cada simulação.

Tabela 6 – MA: Etapas na elaboração do Script em MATLAB

Passo	Descrição
1	Declare as variáveis de falhas e ajuste os valores para a simulação: λ1=0.001, λ2 = 0.0005, λ3= 0.001 and λ4= 0.005;
2	Declare as funções de probabilidade com o tipo simbolico: syms P1(t) P2(t) PF(t);
3	Declare e inicialize as equações diferenciais: edP1 = diff(P1,t) = = -(λ1+λ2+λ3)*P1; edP2 = diff(P2,t) = = (λ2+ λ3)*P1-(λ1+λ4)*P2; edPF = diff(PF,t) = = λ1*P1+(λ1+l λ4)*P2;
4	Declare os vetores com as equações diferenciais: edos = [edP1; edP2; edPF];
5	Entre com os valores das condições iniciais, assumindo que o estado inicial é P1, todos os módulos funcionando: cond = [P1(0) = =1; P2(0) = =0; Pf(0) = =0];
6	Execute a função dsolve para obter os valores das equações diferenciais: [P1sol(t), P2sol(t), PFsol(t)] = dsolve (edos,cond);
7	Some os estados P1 e P2 que representam a condição que o requisito de segurança crítica ainda é atendido, ou seja, o navio mantém-se estável: Safe=P1sol+P2sol;
8	Execute a função para mostrar o gráfico da função Safe (t)

Observação: o caractere λ não é aceito no código MATLAB.

Neste estudo de caso foram levados em consideração os seguintes tipos de cenários para o sistema Anti-Heeling:

- Com contramedidas de segurança cibernéticas implementadas: com aplicação de criptografia e processos de autenticação para a comunicação entre os módulos. Neste caso, o λ1 e λ3 tendem a reduzir drasticamente enquanto o λ2 tende a aumentar devido ao aumento de latência, o λ4 não é influenciado por ser evento de ambiente externo.
- Sem as contramedidas de segurança cibernéticas: todas as comunicações são deixadas em modo claro. Neste caso, λ1 e λ3 tendem a aumentar drasticamente enquanto λ2 tende a diminuir, o λ4 mais uma vez não é influenciado por se tratar de evento do ambiente externo.

Estimar a probabilidade de ataque cibernético é um grande desafio para análises quantitativas. Os valores aqui utilizados foram encontrados em uma faixa onde há antagonismo entre *safety* e *security* e podem representar casos reais. Em qualquer

faixa de valores podemos sempre afirmar que que λ1 e λ3 são inversamente proporcionais a λ2.

Para a simulação A, sem criptografia e sem autenticação, foram arbitrados os valores: λ1=0,0001, λ2=0,0005, λ3=0,0001 e λ4=0,0005, esses valores poderiam ser obtidos a partir das referências apresentados na seção 2.5. O resultado no tempo 1000 da função *Safe* do sistema foi 0,5635 (Figura 20).

Na simulação B, com criptografia e com autenticação, foram atribuídos os valores: λ1=0,000001, λ2=0,001, λ3=0,000001 e λ4=0,0005. Foram considerados que os valores de λ1 e λ3, que representam as chances de M1 e M3 falharem, reduzem drasticamente. Essas relações podem ser extraídas das bases *Repository of Industrial Security Incidentes* (ISI) de Byres, Lowe e Leversage (2015) ao considerar que Sistemas Industrias sem nenhuma medida de proteção cibernética aplicada têm probabilidades de serem atacados significativamente mais altas do que aqueles que adotam proteções. Para o valor de λ2, que representa indisponibilidade do sistema AH, estimou-se que dobraria de valor ao do cenário A devido a implementação das medidas de segurança cibernética. Essa relação está fundamentada no estudo de análise de perda de pacotes em sistemas de controle sem fio, de Santos et al. (2017), onde são constatados que um aumento no tamanho das mensagens de comunicação, influencia nas chances de perda de pacotes. Como resultado da simulação B no tempo 1000 da função “Safe*”* foi de 0,4574 (Figura 21), pior que a simulação A. Neste cenário, as medidas de segurança cibernética aumentaram as chances do navio tombar.

Na simulação C, sem criptografia e sem autenticação, mas com alta probabilidade de ataques cibernéticos foram atribuídos: λ1=0,001, λ2=0,0005, λ3=0,001 e λ4 =0,0005, e o resultado no tempo 1000 da função “Safe” é 0,1162 (Figura *22*). Nesse cenário, as medidas de segurança cibernética seriam eficazes.

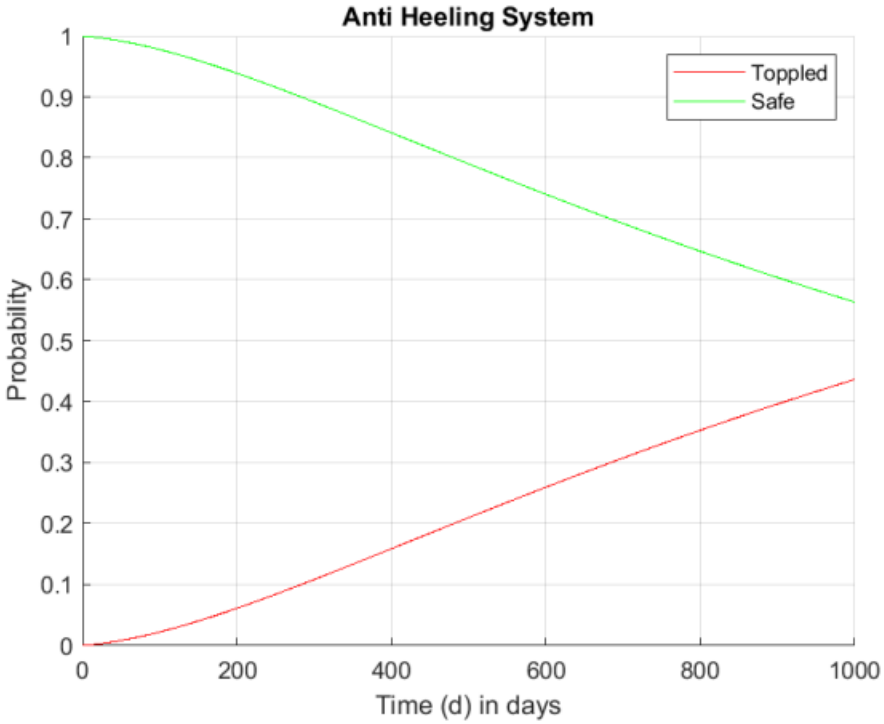

Figura 20 – Markov: Simulação A

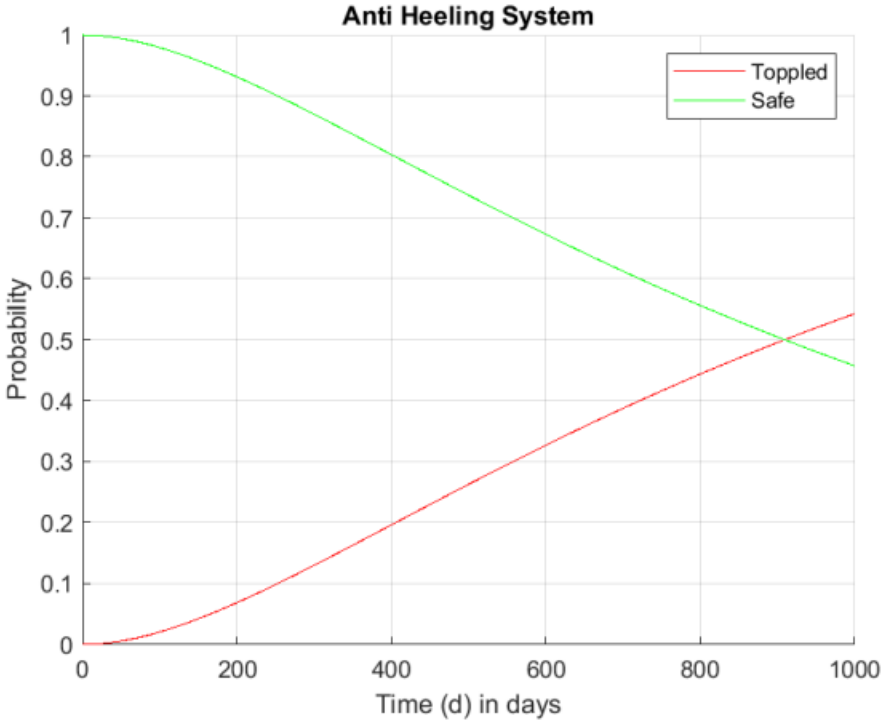

Figura 21 – Markov: Simulação B

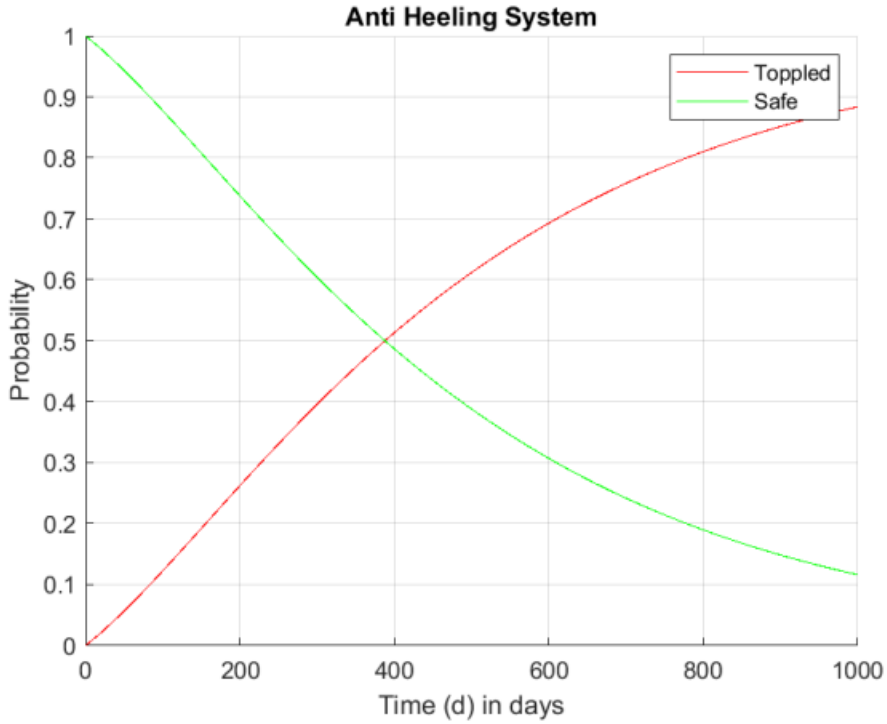

Figura 22 – Markov: Simulação C

Dos resultados obtidos com as simulações, observa-se que nem sempre uma medida de proteção cibernética pode contribuir para melhorar o nível de segurança crítica de um sistema. São destacadas medidas como criptografia e autenticação embora melhorem muito o nível de atendimento as propriedades de *security*, aumentam a latência no sistema que pode ser crucial para a prontidão em um atendimento de medida de *safety,* prejudicando a disponibilidade. Por outro lado, não aplicar tais medidas deixa o sistema extremamente vulnerável a ataques cibernéticos. A melhor decisão para este caso fica na dependência de uma correta estimação de probabilidades para cada evento de *safety* e de *security*.

5.4 APLICAÇÃO BDMP AO ESTUDO DE CASO 1

Na análise do Sistema Anti-Heeling sob a ótica do BDMP, o diagrama pode ser construído em uma abordagem *Top-Down,* semelhante a construção de uma árvore de falha convencional. A Figura 23 ilustra modelo BDMP completo para este estudo de caso. O evento topo é “Navio Tombado”, resultante do evento “Variação de Carga Acima

dos Limites” que inicialmente fica em modo *standby* e pode ser ativada por *triggers* oriundos dos eventos “Opr Maliciosa Tanques” ou do “Sistema AH indisp.”. O evento “Opr Maliciosa Tanques” também é iniciado em modo *standby* e apenas pode ser ativado para o modo *required* se o Sensor, o Computador ou Atuador sofrer um ataque *Spoofing* com sucesso, ou seja, enquanto não houver ataques *spoofing* não haverá possibilidade de ocorrer uma manipulação nos tanques.

O evento “Sistema AH indisp.” ocorre se qualquer um dos componentes, o sensor, o computador ou o atuador sofrerem uma falha acidental ou se tornarem comprometidos por ataques cibernéticos. Esses três componentes têm estrutura que estão representados por uma porta “OR” entre um evento de falha acidental, um evento de “*Jamming* ocorrido” e um evento de “*Spoofing* Ocorrido*”*. Os eventos “*Jamming* Ocorrido” são representados por uma “AND” entre uma contramedida de segurança cibernética denominada “*Anti Jamming*” com *trigger* para “*Jamming*”, ou seja, assume-se que o *Jamming* só possa ocorrer caso a medida *Anti-Jaming* não seja eficiente o bastante ou não seja implementada.

Para a representação do evento “Spoofing ocorrido”, uma primeira barreira corresponde a implementação de um protocolo de autenticação entre os módulos. Enquanto, a autenticação é eficiente então considera-se que não haverá ataques *Spoofing* com sucesso. Porém, quando a autenticação é burlada ainda existe uma segunda contramedida, que se trata de criptografia. Enquanto os dados estão embaralhados um ataque *spoofing* ainda não estará completo. Somente a partir da quebra da criptografia um falso dispositivo poderá ser usado para se passar por um dos módulos do sistema. A partir da ocorrência do *Spoofing* que será possível realizar a manipulação do sistema AH que está representado pelas três setas vermelhas, *triggers,* oriundas do sensor, computador e atuador conforme ilustrado pela Figura 23.

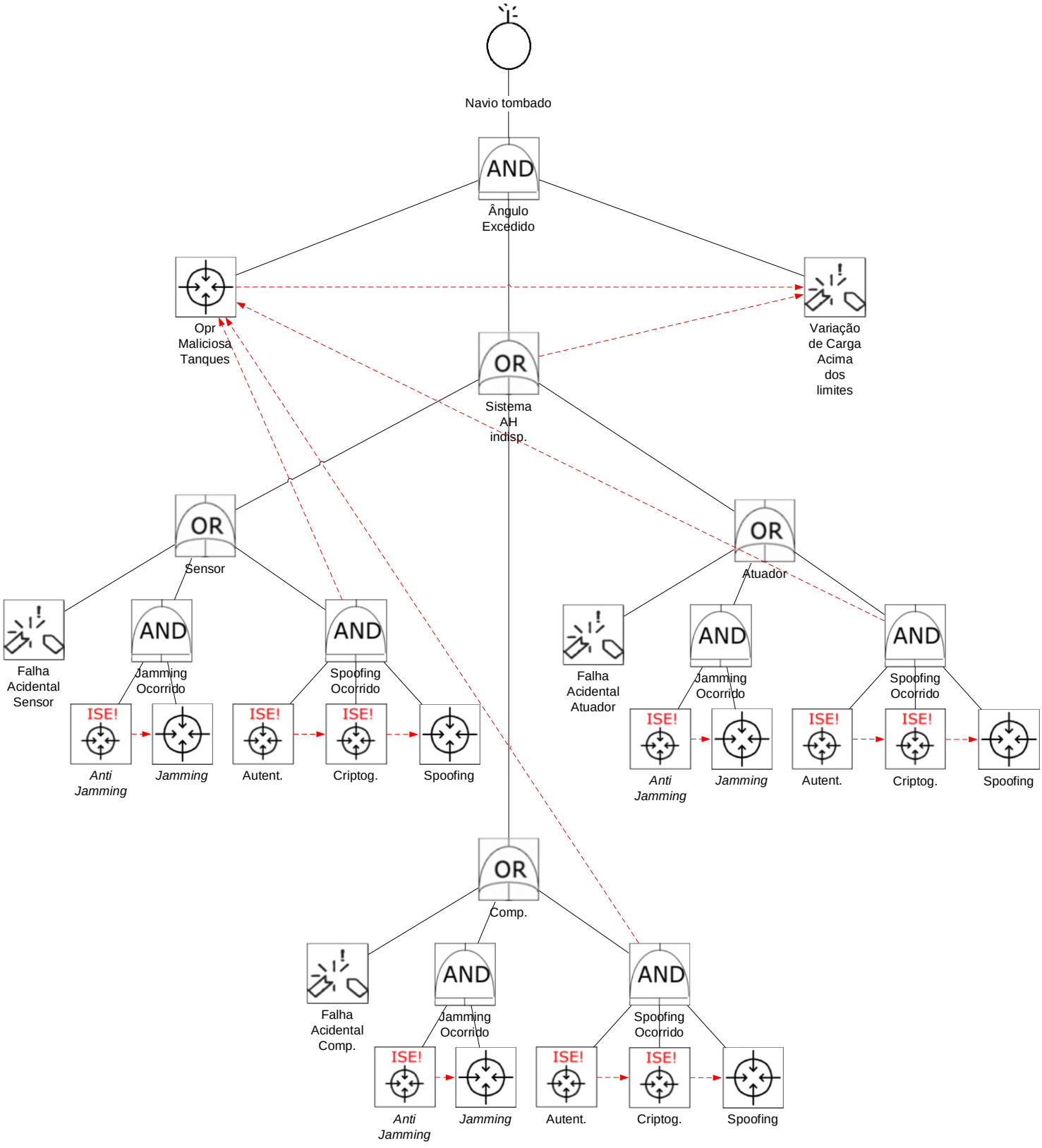

Figura 23 – Sistema *Anti-Heeling* modelado em BDMP

O modelo BDMP foi construído e simulado na ferramenta Risk Spectrum Model *Builder* na versão fornecida para fins acadêmicos pela empresa Risk Spectrum AB (RISK SPECTRUM AB, 2023). A ferramenta permite como entrada, além da árvore BDMP, parâmetros quantitativos para estimativa de ocorrência dos eventos.

Na simulação com o BDMP seguimos a mesma linha mostrada por Cadeias de Markov na seção 5.3, onde são considerados dois cenários, sendo o primeiro sem implementação de medidas de proteção cibernética e são arbitrados os valores de taxas de falhas para os módulos M1, M2, M3 e M4 (λ1=0,0001, λ2=0,0005, λ3=0,0001 e λ4=0,0005). E um segundo cenário considerando medidas de proteção cibernética onde são arbitrados outros valores de taxa de falhas (λ1=0,000001, λ2=0,001, λ3=0,000001 e λ4=0,0005). Neste segundo cenário é considerada que a implementação de autenticação e criptografia reduziriam as chances de ocorrer um ataque cibernético. Logo, λ1 e λ3, que representam as taxas de falhas dos módulos de proteções contra *spoofing* e *jamming,* são arbitrariamente reduzidas (divisão por 100). Assumimos também, que a implementação aumenta a latência de processamento e isso também contribuiria para aumentar as chances do sistema AH falhar de forma não intencional. Logo é arbitrado um incremento para λ2 (multiplicado por 2) em relação ao do cenário anterior. Por último, para o módulo M4, por ser evento externo ao sistema, o valor da taxa de falha permanece igual (λ4=0,0005).

Para o primeiro cenário, sem criptografia e sem autenticação, são ajustados os valores dos parâmetros dos elementos básicos de acordo com a Tabela *7*. São considerados os valores para cada folha base *Spoofing*, Falha Acidental e *Jamming,* iguais entre o sensor, atuador e computador. Para as folhas base de medidas de proteção cibernética, Anti *Jamming*, autenticação e criptografia, são configuradas como não implementadas, ou seja, com ocorrência 100% na referência do BDMP. A folha “Opr Maliciosa Tanques”, que inicialmente fica em standby, também foi configurada com probabilidade de ocorrência 100% quando ativada pelo gatilho.

Ao executar a simulação, obtivemos o valor da probabilidade de ocorrer o evento topo em 0.5772 no tempo 1000 conforme mostrado na Tabela *7*, onde os valores são equivalentes aos obtidos por análises de Markov (MA).

Tabela 7 – BDMP: Cenário A - Entrada de parâmetros AH

Folhas BDMP	**Parâmetros BDMP (taxa de falha)**	**Equivalência em MARKOV**	**Taxas de Falha Cenário (B) MARKOV**
Autenticação e Criptografia (Sensor, Atuador e Comp.)	**1**	M1: Proteção contra Ataque Spoofing	**λ1=0,0001**
Spoofing (Sensor, Atuador e Comp.)	0,0000333		
Falha Acidental (Sensor, Atuador e Comp.)	**0,00001666**	M2: Sistema Anti-Heeling	**λ2=0,0005**
Anti Jamming (Sensor, Atuador e Comp.)	**1**	M3: Proteção contra Ataque Jamming	**λ3=0,0001**
Jamming (Sensor, Atuador e Comp.)	0,0000333		
Variação de carga acima dos limites	0,0005	M4: Variação de Carga Tolerável	λ4=0,0005
Probabilidade do Evento Topo BDMP no tempo 1000	**0.5772**	Probabilidade do Evento Topo MARKOV no tempo 1000	**0,5635**

Para o segundo cenário, com criptografia e autenticação, incluímos os parâmetros de taxa de falha para as folhas de autenticação, criptografia e Anti-Jamming. A implementação das medidas de segurança cibernética por aumentar a latência do sistema incide no aumento das taxas de falhas para as folhas Falha Acidental do sensor, atuador e controlador. Foi considerado a mesma razão da simulação pela Análise de cadeias Markov (duas vezes mais) na seção anterior. Os valores são mostrados na Tabela 8 e são destacadas em negritos os valores alterados.

Tabela 8 – BDMP: Cenário B - Entrada de parâmetros AH

Folhas BDMP	**Parâmetros BDMP (taxa de falha)**	**Equivalência em MARKOV**	**Taxas de Falha Cenário (B) MARKOV**
Autenticação e Criptografia (Sensor, Atuador e Comp.)	**0,01**	M1: Proteção contra Ataque Spoofing	**λ1=0,000001**
Spoofing (Sensor, Atuador e Comp.)	0,0000333		
Falha Acidental (Sensor, Atuador e Comp.)	**0,0000333**	M2: Sistema Anti-Heeling	**λ2=0,001**
Anti Jamming (Sensor, Atuador e Comp.)	**0,01**	M3: Proteção contra Ataque Jamming	**λ3=0,000001**
Jamming (Sensor, Atuador e Comp.)	0,0000333		
Variação de carga acima dos limites	0,0005	M4: Variação de Carga Tolerável	λ4=0,0005
Probabilidade do Evento Topo BDMP no tempo 1000	**0,4222**	Probabilidade do Evento Topo MARKOV no tempo 1000	**0,4574**

Ao executar esta segunda simulação em BDMP, o valor da probabilidade para o evento topo “Navio tombado” ocorrer no instante 1000 unidades é de 0,4222, que é um valor inferior ao cenário A onde simula não haver medidas de proteção cibernética implementadas. As simulações BDMP realizadas no software Risk Spectrum Model Builder foram capazes de identificar o intervalo em que uma medida aplicada para corrigir vulnerabilidade cibernética teve efeito negativo na segurança crítica do sistema.

Os resultados qualitativos obtidos da ferramenta podem ser observados na Tabela 9. Notamos que existem MCS de natureza puramente de *security* que podem permitir o alcance do evento topo.

Tabela 9 – BDMP: MCS do Sistema AH

Tipo	**Nº de Eventos**	**Eventos**
Safety	2	“Falha Acidental Sensor” e “Variação de Carga Acima dos limites”
Safety	2	“Falha Acidental Atuador” e “Variação de Carga Acima dos limites”
Safety	2	“Falha Acidental Computador” e “Variação de Carga Acima dos limites”
Security e Safety	3	“Anti Jamming”, “Jamming” no Sensor e “Variação de Carga Acima dos limites”
Security e Safety	3	“Anti Jamming”, “Jamming” no Atuador e “Variação de Carga Acima dos limites”
Security e Safety	3	“Anti Jamming”, “Jamming” no Computador e “Variação de Carga Acima dos limites”
Security	4	“Autenticação”, “Criptografia”, “Spoofing” no Sensor e “Opr Maliciosa nos Tanques”
Security	4	“Autenticação”, “Criptografia”, “Spoofing” no Atuador e “Opr Maliciosa nos Tanques”
Security	4	“Autenticação”, “Criptografia”, “Spoofing” no Computador e “Opr Maliciosa nos Tanques”

Esses resultados também contribuem para direcionar os analistas a priorizarem a implantação de medidas para ampliar os MCS.

5.5 APLICAÇÃO S-CUBE AO ESTUDO DE CASO 1

A modelagem S-Cube consiste em elaborar a arquitetura do sistema *Anti-Heeling*, relacionando as interações de comunicação entre os elementos e inserindo componentes digitais passivos de ataque cibernético. Uma caraterística dessa modelagem é que não considera a inclusão de elementos externos ao sistema, que para o nosso caso, seria o evento de variação de carga, causado por onda ou por redistribuição de peso interno ao navio. Portanto, para o S-Cube o evento topo seria a indisponibilidade do Sistema AH, e não o “navio tombado”.

O S-cube, assim como o modelo BDMP, foi construído e simulado na ferramenta Risk Spectrum Model *Builder* na versão fornecida para fins acadêmicos pela empresa Risk Spectrum AB (RISK SPECTRUM AB, 2023). Foi instalado uma biblioteca S-Cube KB responsável por carregar as classes dos elementos na linguagem Figaro e S-Cube KB3 HMI responsável por prover a interface gráfica.

A modelagem deste estudo de caso está ilustrada na Figura *24* onde cada elemento é representado de forma decomposta entre a parte física, e parte do *software*, e há uma representação de que estão conectados em uma rede. Portando, o sensor está representado pelos elementos Sensor_SW e Sensor_fisico, o atuador está representado pelo Atuador_SW e Atuador_fisico e o Computador está representado pelo Controlador_SW e Controlador_HW e Estacao_controlador. Ainda é incluído o elemento Rede_Comunicacao interligando os elementos e as setas em azul representando o caminho da informação.

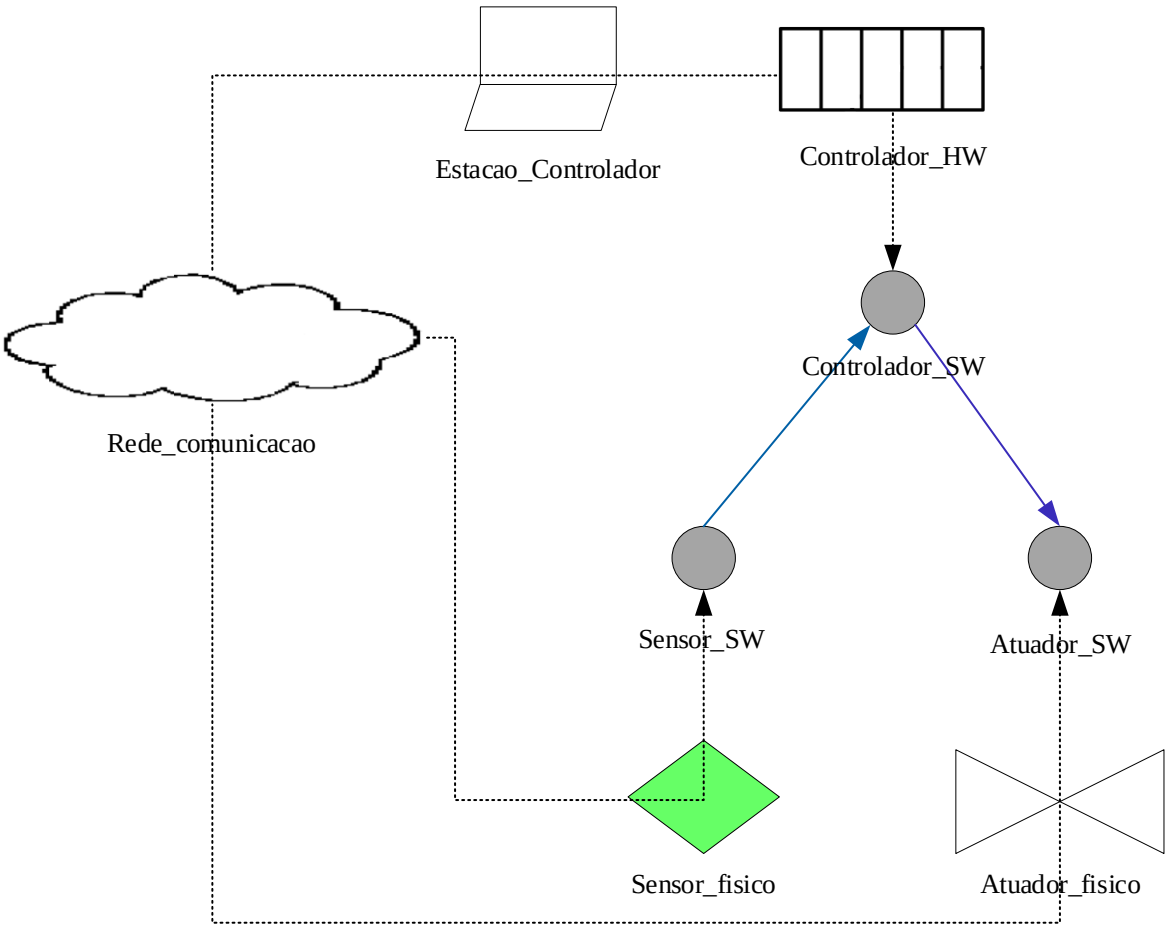

Figura 24 – Sistema *Anti-Heeling* modelado em S-Cube

O S-Cube considera para cada elemento as possibilidades de ataques e de falhas não intencionais que foram catalogadas na base de conhecimento, S-Cube KB. Essa base de conhecimento pode de ser customizada para incluir novos componentes e novos tipos de ataques. Para o nosso estudo de caso, não são alteradas as bases, utiliza-se os componentes nativos. Na Tabela 10 são listados todos os componentes da arquitetura mostrados na Figura *24*, e são relacionados com a classe identificada a base de conhecimento S-Cube e com as ameaças e falhas relacionadas.

Tabela 10 – S-Cube: Classes para o Sistema AH

Componentes do Sistema Anti-Heeling	**Classe no S-Cube**	**Tipo de ameaças catalogadas no S-Cube KB para a classe**
Estacao_Controlador	IT_sys_cpt	**privilege_escalation_attack:** atacante obtém credenciais de administrador **accidental_failure**: falha não intencional; **Access (physical):** atacante em contato com a estação.
Controlador_HW	process_controller_cpt	**send_false_instructions_to_actuator**: um atacante altera instruções no software controlador; **send_no_instructions_to_actuator**: um atacante remove instruções no software controlador.
Controlador_SW	process_controller	**accidental_failure**: falha não intencional; **compromise_comm_link**: falha de comunicação; **ccf_comp**: falhas de causa comum. Ex. Energia.
Sensor_SW	sensor_soft_cpt	**send_false_measures**: um atacante falsifica as medições enviadas pelo sensor; **send_no_measures**: um atacante remove as medidas enviadas pelo sensor.
Sensor_fisico	sensor	**accidental_failure**: falha não intencional; **compromise_comm_link**: falha de comunicação; **ccf_comp**: falhas de causa comum.
Atuador_SW	actuator_soft_cpt	**actuator_does_not_act_properly**: um atacante falsifica ou remove uma ação de controle do atuador.
Atuador_fisico	actuator	**accidental_failure**: falha não intencional; **compromise_comm_link**: falha de comunicação; **ccf_comp**: falhas de causa comum.
Rede_Comunicação	network_zone	**jamming_attack**: negação de serviço da rede; **attacker_scan_network**: um atacante coleta informações de endereços, porta e serviços ativos; **attacker_stablish_connection**: um atacante estabelece uma conexão com algum ativo da rede; **bypass_autentification**: um atacante burla a autenticação.

A base de conhecimento do S-Cube (S-cube KB) disponibilizada não contempla ameaças referentes a quebra de criptografia e não contempla representações de elementos de medidas de proteção cibernética, como criptografia, anti-*jamming* ou

qualquer outra, e ainda conforme mencionado anteriormente não contempla representações de eventos externos. Portanto, devido as estas limitações, não foi possível simular exatamente os mesmos cenários usados em Cadeias de Markov e com o BDMP. Porém, a ferramenta ainda se mostra muito útil por inovar ao permitir como entrada o desenho da arquitetura do sistema e fornecer cenários de ataques e falhas. Os resultados da simulação são mostrados na Tabela *11*.

Tabela 11 – S-Cube: Cenários de ataques e falhas para o Sistema AH

Cenário	Etapas	Descrição
A	1- access (Estacao_Controlador) 2- exploit_Vuln_Priv_Escalator (Estacao_Controlador) 3- attacker_scan_network (Rede_Comunicacao) 4- attacker_stablish_connection (Atuador_SW) 5- Interação com o Atuador: a. send_false_instructions_to_actuator (Estacao_Controlador) b. send_no_instructions_to_actuator (Estacao_Controlador)	O acesso físico à estação do controlador permitiria um atacante manipular as informações enviadas ao atuador.
B	1- access (Rede_Comunicacao) 2- attacker_scan_network (Rede_Comunicacao) 3- bypass_autentification (Rede_Comunicacao) 4- attacker_stablish_connection (Atuador_SW) 5- Controlador sofre *Spoofing*: a. send_false_instructions_to_actuator () b. send_no_instructions_to_actuator ()	O acesso a rede de um dispositivo externo poderia permitir a um atacante se passar pelo controlador e enviar informações falsas ao atuador.
C	1- access (Rede_Comunicacao) 2- attacker_scan_network (Rede_Comunicacao) 3- bypass_autentification (Rede_Comunicacao) 4- attacker_stablish_connection (Controlador_SW) 5- Sensor sofre *Spoofing*: a. send_false_measures () b. send_no_measures ()	O acesso a rede de um dispositivo externo poderia permitir a um atacante se passar pelo sensor e enviar informações falsas ao controlador.
D	1- accidental_failure (Sensor_fisico)	Falhas acidentais no sensor, atuador ou controlador. Uma única falha em qualquer dispositivo já torna o sistema AH indisponível.
E	1- accidental_failure (Estacao_Controlador)	
F	1- accidental_failure (Atuador_fisico)	
G	1- jamming_attack (Rede_Comunicacao)	Um ataque tipo *jamming* realizado com sucesso deixa o sistema AH indisponível.

Foram identificados um total de sete cenários que deixariam o sistema *Anti-Heeling* indisponível. Para os cenários A, B e C da Tabela 11 a ferramenta apontou as

sequências de passos necessários para completar o ataque, isso permite ao analista decidir onde incluir as barreiras de proteção. A ferramenta ainda permite o fornecimento de resultados quantitativos ajustando os valores de taxas de falhas que estão embutidos no código fonte feito em linguagem Figaro (KHAN et al., 2021). Porém, os valores precisam ser estimados por componente ou classes como na seção 2.5. O S-Cube apesar de indicar ataques de segurança cibernética não possui representações de medidas de proteção em sua arquitetura.

5.6 APLICAÇÃO CHASSIS AO ESTUDO DE CASO 1

Na condução desta análise são seguidos os passos da metodologia *Combined Harm Assessment of Safety and Security for Information System* (CHASSIS) mostrados na Figura 8 na seção 4.4. A etapa inicial que corresponde a definição das funções e serviços do sistema AH já está descrita na seção 5.1, cujo objetivo é manter o navio estabilizado controlando os níveis dos tanques de lastro. Então, avançamos para a criação do Diagrama de Caso de Uso ou *Use Case Diagram* (D-UC) inicial em UML (Figura *25*), onde consideramos como atores o sensor, o controlador e o atuador e relacionamos suas atividades.

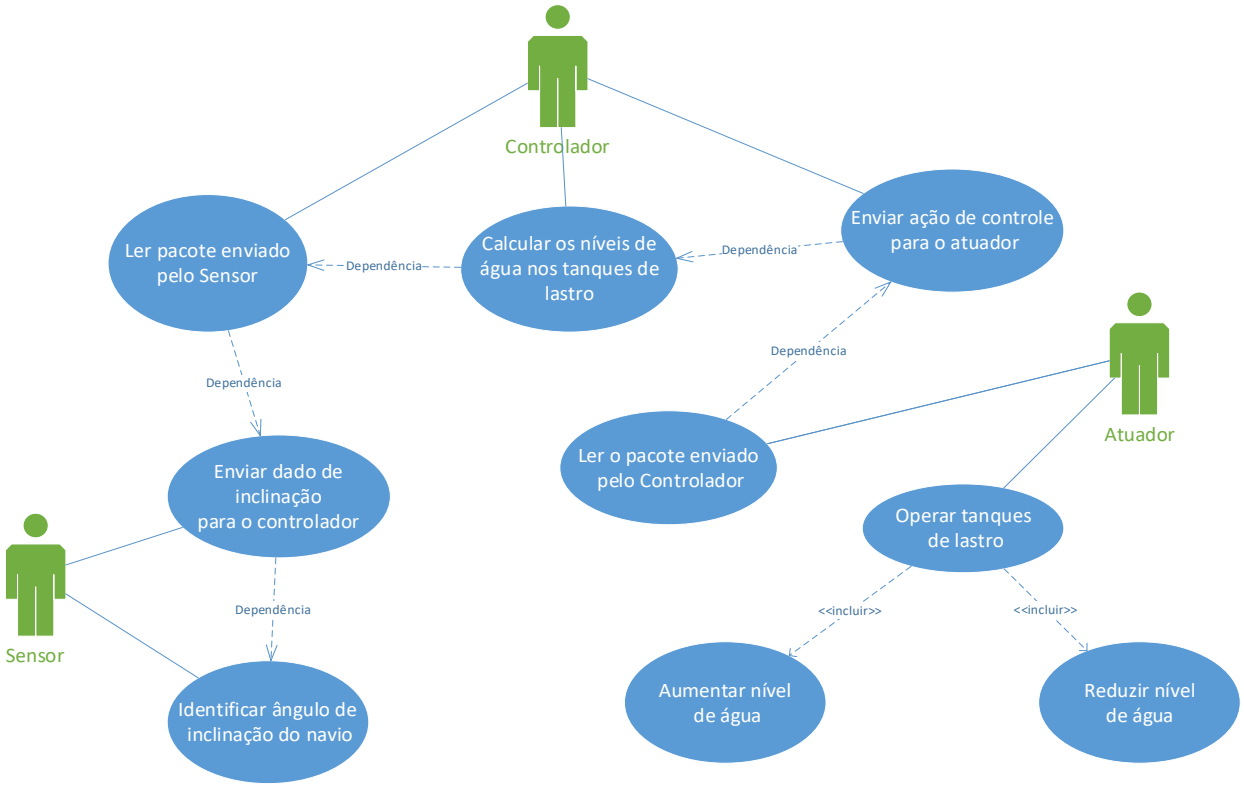

Figura 25 – Diagrama de Caso de Uso (D-UC) inicial para o Sistema AH

A próxima etapa, passo 2 do CHASSIS, consiste em elaborar os casos de uso em modo textual ou *Textual Use Case* (T-UC) então selecionar a ação “Operar tanques de lastro” e incluímos o descritivo T-UC correspondente na Tabela 12.

Tabela 12 – CHASSIS: T-UC do Sistema AH

Campo	**Dado**
Nome	Operar tanques de lastro
Iteração	1
Resumo	Ação de controle responsável pelo ajuste dos níveis de água nos tanques de lastro
Caminho básico	1- Sensor identifica ângulo de Inclinação; 2- Sensor envia o dado para o controlador; 3- Controlador lê o dado; 4- Controlador calcula os níveis de água de cada tanque; 5- Controlador envia as instruções para o atuador; 6- Atuador lê a instrução; e 7- Atuador opera o tanque de lastro a. Aumenta o nível de água do tanque b. Reduz o nível de água do tanque
Caminho alternativo	Inexistente
Gatilhos	Envio de pacotes pelo controlador
Premissas	Os pacotes oriundos do controlador são livres de erros. Os subsistemas responsáveis pela abertura e fechamento de válvula e identificação de nível atingido não são tratados nessa representação.
Pré-condições	O atuador conhece os níveis atuais do tanque
Pós-condições	Inexistente
Regra de Negócio relacionada	Seguir fielmente as instruções do controlador

Para cada atividade deve ser feito uma tabela T-UC, porém, para este estudo de caso apenas uma tabela já pode ser considerada suficiente. Analisando esse resultado podemos observar que a ação Operar Tanques de Lastro não possui caminho alternativo e tem como gatilho a função de envio de pacotes pelo controlador.

Seguindo para o passo 3 do CHASSIS, elaboramos o diagrama de sequência ou *Sequence Diagram* (SD) mostrado na Figura *26*, onde são relacionados os atores e o objeto “tanques de lastro”. O intuito é prover a visualização de toda as etapas que antecede a atividade final do sistema, ajustar o nível dos tanques.

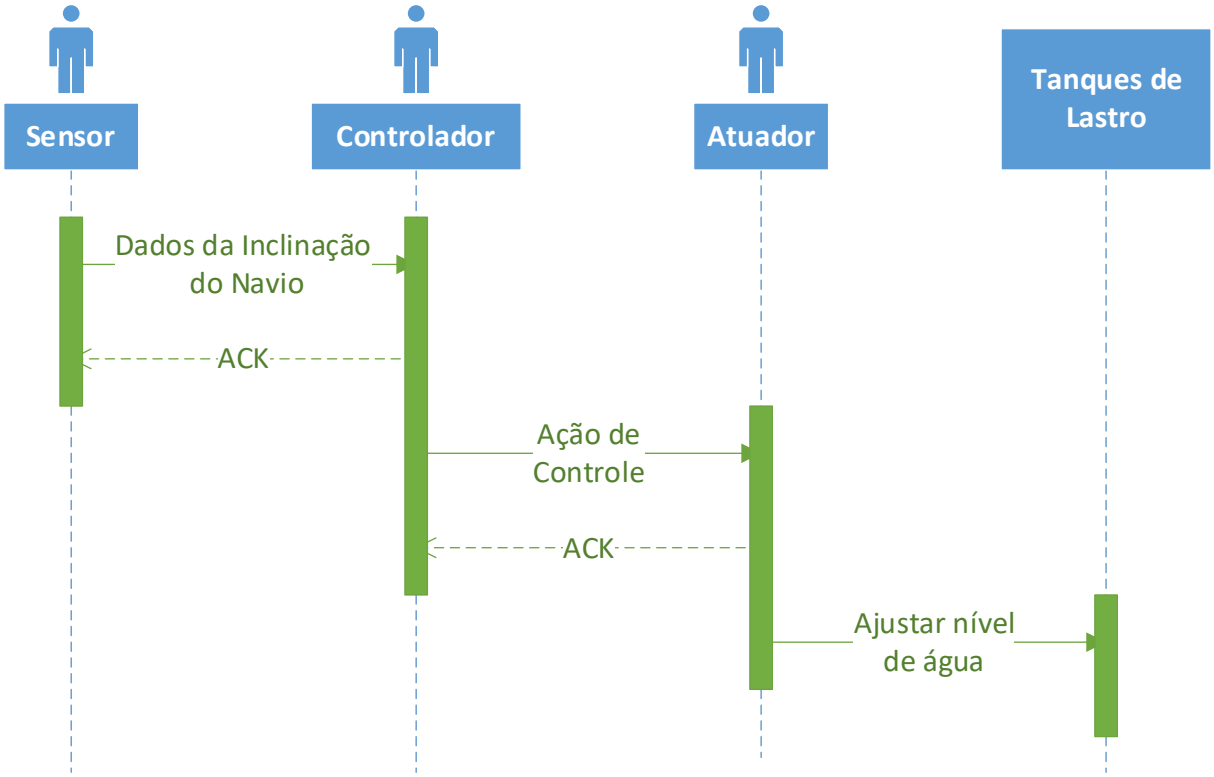

Figura 26 – Diagrama de Sequência (SD) inicial para o Sistema AH

Seguindo para a quarta etapa do CHASSIS, criação dos diagramas de erros de uso ou *Misuse Case Diagram* (D-MUC), o diagrama D-UC da Figura *25* é expandido para o D-MUC da Figura *27*. Os elementos em vermelho representam as ameaças cibernéticas, em cinza representam falhas acidentais e em azul as atividades esperadas para o sistema AH funcionar corretamente.

Adicionamos um ator que representa um atacante com suas possíveis ameaças conectadas as atividades do sistema conforme mostrados por setas com a *tag* “<<ameaçar>>”. Para falhas acidentais adicionamos os atores “sensor defeituoso”, “controlador defeituoso” e “atuador defeituoso” com suas atividades de falhas conectadas às atividades do sistema também por setas com a *tag* “<<ameaçar>>” (Figura *27*).

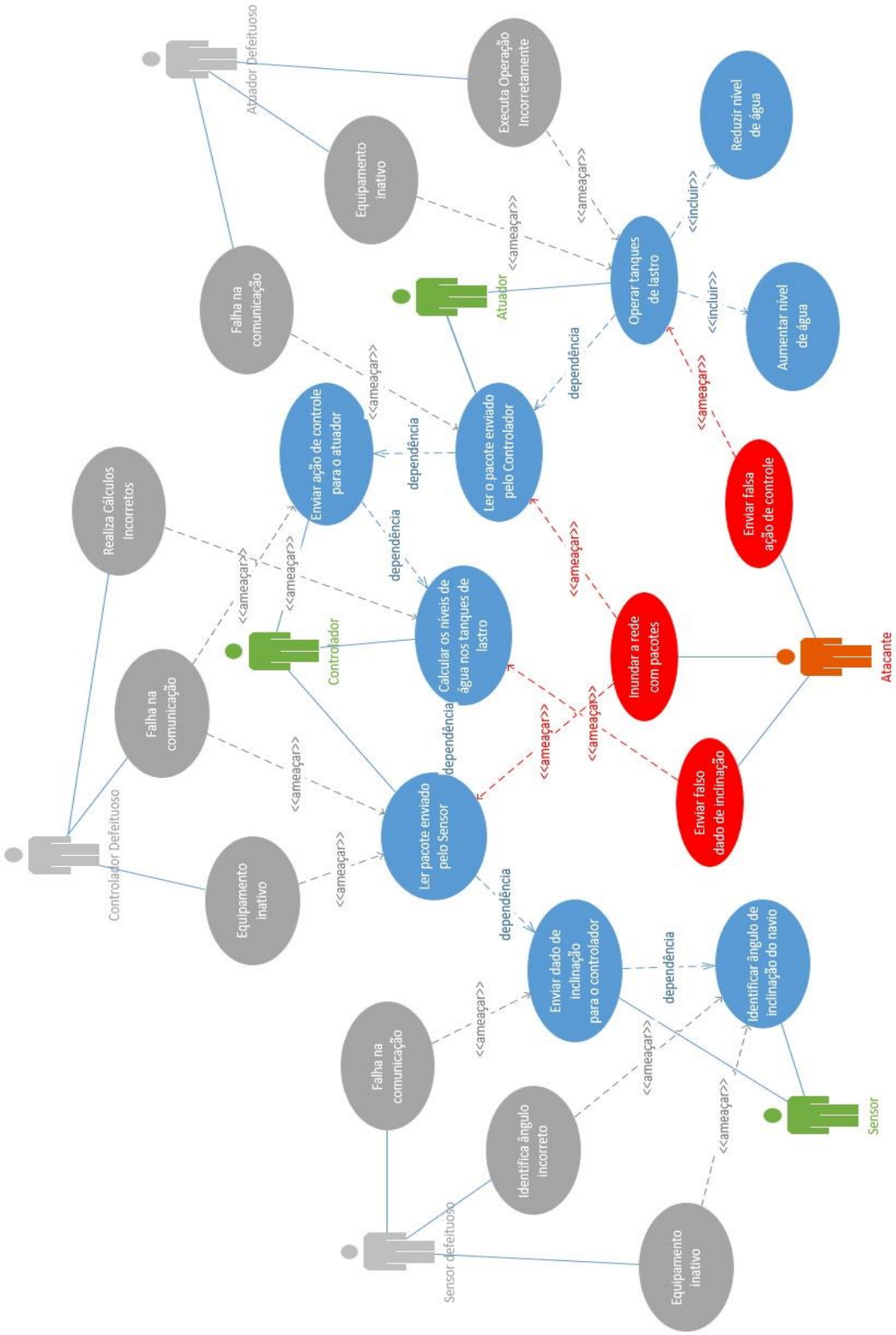

Figura 27 – Diagrama de Erro de Uso (D-MUC) para o Sistema AH

Ainda observando D-MUC (Figura *27*), foram relacionadas três atividades para o atacante. A atividade “Inundar a rede com pacotes” pode comprometer duas outras atividades referentes a capacidade de leitura de dados do controlador (“Ler pacote enviado pelo sensor”) e do atuador (“Ler pacote enviado pelo controlador”). Isso representaria um ataque de *jamming* com sucesso.

A atividade “Enviar falso dado de inclinação” comprometeria a atividade “Calcular os níveis de água nos tanques de lastro” e representa um ataque *spoofing* no sensor. Descrição similar para a atividade “Enviar falsa ação de controle” que compromete “Operar tanques de lastro” e representa um ataque *spoofing* no controlador.

Continuando na Figura *27*, para cada ator do tipo defeituoso, são considerados um evento do tipo “Equipamento inativo” que representa uma indisponibilidade causada por falha acidental. Outro evento do tipo “Falha de Comunicação” representa uma indisponibilidade da rede de comunicação por falha acidental. Um terceiro e último evento que representa um erro na função do equipamento também por falha acidental. Todos esses eventos de falhas estão relacionados com alguma atividade esperada do sistema.

Ainda nesta quarta etapa de elaboração do D-MUC também deve ser gerado uma tabela adaptada com as palavras-guia, *guideword*, de HAZOP (ASPLUND et al., 2019) para cada atividade. Criamos apenas para a função “Ler pacote enviado pelo sensor” mostrado na Tabela 13.

Tabela 13 – CHASSIS: tabela HAZOP para o Sistema AH

Função	Parâmetro	*Guideword*	Causas	Consequência	Dano	Recomendação
Ler pacote enviado pelo sensor	Inclinação do navio	NO	- Falha de Comunicação -Ataque *Jamming*;	Sistema AH indisponível	Navio tombar	-Adicionar redundância de sensor -Incluir medidas anti-jamming
		Other than	-Spoofing no Sensor -Sensor opera com erro			-Implementar autenticação -Implementar verificação do sensor

A função “Ler pacote enviado pelo sensor” extrai o parâmetro com o dado do ângulo de inclinação do navio. Em seguida, aplicamos duas *guideword* do HAZOP, “NO” que representa uma indisponibilidade e “Other than” que representa um valor incoerente com ângulo real do navio. Listamos as possíveis causas que para cada *guideword*, ambas levam a mesma consequência, sistema AH indisponível. O dano relacionado seria o tombamento do navio. Por fim, as recomendações das contramedidas estão na Tabela 13.

A quinta etapa consiste em expandir o D-MUC para o formato textual, *Textual Misuse Case* (T-MUC) que é mostrado na Tabela 14.

Tabela 14 – CHASSIS: T-MUC paro Sistema AH

Campo	**Dado**
Nome	Ler pacote enviado pelo sensor
Iteração	1
Resumo	Responsável por obter o valor do ângulo de inclinação do navio
Caminhos básicos	1. Sensor envia o dado para o controlador; **a. Controlador não recebe pacote;** 2. Controlador lê o dado. **b. Controlador recebe dado incorreto.** **c. Controlador calcula os níveis do tanque incorretamente** **3.** Controlador calcula os níveis do tanque corretamente
Pontos de Mitigação	**a**: Implementar recurso para repetir o envio. **a**: Implementar proteção anti-jamming **a**: Prover redundância de sensor **a**: Prover melhoria na rede de comunicação **b**: Implementar verificação de integridade **b**: Implementar autenticação **c**:Implementar repetição do cálculo
Premissas	-.
Pré-condições	-
Perfil de Usuário Indevido	Atacante, Sensor Defeituoso, Controlador Defeituoso
Riscos e Partes Afetadas	Tombamento do navio: tripulação, passageiros, carga material e meio ambiente.

Na Tabela 14, os erros de uso são representados em vermelho e incluídos no campo de caminhos básicos. No próximo campo, Pontos de Mitigação, são descritas as contramedidas para os erros de uso. Ainda são relacionados os perfis dos usuários indevido, *misusers*, que para o nosso caso corresponde aos atores vinculados diretamente à função. Por fim, são relacionados os riscos e as partes afetadas resultante da execução indevida da atividade.

A sexta etapa do CHASSIS consiste na elaboração dos diagramas sequenciais de falhas ou *Failure Sequence Diagram* (FSD) e diagramas sequenciais de erro de uso ou *Misuse Sequence Diagram* (MUSD). Na Figura 28 temos a representação de um FSD para um perigo, *hazard*, da função Ler Pacote enviado pelo sensor. Na Figura 29 temos um MUSD para a ameaça de interceptação de pacotes e reprodução com dados adulterados.

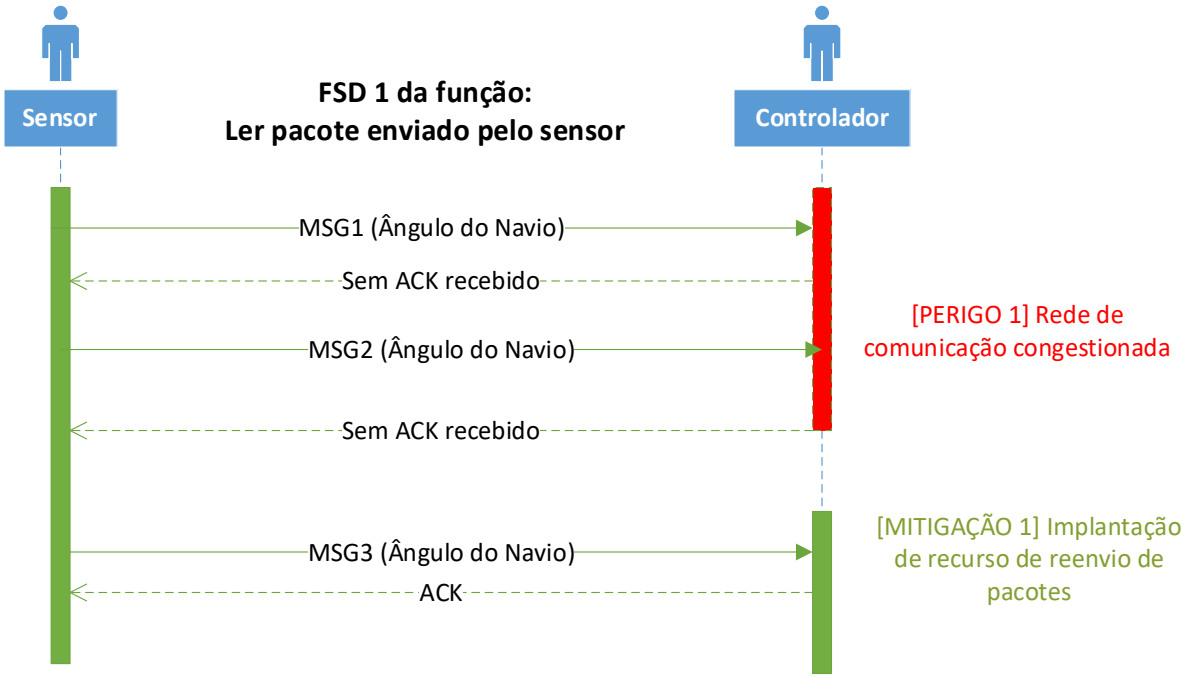

Figura 28 – Um Diagrama Sequencial de Falha (FSD) do Sistema AH

Para o FSD representado, é considerado a hipótese do perigo de rede congestionada que pode ser mitigada por uma implementação de reenvio de pacotes caso não recebe uma confirmação, *Ack*, do destinatário (Figura *28*).

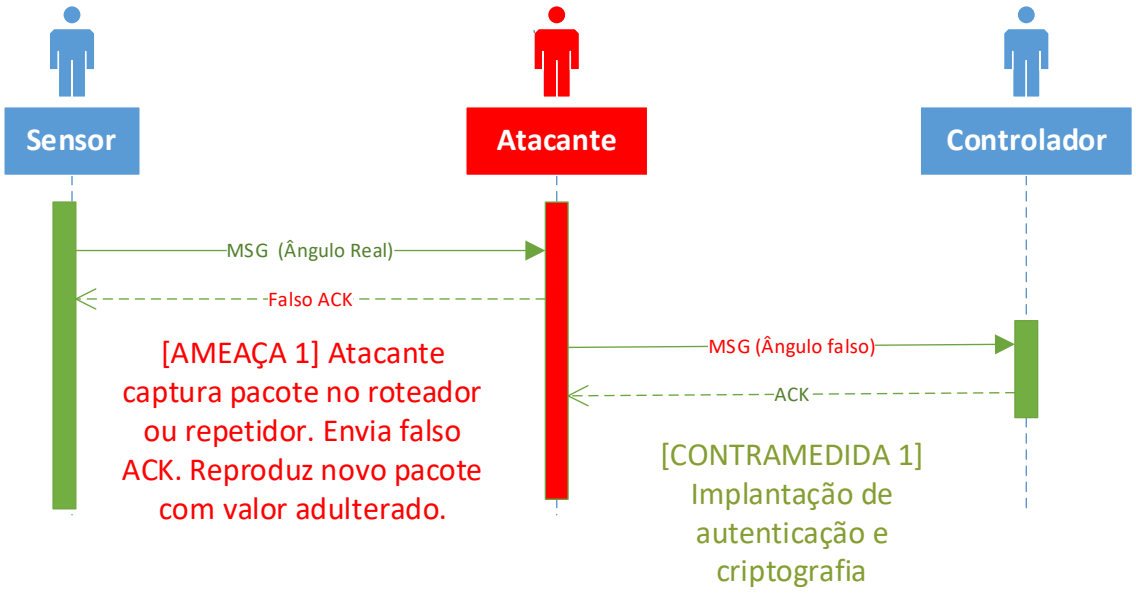

Figura 29 – Um Diagrama Sequencial de Erros de Uso (MUSD) do Sistema AH

Para o MUSD (Figura *29*), é representado a ameaça de um atacante interceptar os pacotes oriundo do sensor, alterar o dado e enviar um valor falso para o ângulo de inclinação. Uma contramedida pode ser a implantação de procedimentos de autenticação e criptografia nas trocas de pacotes.

As saídas desta sexta etapa (FSD e MUSD) podem realimentar a quinta etapa (T-MUC) e retornar à primeira etapa (D-UC) conforme visualizado no Figura *8*. A etapa 7 corresponde a extração de tabelas HAZOP no mesmo formato da Tabela 13. A última etapa consiste na especificação dos requisitos de *safety*/*security* identificados ao longo do ciclo.

Tabela 15 – CHASSIS: Requisitos especificados para o sistema AH

Nº	Requisitos
1	O sistema deve ter protocolos de autenticação.
2	O sistema dever ter redundância de sensor, controlador e atuador.
3	O sistema deve detectar quando um componente estiver defeituoso.
4	O sistema deve prover mecanismo *anti-jamming.*
5	A rede de comunicação deve ter largura de banda suficiente para a comunicação entre os módulos.
6	O sistema deve implementar criptografia desde que não consuma largura de banda excessiva.

A abordagem CHASSIS, portanto, tem como saída os requisitos de forma refinada já que em sua proposta é estimulado a repetição do ciclo. No entendimento dos autores

do método (RASPOTNIG; KARPATI; OPDAHL, 2018), o atendimento a esses requisitos aumentaria o nível de segurança da aplicação, tornando apropriada para execução em ambiente crítico.

5.7 APLICAÇÃO GTST-MLD AO ESTUDO DE CASO 1

Na condução da análise do sistema Anti-Heeling seguindo os preceitos do método *Goal Tree Succes Tree Master Logic Diagram* (GTST-MLD), seguimos primeiro elaborando árvore *Goal Tree* que representa a função objetivo decomposta em subfunções. Definimos a função topo como “Prevenir o tombamento do navio” e pela recomendação do método, vamos questionando o “como" podemos atende-la. Uma das subfunções então deve ser “Evitar a manipulação no controle dos tanques de lastros”. Caso essa subfunção seja comprometida, então a função topo também será comprometida, logo, esta subfunção deve estar conectada a um porta AND. Outra subfunção para atender a função topo é que o Sistema AH continue com a capacidade de “Controlar os níveis de água nos tanques de lastro”, mais ainda se esta subfunção for comprometida ainda seria necessário uma variação de carga, como ondas marítimas intensas ou redistribuição de peso dentro do navio, logo teremos uma conexão de porta OR entre as subfunções “Controlar os níveis de água nos tanques de lastro” e “Evitar Exposição a Variação de Carga”. Na Figura 30, a árvore GT pode ser visualizada na disposição vertical.

A segunda etapa consiste na criação da árvore *Success Tree* (ST) que será disposta horizontalmente e deverá conter os sistemas e componentes necessários para atender as funções da árvore GT. Representamos o Sistema Anti-Heeling composto de uma porta AND entre os componentes Sensor, Controlador, Atuador e Tanques. No mesmo nível incluímos o sistema de Rede de Comunicação e Gestão Náutica. Este último, apesar de não ter sido representado nos modelos anteriores, definimos sua inclusão para termos uma representação mais rica do diagrama e assim contribuir para um entendimento melhor da ferramenta (Figura 30).

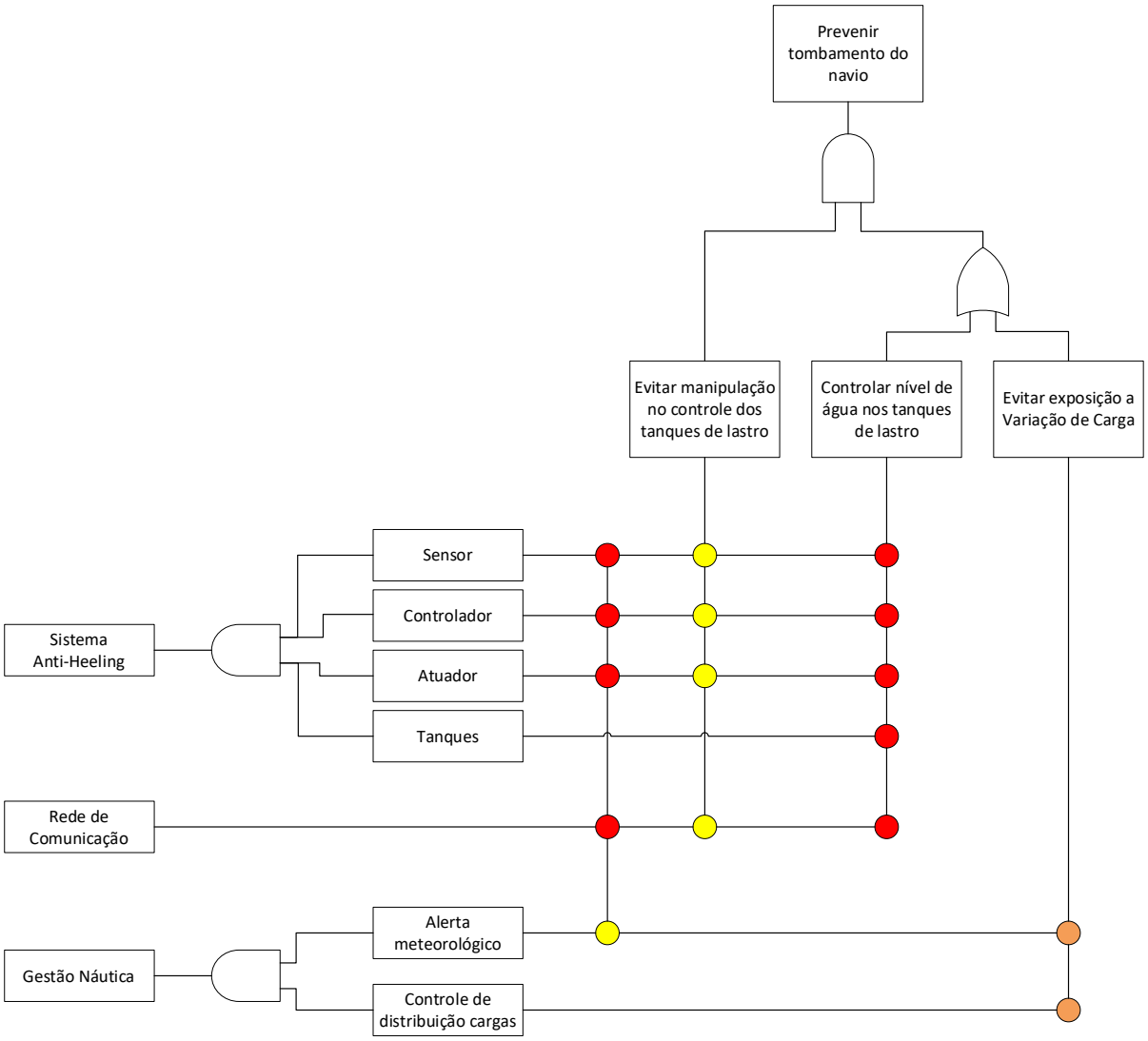

Figura 30 – GTST-MLD do Sistema AH

A última etapa para conclusão do diagrama é realizar as conexões *Master Logic Diagram* (MLD) entre as duas árvores, GT e ST, e entre componentes que fazer parte de outro sistema. O método permite dimensionar três níveis de conexões: alto (vermelho), médio (laranja) e baixo (amarelo). Esses níveis podem ser usados para os cálculos quantitativos e ainda podem ser expandidos para mais faixas (DI MAIO; MASCHERONA; ZIO, 2020).

Para o componente “**sensor**” observamos na Figura 30 que está:

a. Fortemente acoplado ao sistema “rede de comunicação”: significa que uma falha na rede de comunicação afetará a missão do sensor.

b. Fracamente acoplado a função “Evitar Manipulação do Controle dos Tanques de Lastro”: significa que mesmo que o sensor venha ser comprometido ainda não poderia ser suficiente para comprometer a função associada. Uma vez que o sensor envia o dado para o controlador, uma implementação de autenticação no lado do controlador poderia combater um *spoofing* sucedido no sensor.
c. Fortemente acoplado a função “Controlar Nível de Água nos Tanques”: como o sistema possui apenas um sensor então seu comprometimento reflete diretamente o atendimento a função.

Os componentes controlador, atuador e o sistema rede de comunicação têm explicações equivalente aos acoplamentos do sensor. O componente tanque está fortemente acoplado somente a função “Controlar Nível de Água nos Tanques”.

A função “Evitar Exposição a Variação de Carga” representa um evento externo, onda marítima ou movimentação de carga interna, fora do controle do Sistema AH. Para não deixar a função no vazio adicionamos um outro sistema “Gestão Náutica” que corresponde a uma porta OR entre os subsistemas “Alerta Meteorológico” e “Controle de Distribuição de Cargas”. Ambos subsistemas relacionamos com o nível médio de acoplamento por considerar que são necessários para o atendimento da função, porém assumimos que possuem execuções manuais. Para o subsistema de “Alerta Meteorológico” também o acoplamos fracamente à “Rede de Comunicação” por assumirmos que haverá outros meios redundantes como rádio e telefone para receber os alertas.

Ainda podemos representar para o Sistema “Rede de Comunicação” e para os componentes “Sensor”, “Atuador” e “Controlador” em um outro diagrama mostrado na Figura 31 relacionando com os tipos de ameaças cibernéticas. Adaptamos para aplicar o modelo STRIDE conforme Tabela 1. Podemos observar que o sensor, atuador e controlador estão sujeitos a todos os cincos primeiros tipos de ameaças do modelo STRIDE e somente o Controlador pode estar sujeito a ameaça de elevação de privilégio por presumirmos que está instalado em uma estação de trabalho completa com sistema operacional e com perfis de usuários. O sensor e atuador são dispositivos embarcados.

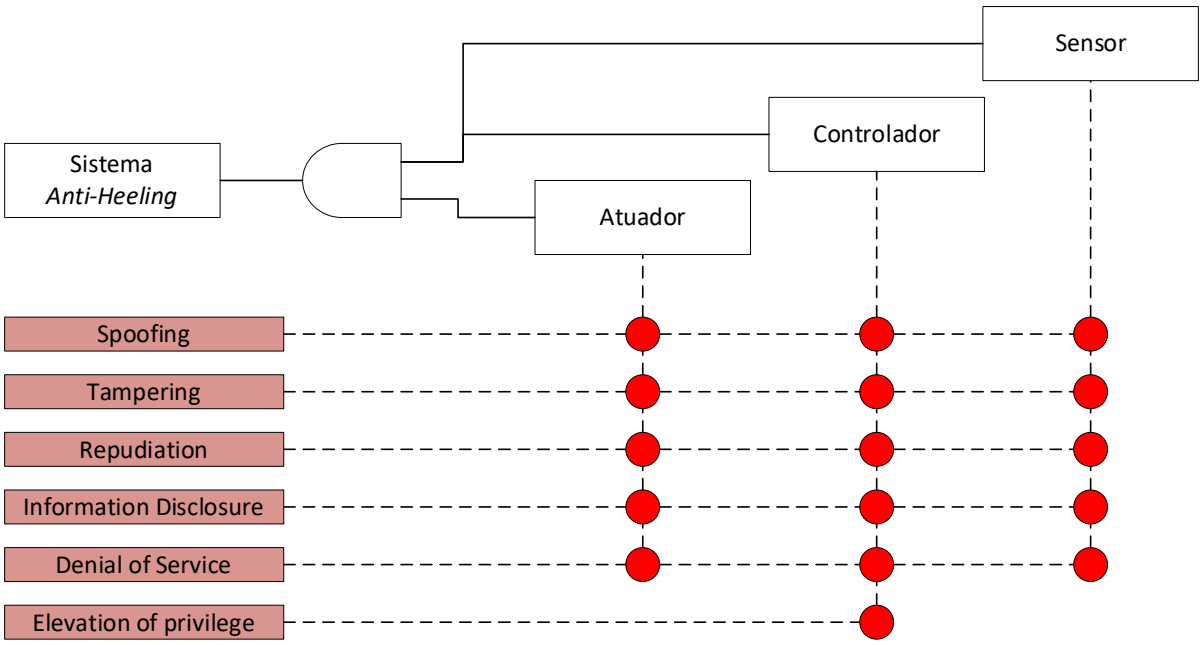

Figura 31 – GTST-MLD: Ameaças cibernéticas ao Sistema AH

Ambos diagramas contribuem para uma visualização ampla da relação componentes, funções e ameaças cibernéticas. Permite ao analista identificar alta dependência de uma função a um componente e adicionar alternativas. Permite também identificar quando uma função não é plenamente atendida para que seja providenciado um sistema correspondente. Sobre o ponto de vista de segurança cibernética é possível rastrear o caminho de um ameaça cibernética até a função ou sistema afetado.

Para o GTST-MLD não foi disponibilizado uma ferramenta própria, os diagramas podem ser montados em qualquer ferramenta MLD. Sobre resultados quantitativos a ferramenta deixa aberto a possibilidade de associar a modelos matemáticos como probabilidade de eventos dependentes e independentes, simulação de modelos de Monte Carlo e redes bayesianas (GEORGE; RENJITH, 2021).

Como resultado qualitativo da análise para o Sistema *Anti-Heeling* pelos diagramas mostrados na Figura 30 e na Figura 31, podemos concluir que é necessário prover redundância do sensor, controlador e atuador, implantar mecanismo de autenticação, prover meios alternativos a indisponibilidade da rede de comunicação como acionamento manual. Por último, a inclusão da função “Evitar a exposição a variação de carga” contribuiu para a rápida visualização da necessidade de adicionar um

sistema “Gestão Náutica”. Além ainda, de que a subfunção “Alerta Meteorológico” está acoplada a rede de comunicação.

5.8 APLICAÇÃO STPA-SEC AO ESTUDO DE CASO 1

Para esta análise são seguidos os processos do *System-Theoretic Process Analysis* (STPA) estão ilustradas na Figura 10 e adicionando as etapas referente a cibersegurança da versão STPA for *Security (*STPA-Sec*)*. Uma grande vantagem deste método é a existência de diversas ferramentas como STAMP Workbench e SafetyHAT de livre acesso e Visual PRO que é comercial com possibilidade de suporte especializado (MIT PSASS GROUP, 2023).

O primeiro passo consiste em identificar as perdas da missão do sistema e levantamos inicialmente um total de quatro perdas conforme observado na Tabela 16. As perdas L-1, L-2 e L-3 associadas a acidentes e a perda L-4 está associada a satisfação de clientes. As perdas não são consideradas hierárquicas para o sistema.

Tabela 16 – STPA-Sec: Perdas levantadas para o sistema AH

Identificação	**Perdas (ou *Losses*)**	**Descrição**
L-1	Violação da Integridade física de pessoas pelo tombamento do navio.	Associada a possibilidade de ocorrência de acidentes que resultem em pessoas feridas ou óbitos.
L-2	Dano a meio ambiente pelo tombamento do navio.	Associada a possibilidade de agressão ao meio ambiente como derramamento de óleo ou outras cargas poluentes.
L-3	Prejuízo Material pelo tombamento do navio.	Associada a possibilidade de perdas de cargas ou danos aos equipamentos.
L-4	Desconforto dos passageiros	Associada a possibilidade de passageiros vivenciarem uma má experiência devido a movimentação do navio nas ondas marítimas

Partimos para a identificação dos perigos e ameaças que são mostrados na Tabela *17*. Observamos que os perigos H-1, H2, H3 e H4 foram associados a todas as perdas e o H-5 apenas deixou de ser associada a perda L-4.

Tabela 17 – STPA-Sec: Perigos e Ameaças levantadas para o sistema AH

Identificação	**Perigos ou ameaças (*Hazards* e *Threats*)**	**Perdas associadas**	**Descrição**
H-1	Sistema AH Indisponível	L-1, L-2, L-3 e L-4	Representa um perigo associado quando o sistema de AH está inoperante.
H-2	Sistema AH Opera Incorretamente	L-1, L-2, L-3 e L-4	Representa um perigo associado quando o sistema de AH opera com erros.
H-3	Sistema AH Manipulado	L-1, L-2, L-3 e L-4	Representa uma ameaça quando o sistema de AH está sob controle de atacantes cibernéticos.
H-4	Navegação em mar agitado	L-1, L-2, L-3 e L-4	Representa um perigo quando o navio estar exposto às ondas marítimas intensas.
H-5	Ingerência na operação de cargas	L-1, L-2, L-3	Representa um perigo para um navio cargueiro nas operações de carga e descarga de forma não coordenada.

A próxima atividade é a definição das restrições do sistema, ou *System Constraints* (SC) que é mostrado na Tabela 18 e estão relacionados com Perigos ou Ameaças.

Tabela 18 – STPA-Sec: Restrições do Sistema AH

Identificação	**Restrições do Sistema (*System Constraints*)**	**Perigos ou Ameaças associadas**
SC-1	Sistema AH deve operar continuamente durante a navegação	H-1, H-2, H3
SC-2	Sistema AH deve ser testado periodicamente	H1, H-2
SC-3	Sistema AH deve implementar proteções cibernéticas	H-1, H-2, H3
SC-4	O navio deve receber alertas meteorológicos	H-4
SC-5	A movimentação de carga deve ser feita com o Sistema AH ativo	H-5

Partindo para o processo 2 do STPA-Sec, desenhamos a estrutura de controle para o Sistema Anti-Heeling e pode ser visualizada na Figura 32. A estrutura de controle deve ser desenhada visando atender as SC identificadas na etapa anterior.

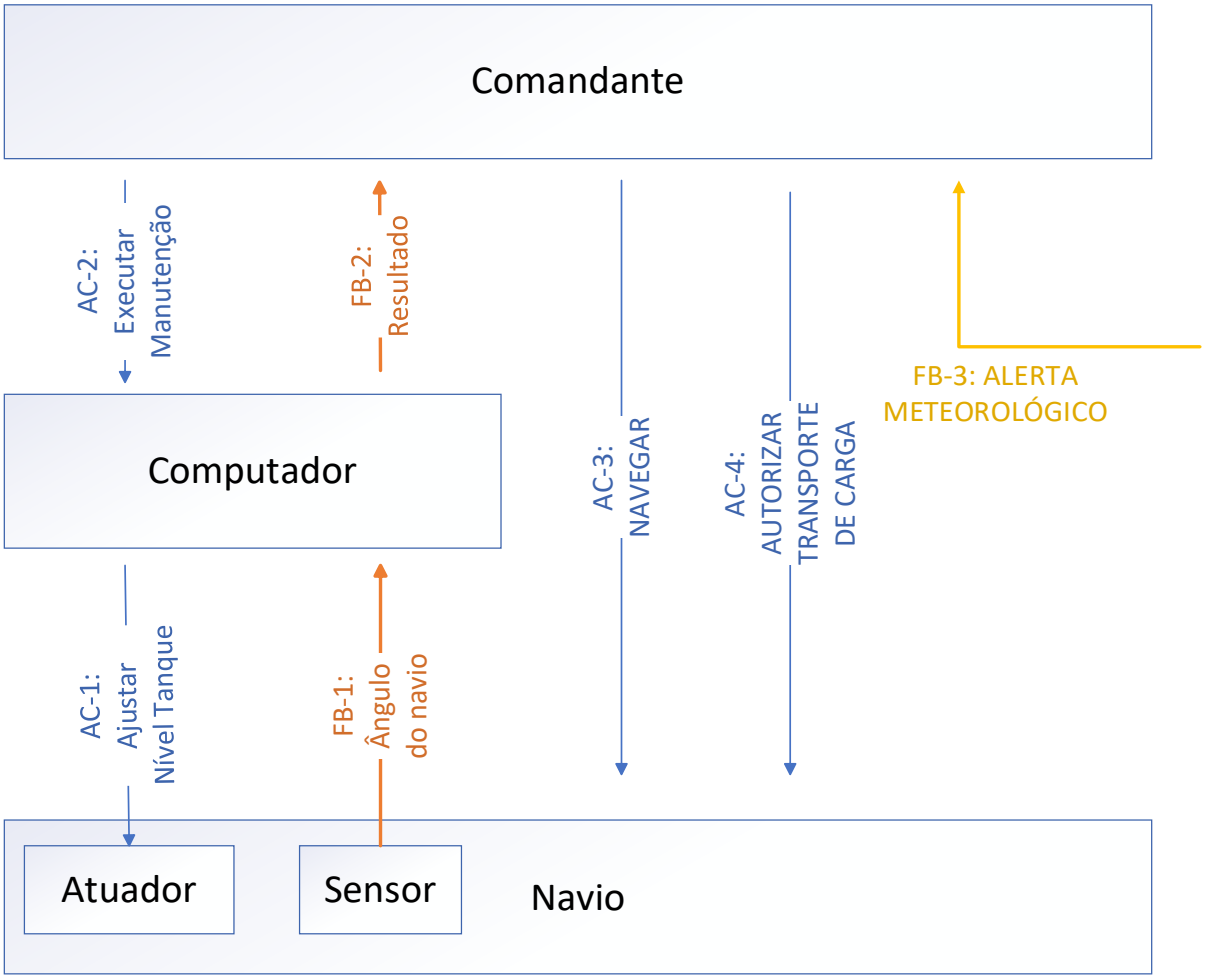

Figura 32 – STPA-Sec: Estrutura de Controle do Sistema AH

Observando a Figura 32, adicionamos o módulo “Comandante” que se trata de outro controlador que foi necessária para atender as SC-2, SC-4 e SC-5 já que o Sistema Anti-Heeling por si só não contempla funções para estas restrições. Ainda devemos estabelecer as responsabilidades de cada controlador da estrutura, no caso, o Computador de Bordo do sistema Anti-Heeling e o Comandante (Tabela 19).

Tabela 19 – STPA-Sec: Responsabilidades dos Controladores AH

ID	Controlador	Responsabilidade	Restrição
R-1	Computador	Ajustar os níveis dos tanques quando o navio sofrer inclinações.	SC-1, SC-3
R-2		Identificar a legitimidade das informações trocadas entre os módulos.	SC-3
R-3	Comandante	Programar manutenções periódicas para o Sistema AH.	SC-2
R-4		Desviar a rota do navio de zonas de mal tempo.	SC-4
R-5		Certificar que o Sistema AH está ativo antes de iniciar operações de carga e descarga	SC-5

Partimos para o processo 3 que consiste em Identificar Ações de Controles Inseguras (UCA). Cada ação de controle deve ser analisada nas categorias de comportamento de UCA. Será analisado a AC-1, “Ajustar Nível Tanque” mostrado na Tabela 20.

Tabela 20 – STPA-Sec: Ações de Controle Inseguras AH

Ação de Controle / Comportamento	**AC-1: Ajustar Nível do Tanque**	**Perigo**
Não executada quando requisitada	UCA-1: Sistema AH não ajusta tanque de lastro quando o navio está inclinado	H-1
Executada quando não requisitada	UCA-2: Sistema AH altera nível do tanque de lastro mesmo com o navio balanceado	H-2 H-3
Executada Muito Cedo, Muito Tarde ou Fora de Ordem	UCA-3: Sistema AH opera os tanques com muito atraso	H-2
Interrompida muito cedo, aplicada por muito tempo	UCA-4: Sistema AH não atinge o nível de água solicitado. UCA-5: Sistema AH ultrapassa o nível de água solicitado	H-2 H-3
Executada incorretamente	UCA-6: Sistema AH não ajusta o nível de água para o valor correto.	H-2, H-3

Foram levantadas 6 Ações de Controle Inseguras e estão relacionadas um tipo de Perigo e para cada UCA deve ser estabelecido as *Controller Constraints* (CC) ou Restrições do Controlador que visa especificar a prevenção as UCA (LEVESON; THOMAS, 2018) mostrados na Tabela 21.

Tabela 21 – STPA-Sec: Restrições do Controlador AH

UCA do AC-1	**Restrição do Controlador (CC)**
UCA-1	CC-1: O computador de bordo deve coletar a inclinação real e garantir que o atuador ajuste os níveis dos tanques
UCA-2	CC-2: O computador de bordo deve certificar que está recebendo o valor de inclinação do sensor legítimo.
UCA-3	CC-3: O computador de bordo deve operar em rede de comunicação de baixa latência
UCA-4	CC-4: O computador de bordo reenviar a instrução para o atuador até que os tanques estejam plenamente ajustados
UCA-5	
UCA-6	

Partimos então para o último processo do STPA-Sec, identificação dos cenários de perda, que consiste no levamento dos fatores relacionados na ocorrência das UCA. Nesta fase também pode ser aplicado o *Wargaming* (DE SOUZA et al., 2020) (YOUNG; LEVESON, 2014) que terá como objetivo uma disputa entre duas equipes, uma

realizando os ataques cibernéticos e a outra implementando as defesas. Os resultados são incluídos na tabela de fatores relacionados as UCA.

Foram realizados o levantamento de alguns fatores e cenários de perdas para UCA-1, “Sistema AH não ajusta tanque de lastro quando o navio está inclinado” (Tabela 22).

Tabela 22 – STPA-Sec: Cenários de Perda AH

Fatores para a UCA-1	**Cenários de perda para a UCA-1**	**Perigo**
Falha física no controlador (aquecimento)	S-1: a exposição do controlador a faixas de temperaturas não toleráveis o torna inoperante.	H-1, H-2
Falha de energia	S-2: Instalações elétricas inadequadas comprometem o funcionamento do controlador.	H-1, H-2
Erros do algoritmo	S-3: Software não testado o suficiente pode levar a erros de processamento e disponibilidade.	H-1, H-2, H-3
Erro de comunicação do controlador	S-4: excesso de tráfego na rede impede o recebimento e envio de pacotes	H-1, H-2, H-3
Controlador não recebe feedback	S-5: o controlador fica impossibilitado de calcular novo nível para os tanques.	H-1, H-2, H-3
Controlador recebe feedback incorreto	S-6: o controlador calcula novo nível para os tanques equivocadamente.	H-2, H-3

Através dos cenários de perdas identificados mostrados na Tabela 22, novos requisitos podem ser estabelecidos e novas estruturas de controle elaboradas. Podemos tomar como exemplo o cenário S-1, no qual considera a possibilidade do controlador ficar exposto às altas temperaturas, e isso induz a criação de um requisito que pode ser “SC: os equipamentos devem estar localizados em ambiente cuja temperatura não ultrapasse os valores especificados pelo fabricante”. Com essa nova SC, pode-se gerar uma nova estrutura de controle como monitoramento de temperatura ambiente e emissão de alertas.

Para os cenários de segurança cibernética, podemos mencionar S-6 onde considera um recebimento incorreto do valor e isto pode estar atrelado tanto a uma má operação do sensor como também a um ataque *spoofing*. Isto pode gerar outro requisito e uma nova estrutura de controle para proteção contra este tipo de ataque cibernético. No entanto, tais medidas que seriam autenticação e criptografia aumentaria a latência nas comunicações contribuindo para ocorrência da perda S-4 que está associada a alta latência.

6 ESTUDO DE CASO 2 – SISTEMA DE TORRE DE CONTROLE DE AERÓDROMO (TWR)

O estudo de caso analisado neste capítulo representa um subconjunto de funcionalidades de um sistema de controle de tráfego aéreo em torres de controle de aeroportos. O sistema apresentado é responsável pelas autorizações de pouso e decolagem de aeronaves. A análise de segurança crítica e cibernética será feita sob as óticas das metodologias BDMP, S-Cube, CHASSIS, GTST-MLD e STPA-Sec. As análises levam em consideração as possibilidades de falhas dos componentes e de ataques cibernéticos.

6.1 DESCRIÇÃO DO SISTEMA TWR

A finalidade de um sistema de torre de controle de aeródromo, também conhecido como *Aerodrome Control Tower* (TWR), é permitir o gerenciamento do tráfego aéreo de um aeroporto nas operações de pouso, decolagem, manobras em solo e ainda controle de voos no circuito de tráfego visual do aeródromo (DECEA, 2023).

Para este estudo de caso são consideradas somente as funções de pouso e decolagem. O Controlador de tráfego aéreo, conhecido como *Air Traffic Controller* (ATCO) da Torre, é o responsável pela autorização de pousos e decolagens. No entanto, o controlador deve observar os seguintes requisitos antes de conceder uma autorização:

a) A condição meteorológica deve estar favorável.
b) O espaço aéreo deve estar livre de outras aeronaves ou qualquer objeto voador.
c) O pátio deve estar livre de obstáculos terrestres.
d) Não deve haver outra autorização em curso.

Após o piloto receber a autorização, a aeronave e a tripulação devem ser preparadas para o pouso ou decolagem.

Consideraremos então que, não ocorrerão acidentes desde que as autorizações estejam de acordo com os requisitos e os pilotos preparem corretamente suas aeronaves para o pouso ou decolagem.

A Figura 33 ilustra o diagrama que representa os elementos e interações para o estudo de caso. O elemento Torre corresponde ao Controlador de Tráfego e o Sistema TWR. As entradas têm origens no Fiscal de Pátio via Rádio, do Radar e do Sistema Externo via rede de dados. O Controlador interage com o Piloto via Rádio e também por enlace de dados. O Piloto prepara a aeronave operando os sistemas ciber físicos que ajustam os *flaps, slats* e reversos e atualizam os estados na console de monitoramento.

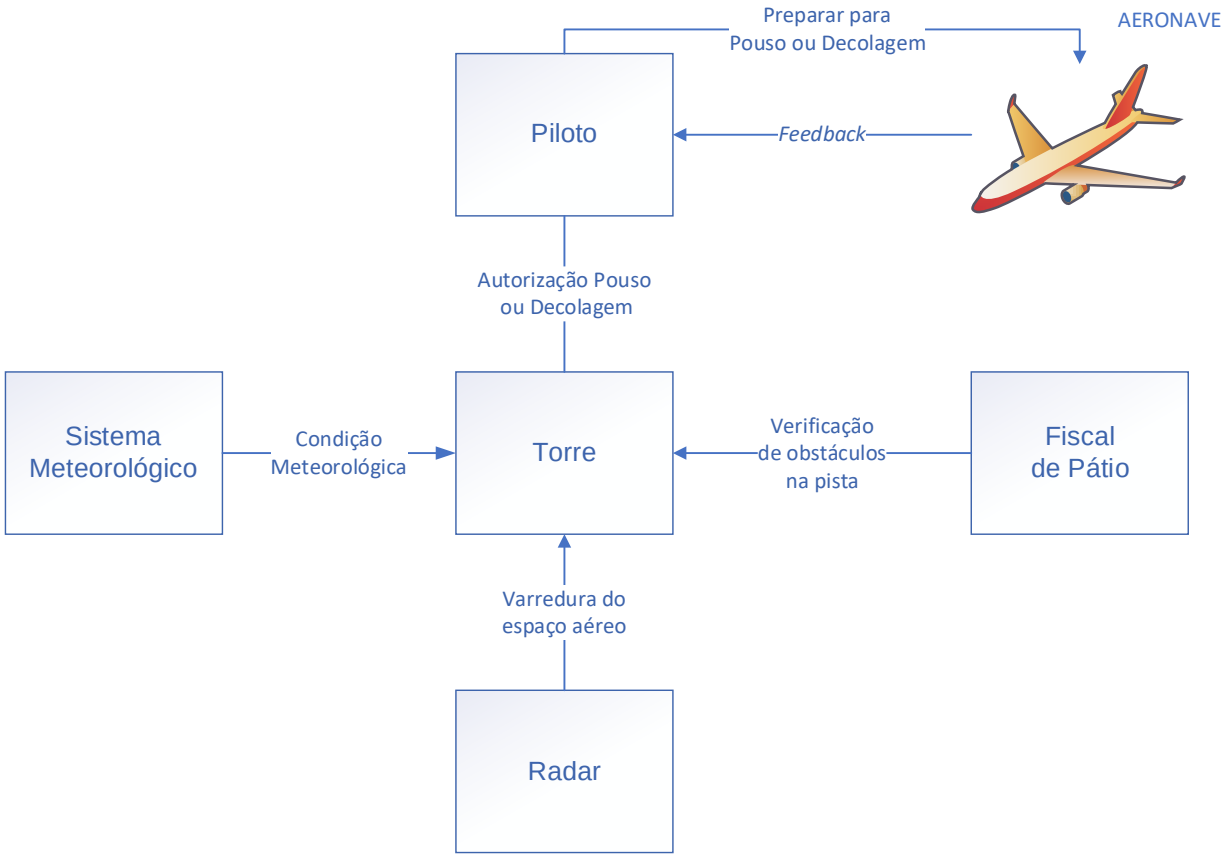

Figura 33 – Diagrama do Sistema de Torre de Controle de Aeródromo

Com o sistema especificado seguimos para as etapas de análise nas metodologias integradas BDMP, S-Cube, GTST-MLD e STPA-Sec.

6.2 APLICAÇÃO BDMP AO ESTUDO DE CASO 2

Uma possível representação *em Boolean logic Driven Markov Process* (BDMP) para o Sistema de Torre de Controle de Aeródromo (TWR) deste estudo de caso está ilustrado na Figura 34. O evento topo é “Aeronave colidiu” que considera sua ocorrência

entre uma combinação “OR” de duas sub árvores, uma representando “Colisão por falha de Controle” e outra “Colisão por falha Mecânica”.

A “Colisão por falha de Controle” ocorre se houver uma “Autorização Indevida” com “Desvio sem sucesso” caso haja “Objeto na rota de colisão”. Foi adicionado um *trigger* de “Autorização Indevida” para “Objeto na rota de colisão” que significa que a folha “Objeto na rota de colisão” está em modo *standby*, ou seja, não é considerada no estado inicial do sistema. Presume-se que se a autorização for concedida de forma coerente então são descartadas qualquer possibilidade de presença de obstáculos na rota da aeronave.

A “Autorização indevida” ocorre se houver “Erros de detecções” ou “Ataque cibernético”. Os ataques podem ser da classe *Tampering,* que altera os dados no sistema TWR consultado pelo controlador de tráfego aéreo, induzindo-o a uma decisão equivocada. A folha da classe *Spoofing* seria transmissão de mensagens para as aeronaves partindo de um falso Sistema TWR. Para proteção contra ataques *Tampering* são adicionados dois elementos tipo *Timed Security Event* (TSE). O primeiro “*Access Control List*” representa uma contramedida para ações de tentativa de acesso a um sistema. Caso essa defesa seja violada e um atacante consiga credenciais com privilégios, um trigger passa a considerar uma segunda linha de defesa, “*Intrusion Prevent” System*”. Esta representa uma contramedida para comportamentos de suspeitos de usuários com varredura da rede e arquivos do sistema. Com um possível avanço do atacante é ativado o *trigger* para a possibilidade do ataque de *Tampering*. Esta parte, portanto, modela as possibilidades de progressão de ataque em curso.

Não foram representadas proteções para a folha “Spoofing” por consideramos inicialmente que as comunicações entre piloto e controlador estão em modo claro, sem criptografia ou autenticação. Com o intuito de mitigar efeitos de falsas autorizações partindo de terceiros, uma contramedida seria transmitir uma autorização para todas as aeronaves na área de controle. Deste modo, todos os pilotos estariam cientes qual a aeronave irá pousar ou decolar. Porém, ao mesmo tempo essa medida também facilita uma ação maliciosa.

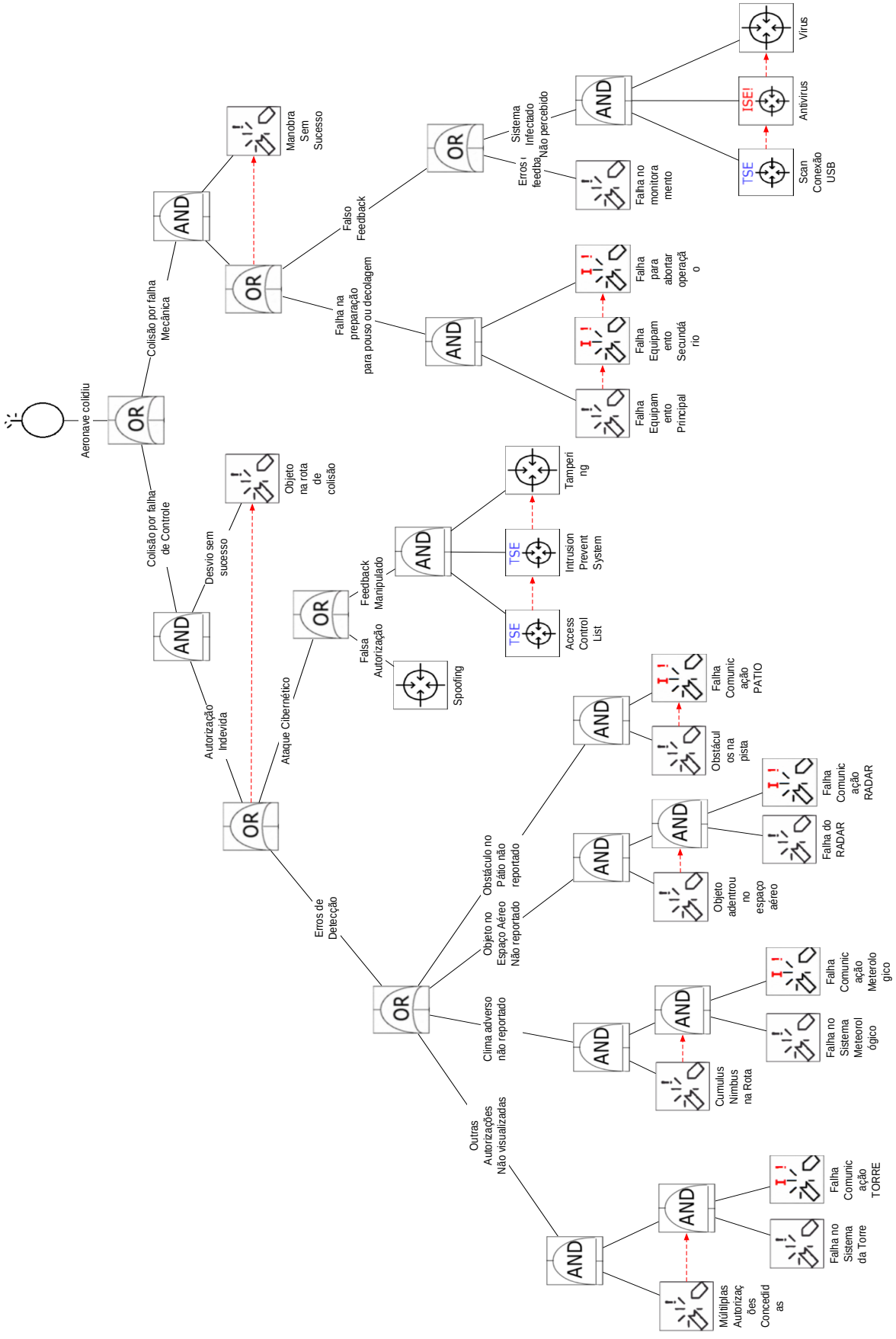

Figura 34 – BDMP: Torre de Controle de Aeródromo

Ainda para “Autorização indevida”, a sub árvore que representa “Erros de detecção” ocorre se qualquer um dos eventos “Outras autorizações não visualizadas”, “Clima Adverso Não Reportado”, “Objeto no Espaço Aéreo Não Reportado” e “Obstáculo no Pátio Não Reportado” ocorrerem. Todos os eventos com estruturas simulares que considera primeiramente a presença do obstáculo, caso ocorra, há triggers para considerar a possibilidade de falhas no respectivo sistema responsável e para falhas nos canais de comunicação.

Partindo para o lado mais a direita da árvore BDMP, ontem temos a sub árvore para representação “Colisão por falha Mecânica”, ocorre por uma “Falha na preparação para pouso ou decolagem” ou por “Falso Feedback” que induz o piloto a erros. E ainda, é inserido um “trigger” para representar um evento como última esperança “Manobra Sem Sucesso”.

No ramo “Falha na preparação para pouso ou decolagem”, o primeiro evento representa “Falha Equipamento Principal” no qual está associado a um *trigger* para “Falha Equipamento Secundário”, que por sua vez conta com outro *trigger* para “Falha para Abortar Operação”. Os dois últimos são representados por elementos BDMP do tipo *Failure On Demand* (ver Tabela 4). Este tipo de folha representa ocorrência de falhas no monitoramento e chaveamento.

No último ramo a direita, “Falso Feedback”, está representado como a possibilidade entre a ocorrência “Falha no monitoramento” ou “Sistema Infectado Não Percebido”. Como proteções cibernéticas foram representados “Scan Conexão USB” com *trigger* para “Antivirus”, com um último *trigger* para “Virus”. Essa proteção foi baseada no *modus operandi* do ataque Stuxnet (KRIAA; BOUISSOU; PIÈTRE-CAMBACÉDÈS, 2012).

Após executar a simulação, são apresentados na Tabela *23*, alguns *Cut Sets* que podem levar a ocorrência do evento topo indesejável. O *cut set* A da tabela refere-se à ocorrência de um ataque “Spoofing” seguido pela existência de “Objeto na rota de colisão”. Esse *cut set* com origem de *security* ainda precisa de um evento de *safety* para ocorrer. Esse resultado sugere a inclusão de medidas de proteção, como criptografia e

autenticação. Porém, tais medidas conflitam com a ação de notificar todas as aeronaves da área de controle para estarem cientes da autorizações em curso.

Tabela 23 – BDMP: MCS do Sistema TWR

ID	Tipo	Nº de Eventos	Eventos
A	Security/Safety	2	“Spoofing” seguido por “Objeto na rota de colisão” não desviado.
B	Safety	2	“Falha no Monitoramento” com falso feedback seguido por “Manobra Sem Sucesso”.
C	Safety	3	Formação de nuvens “Cumulus Nimbus”, seguido por “Falha do Sistema Meteorológico” e seguido por “Objeto na rota de colisão” não desviado.
D	Security/Safety	4	“Access Control List” violado, seguido por “Intrusion Prevent System”, e seguido por “Tampering” com autorização indevida e seguido por “Objeto na rota de colisão” não desviado.
E	Security	4	“Scan Conexão USB” violado, seguido por “Antivirus” violado, seguido por ataque de “Virus” com sucesso e seguido por “Manobra Sem Sucesso”.
F	Safety	4	“Falha Equipamento Principal”, seguido por “Falha Equipamento Secundário”, seguido por “Falha para abortar operação” e seguido por “Manobra Sem Sucesso”.

O *cut set* B refere-se à possibilidade de “Falha no Monitoramento” da Aeronave com falso feedback que induz o piloto a acreditar que acionou, por exemplo, os flaps e reversos para posição de decolagem. Seguido por “Manobra sem sucesso” que representa uma última tentativa de salvar a aeronave. Esse cut set é classificado somente do tipo *Safety*, porém, o cut set E que tem origem em *security* resultaria na mesma situação de perigo. Todos os cut sets com ligações *triggers* precisariam ocorrer na mesma ordem para que o evento topo ocorra, sendo isto, outra vantagem de representações BDMP.

Para resultados quantitativos, uma vez que, assertivamente estimados os parâmetros de entrada, a ferramenta calcularia o valor corresponde para o evento topo. Além disso, para o exemplo do cut set E, os parâmetros de “Scan Conexão USB” e “Antivirus” influenciariam na probabilidade de ocorrer “Virus”.

6.3 APLICAÇÃO S-CUBE AO ESTUDO DE CASO 2

A arquitetura para o Sistema da Torre de Controle de Aeródromo elaborada com os componentes disponíveis do S-Cube está representada na Figura 35. Foi possível modelar as comunicações da torre com o sistema meteorológico, com o sistema de radar e a comunicação via dados entre o piloto e controlador. Devido a indisponibilidade de componentes S-Cube para comunicação em voz como VHF, não foi possível modelar outros canais de comunicações entre o piloto e controlador e entre fiscal de pátio e controlador. Também não foi possível modelar a funcionalidade preparação para pouso ou decolagem por não encontrar componentes que representam uma ação de chaveamento para um módulo em *standby*.

A comunicação piloto e controlador é representada pelos componentes “Console_Aeronave” e “Console_ATCO_1” ambos da classe do S-Cube “IT_sys_cpt” que representam estações de trabalho com funcionalidades como sistema operacional Windows ou Linux. Em seguida temos os componentes “Console_Aeronave_SW” e “Console_ATCO_SW_1” ambos da classe “IT_soft_cpt” que representa os softwares executados em componentes “IT_sys_cpt”. Por fim, temos o componente “Data_Link_Aeronave” da classe “network_zone” interligando os componentes já mencionados (Figura *35*).

De maneira similar a anterior, representamos a conectividade do Controlador com o Servidor TWR pelos componentes “Console_ATCO_2” (IT_sys_cpt), “Console_ATCO_SW_2” (IT_soft_cpt), “Servidor_TWR” (IT_sys_cpt), “Servidor_TWR_SW_1” (IT_soft_cpt) e “Rede_Sistema_Torre” (network_zone). O componente “Servidor_TWR” também se conecta ao “Srv_Metero”, que representa um serviço de dados meteorológicos, na mesma rede “Rede_Sistema_Torre”. O “Servidor TWR” recebe os dados do radar pela conexão via “Scada_Server_Soft” da classe “scada_server_soft_cpt” que representa um sistema SCADA aplicado ao radar. Para o lado do sistema radar elaboramos uma arquitetura básica de sistema de controle com sensor, atuador e controlador. A estrutura é similar ao estudo de caso anterior. Neste sistema os módulos se conectam a uma rede interna para trocar informações.

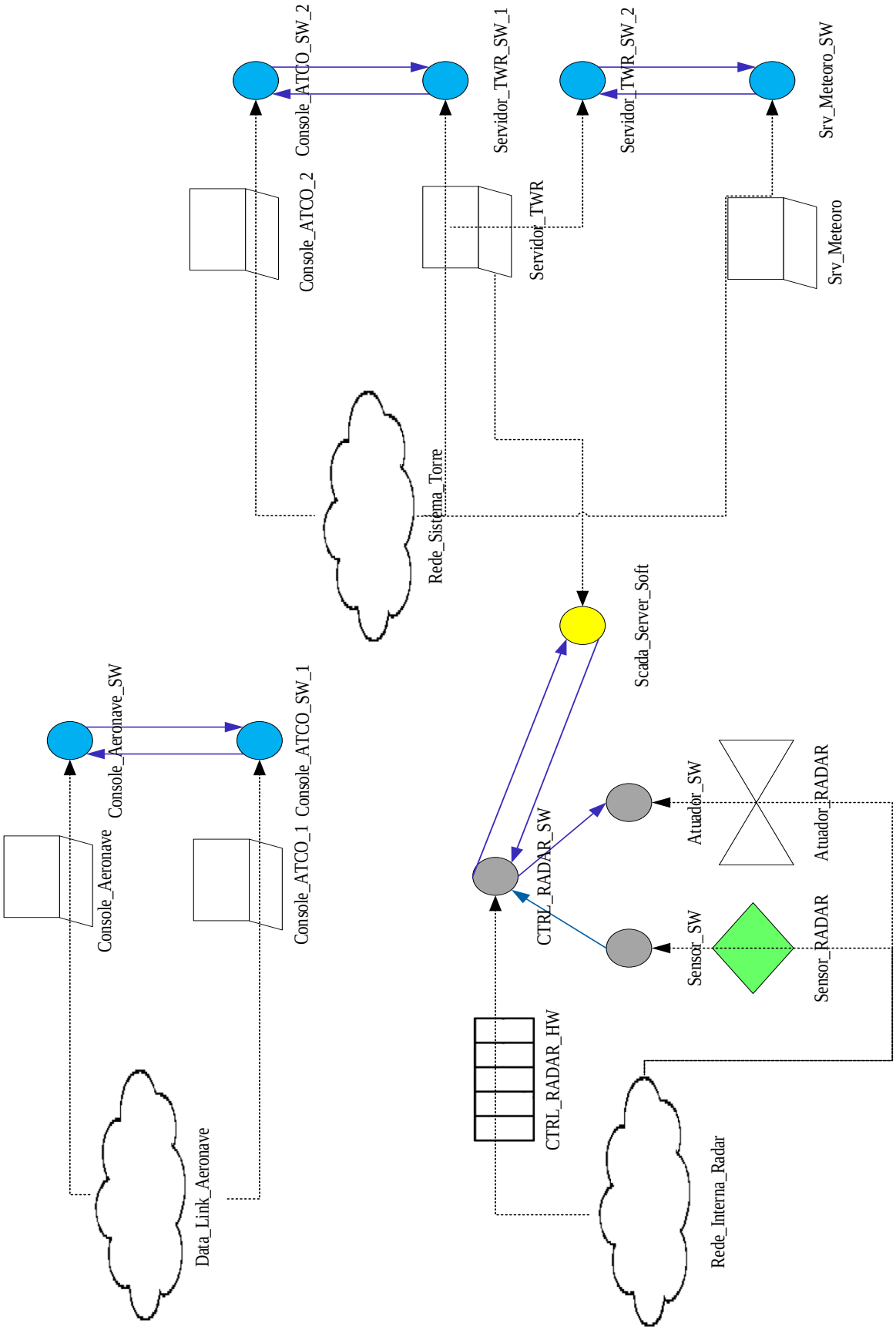

Figura 35 – S-Cube: Torre de Controle de Aeródromo

Tabela 24 – S-Cube: Classes para o Sistema TWR

Componentes TWR	**Classes S-Cube**	**Alguns tipos de ameaças catalogadas no S-Cube KB associadas a classe**
Console_Aeronave, Console_ATCO_1, Console_ATCO_2, Servidor_TWR, Srv_Meteoro	IT_sys_cpt	**privilege_escalation_attack:** atacante obtém credenciais de administrador. **accidental_failure**: falha não intencional; **Access (physical):** atacante em contato com a estação.
Console_Aeronave_SW, Console_ATCO_SW_1, Console_ATCO_SW_2, Servidor_TWR_SW_1, Servidor_TWR_SW_1, Srv_Meteoro_SW	IT_soft_cpt	**bypass_autentification**: atacante burla a autenticação. **Exploit vuln priv escalation:** atacante obtém credenciais de administrador. **Exploit vuln integrity loss:** atacante altera dados. **Exploit vuln denial of service:** atacante indisponibiliza o serviço. **Exploit vuln confidentiality loss:** atacante obtém dados sigilosos.
“Scada_Server_Soft””	scada_server_soft_cpt	**send_false_instructions_to_actuator**: atacante altera instruções no software SCADA; **send_no_instructions_to_actuator**: atacante remove instruções no software SCADA.
CTRL_RADAR_HW	process_controller_cpt	**send_false_instructions_to_actuator**: atacante altera instruções no software controlador; **send_no_instructions_to_actuator**: atacante remove instruções no software controlador.
CTRL_RADAR_SW	process_controller	**accidental_failure**: falha não intencional; **compromise_comm_link**: falha de comunicação; **ccf_comp**: falhas de causa comum. Ex. Energia.
Sensor_SW	sensor_soft_cpt	**send_false_measures**: atacante falsifica as medições enviadas pelo sensor; **send_no_measures**: atacante remove as medidas enviadas pelo sensor.
Sensor_RADAR	sensor	**accidental_failure**: falha não intencional; **compromise_comm_link**: falha de comunicação; **ccf_comp**: falhas de causa comum.
Atuador_SW	actuator_soft_cpt	**actuator_does_not_act_properly**: atacante falsifica ou remove uma ação de controle do atuador.
Atuador_RADAR	actuator	**accidental_failure**: falha não intencional; **compromise_comm_link**: falha de comunicação; **ccf_comp**: falhas de causa comum.
Data_Link_Aeronave, Rede_Sistema_Torre, Rede_Interna_Radar	network_zone	**jamming_attack**: negação de serviço da rede; **attacker_scan_network**: atacante coleta informações de endereços, porta e serviços ativos; **attacker_stablish_connection**: atacante estabelece uma conexão com algum ativo da rede; **bypass_autentification**: atacante burla a autenticação.

As classes de todos componentes da arquitetura S-Cube na Figura 35 são mostradas na Tabela 24 junto com algumas falhas e vulnerabilidades associadas. Na simulação mostramos na Tabela 25 apenas alguns cenários de ataques e de falhas para a arquitetura deste estudo de caso.

Tabela 25 – S-Cube: Cenário de Ataques e Falhas TWR

Cenário	Etapas	Descrição
A	1. access_network (Rede_Sistema_Torre) 2. attacker_scan_network (Rede_Sistema_Torre) 3. bypass_autentification (Servidor_TWR) 4. attacker_stablish_connection (Servidor_TWR) 5. exploit_Vuln_Priv_Escalator (Servidor_TWR) 6. Exploit_Vuln_Integrity_loss (Servidor_TWR**)**	Atacante conecta um dispositivo externo na rede, realiza a varredura, contorna uma autenticação e estabelece conexão com o servidor da Torre. Nas etapas seguintes obtém credenciais de administrador e altera um dado no sistema TWR.
B	1. Access (Data_Link_Aeronave) 2. attacker_stablish_connection (Console_Aeronave) 3. Envia Falsa Autorização	Atacante acessa a rede de comunicação com a aeronave e envia falsa autorização.
C	1. access (Rede_Interna_Radar) 2. attacker_scan_network (Rede_Interna_Radar) 3. attacker_stablish_connection (CTRL_RADAR_SW) 4. send_false_feedback (CTRL_RADAR_SW)	Atacante falsifica dados do sistema de detecção do radar.

O cenário A mostra os passos de um ataque *Tampering* ao Sistema da Torre para induzir o controlador conceder uma autorização equivocada. O cenário B mostra os passos de um ataque *Spoofing* onde se envia diretamente uma mensagem para a aeronave de uma autorização ilegítima. O cenário C mostra os passos para um ataque ao sistema SCADA do RADAR falsificando feedback de detecção de objetos em rota de colisão.

Os resultados dos cenários de ataques destacados estimulam os projetistas a incluir barreiras para mitigar ou anular os passos. Além disso, um projetista pouco experiente em análise de segurança crítica e segurança cibernética poderia ser beneficiado com os resultados fornecidos pela ferramenta.

6.4 APLICAÇÃO CHASSIS AO ESTUDO DE CASO 2

Realizando a análise do Sistema TWR sob a ótica do *Combined Harm Assessment of Safety and Security for Information System* (CHASSIS), são elaborados o Diagrama de Casos de Uso (D-UC) ilustrado na Figura 36. Para o D-UC são definidos os atores como sendo o Piloto, Sistema da Aeronave, Sistema Meteorológico, Radar, ATCO TWR (representa o Controlador de Tráfego e o Sistema da Torre) e o Fiscal de Pátio. São relacionados os casos de uso com os atores, e as dependências entre casos de uso. Na Figura 36 pode ser observado que o caso de uso “Autorizar Pouso ou Decolagem” possui como dependência os outros casos de uso “Alertar Condições Climáticas”, “Informar Situação do Espaço Aéreo” e “Comunicar Estado do Pátio”. O caso de uso “Preparar Aeronave para pouso e decolagem” é de responsabilidades de dois atores “Piloto” e “Sistema da Aeronave”. Este caso de uso possui “Autorizar Pouso ou Decolagem” como dependência.

.

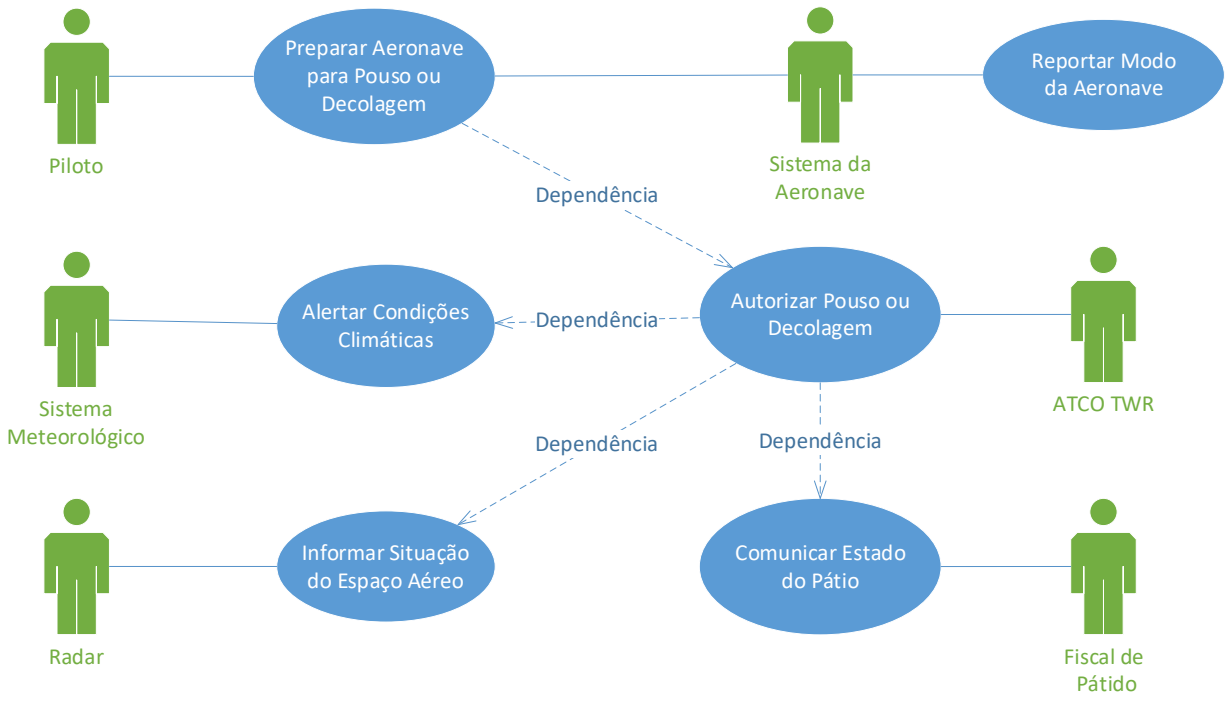

Figura 36 – CHASSIS: Diagrama de Caso de Uso (D-UC) Sistema TWR

Os diagramas de caso de uso auxiliam na visualização das relações entre atores e sistemas sem que necessariamente estejam conectados em uma rede.

Avançando para o processo 2 do CHASSIS, que corresponde a criação do Caso de Uso Textual (T-UC), é elaborado o T-UC para o “Autoriza Pouso ou Decolagem” mostrado na Tabela 26. No T-UC está descrito o caminho básico e também um alternativo que corresponde a uma comunicação por voz com o piloto. O T-UC portanto explica e adiciona detalhes ao D-UC e induz o analista a pensar nas vulnerabilidades e perigos envolvidos para que as contramedidas sejam implantadas ou aprimoradas.

Tabela 26 – CHASSIS: Um Caso de Uso Textual para o Sistema TWR

Campo	**Dado**
Nome	Autorizar Pouso ou Decolagem
Iteração	1
Resumo	Ação responsável para permitir ao piloto a preparação da aeronave para o pouso ou decolagem.
Caminho básico	1- ATCO TWR verifica a existência de outra autorização em curso; 2- ATCO TWR verifica alerta de condições climáticas; 3- ATCO TWR verifica Situação do Espaço Aéreo; 4- ATCO TWR Verifica Estado do Pátio; 5- ATCO TWR Envia Informação ao piloto.
Caminho alternativo	Não possui.
Gatilhos	Solicitação do piloto
Premissas	Assumimos que o controlador de tráfego atua fielmente com os *feedbacks* recebidos, ou seja, não há erros por parte do controlador.
Pré-condições	-
Pós-condições	-
Regra de Negócio relacionada	Atender aos requisitos: não haver outra autorização em curso, condição climática favorável, pátio e espaço aéreo livre de obstáculos.

Ainda na Tabela 26, são adicionadas as premissas consideradas neste estudo de caso e ignoramos erros oriundos de imperícias, imprudências e negligências partindo do Controlador de Tráfego, com o objetivo de conter uma expansão da análise e comprometer a didática. No último campo é adicionado a regra de negócio vinculada ao caso de uso.

Prosseguindo para a terceira etapa do CHASSIS, elaboramos o Diagrama de Sequência (SD) ilustrado na Figura 37 representando as trocas de mensagens entre o

Piloto e o Controlador da Torre (ATCO TWR) com origem no pedido do piloto para pouso ou decolagem. No SD são representadas as consultas do ATCO TWR ao Sistema Meteorológico, Radar, Fiscal de Pátio e uma consulta interna para verificação de existência de outra autorização em curso.

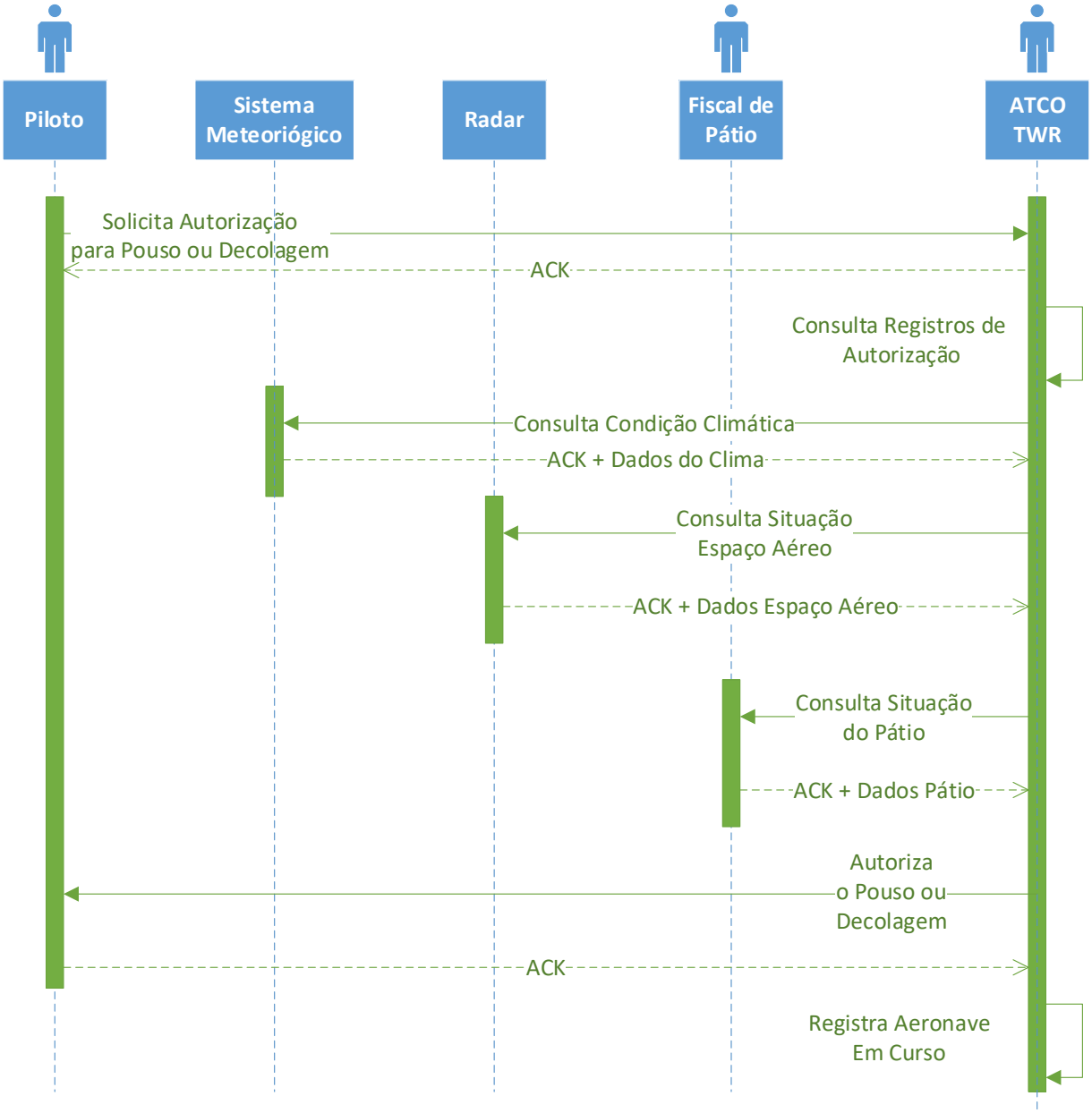

Figura 37 – CHASSIS: Diagrama de Sequência (SD) para o sistema TWR

Os três primeiros processos do CHASSIS, Diagramas de Casos de Uso (D-UC) Casos de Uso Textual (T-UC) e Diagramas de Sequência (SD) fazer parte do grupo de

Levantamentos de Requisitos Funcionais conforme mostrado na Figura 8 que servem como subsídios para as etapas seguintes que são parte do grupo de Levantamentos de Requisitos de *Safety e Security* .

A partir do D-UC, mostrado na Figura 36, já pode ser criado o Diagrama de Casos de Erros de USO (MUSD), mostrado na Figura 38, corresponde ao quarto processo CHASSIS.

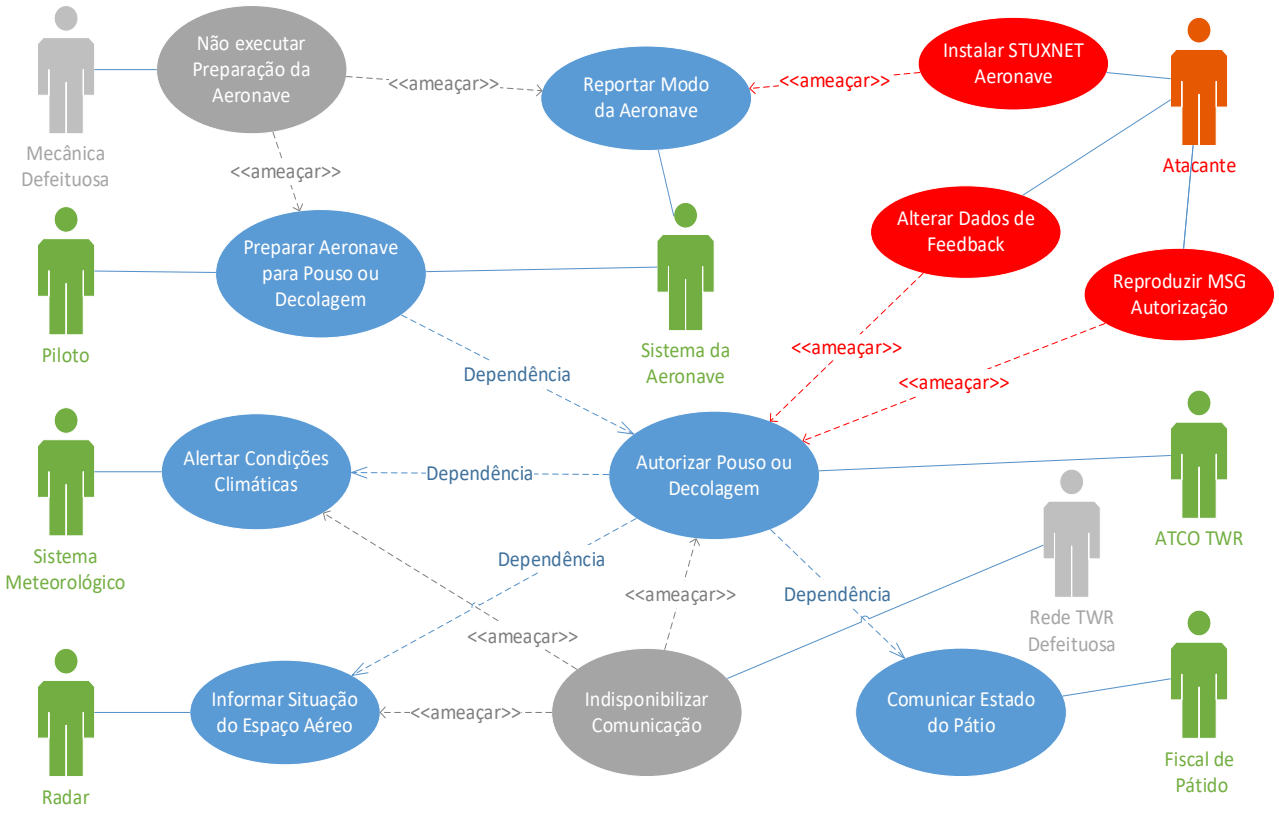

Figura 38 – CHASSIS: Erros de Uso (MUSD) para o sistema TWR

Continuando na Figura 38, são adicionados os Erros de Uso que podem ameaçar os casos de uso. Em vermelho as ameaças partindo de possibilidades de ataques cibernéticos e em cinza as ameaças de falhas não intencionais. Foram adicionadas somente algumas delas para conter expansão. As ameaças cibernéticas retratam a possibilidade de um atacante alterar dados de *feedback* e reprodução de mensagens falsas de autorização que comprometem o caso de uso “Autorizar Pouso ou Decolagem”. Também retrata a possibilidade de um ataque similar ao Stuxnet que

fraudaria os dados de visualização de *status* de componentes e subsistemas da aeronave. O MUSD estimula os projetistas a reportarem e incluir defesas para os ataques levantados e ainda nesta etapa se iniciam a elaboração das tabelas HAZOP para cada caso de uso. A Tabela 27 apresenta um registro para o caso de uso “Autorizar Pouso ou Decolagem”. Nesse registro já deve ser incluído recomendações de defesa para o erro de uso apresentado.

Tabela 27 – CHASSIS: um registro adaptado de HAZOP para o Sistema TWR

Função	Parâmetro	*Guide word*	Causas	Consequência	Dano	Recomendação
Autorizar Pouso ou Decolagem	-Condição Climática; -Varredura Espaço Aéreo; -Fiscalização do Pátio; e -Verificação de autorizações.	Other than	-Alteração de dados de feedback	Existência de obstáculos	Colisão de aeronave	- Incluir Autenticação - Incluir redundância de visualização
		More	-Captura de pacotes com Reprodução de MSG de autorização			-Criptografia de pacotes; -Confirmação por VHF

Na Tabela 27 temos o primeiro registro associado a *guideword Other than* que corresponderia a uma falsificação dos *feedbacks* recebidos. Como consequência, poderiam haver obstáculos no ar, em solo ou formação de nuvens *Cumulus Nimbus* na rota da aeronave. O dano seria a colisão da aeronave. No último campo, recomendação, o analista deve incluir as defesas para a ameaça relatada na tabela.

Seguindo para o quinto processo, que se trata da elaboração dos casos de erro de uso textual (T-MUC), a Tabela *28* apresenta a descrição das possibilidades de ataques (em vermelho) em cada passo do caminho básico. Por exemplo, é incluído a possibilidade de um atacante capturar um pacote em um roteador, impedir a transmissão do pacote original e transmitir outro pacote com dados alterados. No campo Pontos de Mitigação está descrito as medidas para combater as ações maliciosas, entre elas, autenticação, criptografia e verificação em canal redundante.

Além disso, são descritos ainda os perfis dos atacantes, os riscos e os afetados que podem sugerir uma atenção especial a implementação dos pontos de mitigação.

Tabela 28 – CHASSIS: um Erro de Uso Textual (T-MUC) o Sistema TWR

Campo	**Dado**
Nome	Autorizar Pouso ou Decolagem
Iteração	1
Resumo	Ação responsável para permitir ao piloto a preparação da aeronave para o pouso ou decolagem.
Caminho básico	1- ATCO TWR verifica a existência de outra autorização em curso; **a. Atacante obtém credenciais de administrador e altera a base de dados do sistema TWR** 2- ATCO TWR verifica alerta de condições climáticas; **a. Atacante captura pacote, altera conteúdo e encaminha para Sistema TWR** 3- ATCO TWR verifica Situação do Espaço Aéreo; **a. Atacante captura pacote do Radar, altera conteúdo e encaminha para Sistema TWR** 4- ATCO TWR Verifica Estado do Pátio; 5- ATCO TWR Envia Informação ao piloto. **a. Autorização encaminhada com base em *feedbacks* adulterados** **b. Atacante captura pacote reproduz mensagem de autorização**
Caminho alternativo	Não possui.
Pontos de Mitigação	1- Para manter a integridade dos f*eedbacks*: **a. Incluir redundância na visualização; e** **b. Receber alertas de sistemas indisponíveis.** 2- Para evitar falsas autorizações: **a. Incluir segundo fator de confirmação via Rádio VHF;** **b. Broadcast para todas as aeronaves para estarem cientes de qual está com autorização em curso.** 3- Para ambos: **a. Implementar autenticação entre os sistemas consultados;** b. **Criptografia de pacotes.**
Premissas	-.
Pré-condições	-
Perfil de Usuário Indevido	Atacante com motivações de terrorismo, assassinato de um passageiro específico, operações entre países em conflito por exemplo.
Riscos e Partes Afetadas	Colisão de aeronave com provável óbito da tripulação, de passageiros e pessoas em solo. Inutilização do aeroporto de uma cidade. Abalo moral da população e autoridades do país afetado.

Seguindo adiante, no sexto processo, elaboramos o Diagramas Sequencial de Erros de Uso (MUSD) para “Autorizar Pouso ou Decolagem” e está ilustrado na Figura

39. A imagem do atacante (em vermelho), representa a operação de escuta de pacote enviado pelo ATCO TWR, mas com retornos de respostas capturados com anulação de suas transmissões. Com a posse dos pacotes, o atacante altera o conteúdo e encaminha para o piloto.

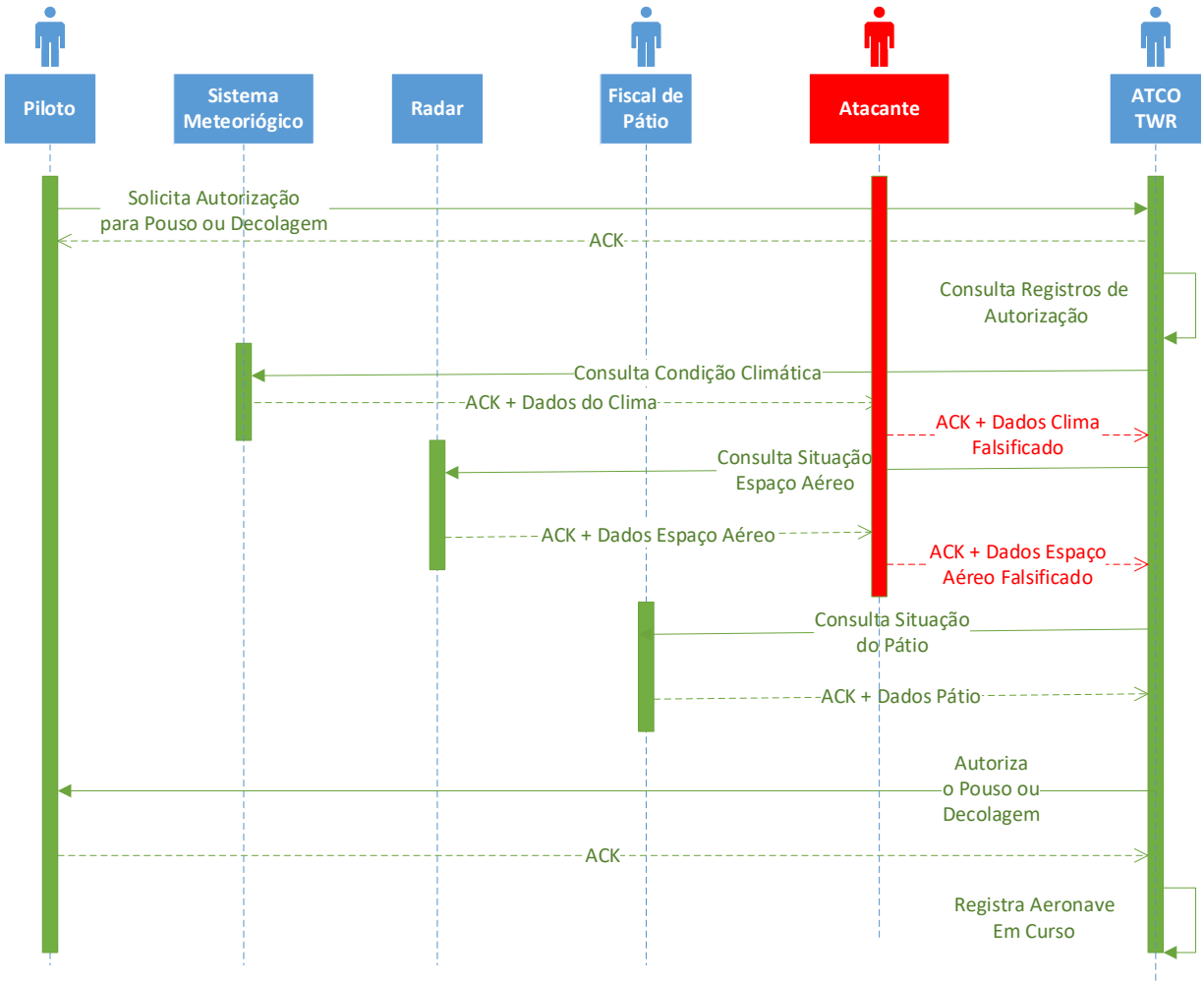

Figura 39 – CHASSIS: Diagrama Sequencial de Erros de Uso (MUSD) do TWR

Após finalizados os processos quatro, cinco e seis do CHASSIS, Criação dos Diagramas de Erros de Uso (D-MUC), Erros de Uso Textual (T-MUC) e Digramas Sequenciais de Erros de Uso (MUSD), conclui-se o segundo grupo, Levantamento de Requisitos de *Safety/Security*, mostrado na Figura 8. A partir daí, pode-se retornar ao primeiro processo e realizar novas iterações de diagramas e modos textuais ou avançar para o último grupo, Especificação dos Requisitos de *Safety/Security.* Esta etapa final

corresponde aos processos de criação da tabela HAZOP (sétimo processo) selecionando os registros criados anteriormente. E por último, especificação dos requisitos finais para o desenvolvimento do sistema.

Como saída da aplicação completa dos processos CHASSIS para este estudo de caso, Sistema TWR, são obtidos alguns requisitos mostrados na Tabela 29. Os requisitos 4 e 5 são conflitantes. Foram aplicadas somente para um dos casos de uso da Figura 36, que se refere ao “Autorizar Pouso ou Decolagem”. Para os resultados completos deve-se aplicar a todos os casos de uso levantados, o que tornaria extenso e não adicionaria outra contribuição no entendimento e avaliação da metodologia CHASSIS.

Tabela 29 – CHASSIS: Requisitos Especificados para o Sistema TWR

Nº	**Requisitos**
1	O sistema TWR deve incluir redundância na visualização. Outra console obtendo dados por canais alternativos.
2	O sistema TWR deve receber alertas de indisponibilidade dos sistemas consultados.
3	A operação de autorização deve ser confirmada também via Rádio VHF;
4	As autorizações devem ser transmitidas para todas as aeronaves para estarem cientes das autorizações em curso.
5	A comunicação pelo enlace de dados entre o sistema TWR e a aeronave deve ser criptografada e seguir protocolos de autenticação.
6	O sistema TWR deve ter criptografia de pacotes com os sistemas consultados.

6.5 APLICAÇÃO GTST-MLD AO ESTUDO DE CASO 2

A análise do sistema TWR conduzida de acordo com o *Goal Tree Succes Tree Master Logic Diagram* (GTST-MLD) obtemos a representação na Figura 40. A função topo da *Goal Tree* (GT) foi definida em “Prevenir Colisão de Aeronaves”, que é uma junção das subfunções “Controlar Tráfego no Aeródromo” e “Garantir Funcionamento da Aeronave”. Essa definição foi concluída a partir da premissa de que se o tráfego no aeródromo for controlado corretamente e as aeronaves estiverem livres de defeitos então não haverá colisão entre aeronaves.

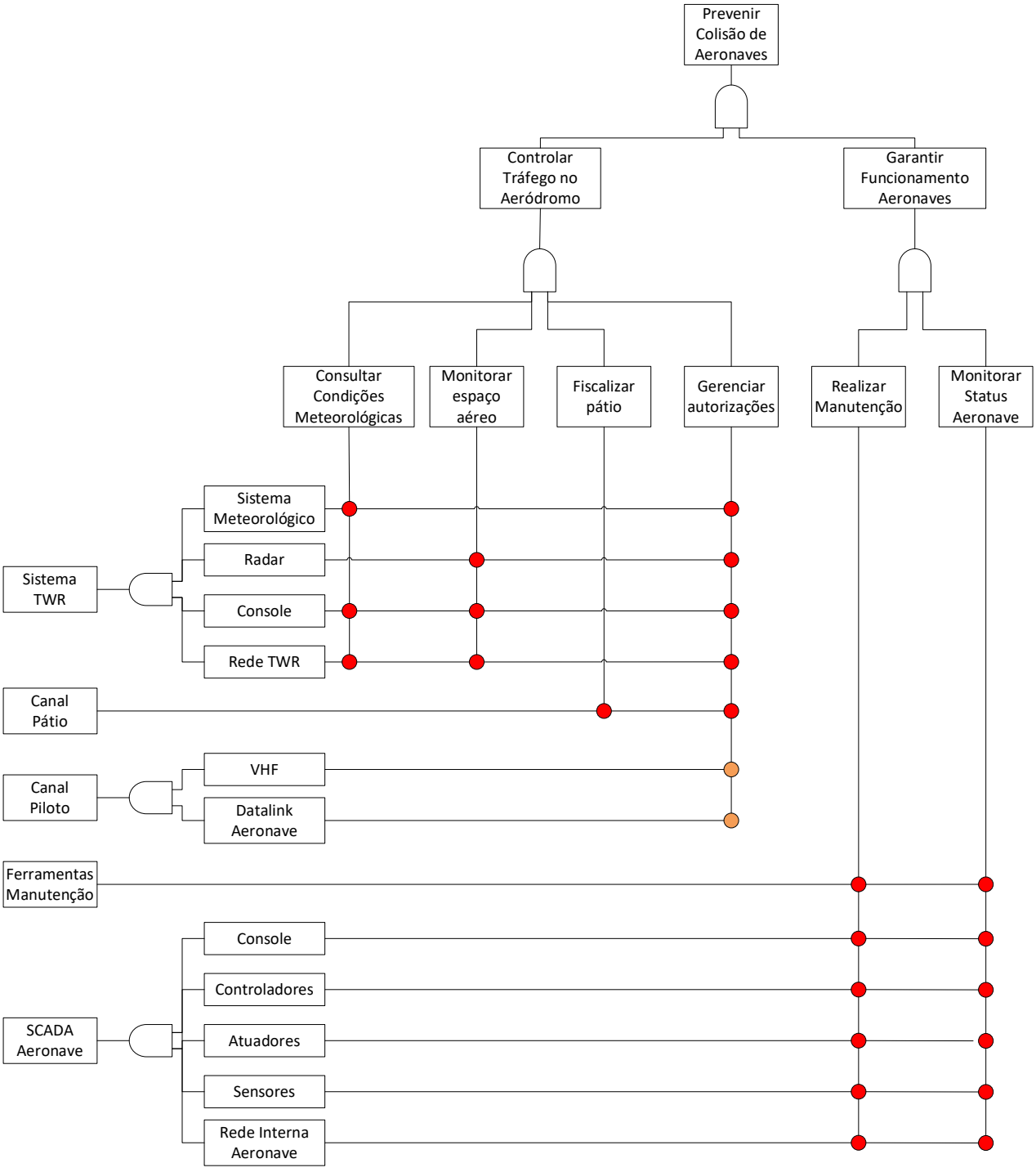

Figura 40 – GTST-MLD do sistema TWR

Continuando na Figura 40, a subfunção “Controlar Tráfego no Aeródromo” é uma junção das outras subfunções “Consultar Condições Meteorológicas”, “monitorar espaço aéreo”, “Fiscalizar pátio” e “Gerenciar autorizações” que correspondem aos requisitos para uma autorização de pouso ou decolagem. No lado direito da GT, a função “Garantir Funcionamento Aeronave” é uma junção de “Realizar Manutenção” e “Monitorar Status da Aeronave”. Ambas estão relacionadas com os componentes do sistema “SCADA

Aeronave” na *Success Tree* (ST) e ainda foi definido um novo componente “Ferramentas Manutenção”. A construção do diagrama levou a definir esse componente de forma intuitiva, partindo da lógica de que para garantir o funcionamento de algo é necessário realizar manutenção. E os componentes para se realizar manutenção seriam ferramentas de manutenção.

Ainda em continuação do GTST-MLD deste estudo de caso, são selecionados todos os componentes que são passíveis de sofrer diretamente um ataque cibernético, ou seja, os componentes digitais. Na Figura 41 são apresentadas a extensão do diagrama relacionando os componentes com as possíveis ameaças classificadas pelo modelo STRIDE.

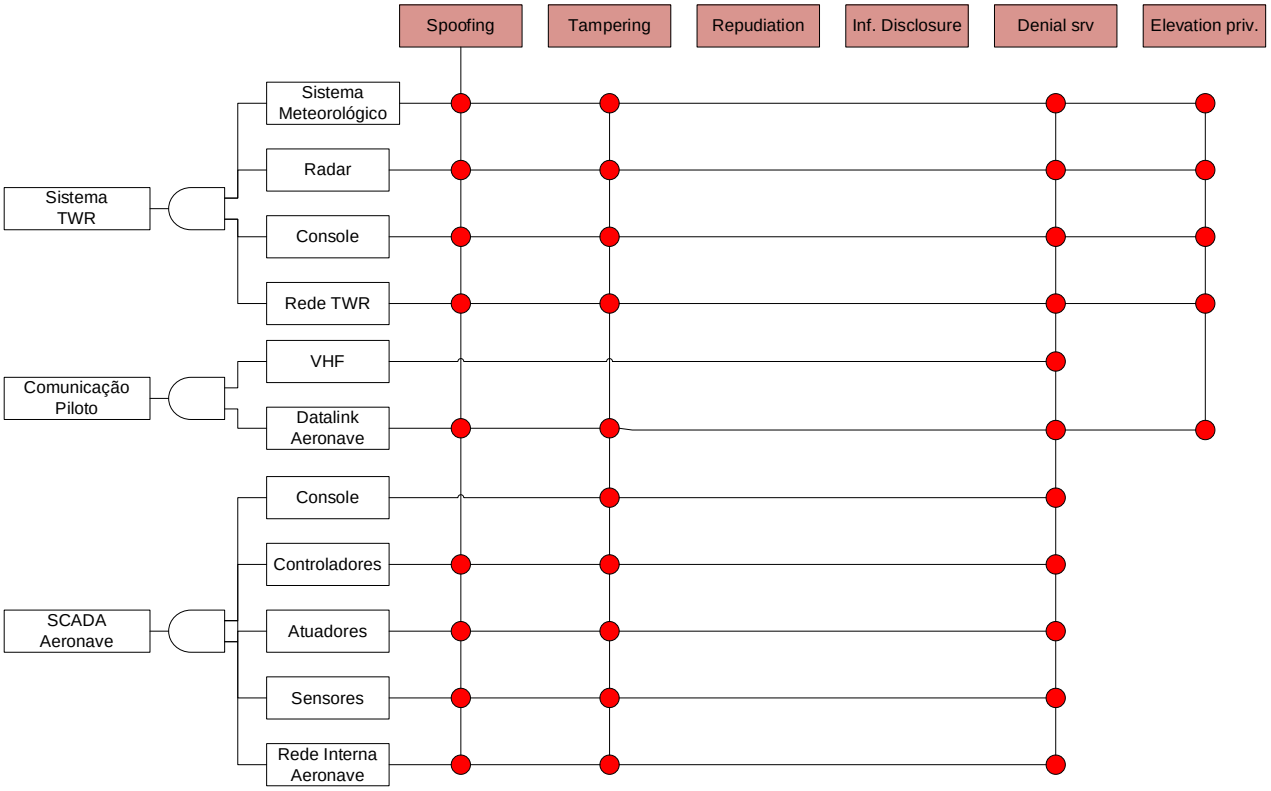

Figura 41 – GTST-MLD: ameaças cibernéticas ao sistema TWR

A Figura 40 e Figura 41 fazem parte do mesmo diagrama e estão conectadas, porém, foram divididas apenas para facilitar a visualização. Pelo diagrama torna-se possível verificar quais funções podem ser afetadas por ameaças cibernéticas e de qual tipo. Além disso, para cada relação de ameaça ao componente deve ser especificado a possível forma de ataque e a contramedida empregada.

Como contribuição do GTST-MLD para esta análise temos a definição do componente “Ferramentas de Manutenção” já mencionado. A representação da função “Fiscalizar Pátio” está livre de influência cibernética por presumir que pode ser feita por outros meios como visual, telefones ou presencial. Na comunicação com o piloto, o canal VHF está livre de ataque de *Tampering*, mas ainda associado a negação de serviço. Isto significa que, o estabelecimento do uso de dois canais para uma autorização poderia reduzir drasticamente as chances de sucesso de uma retransmissão de mensagens por um atacante via enlace de dados.

6.6 APLICAÇÃO STPA-SEC AO ESTUDO DE CASO 2

Na realização da análise para *System-Theoretic Process Analysis for Security* (STPA) no Sistema TWR, são consideradas as perdas da missão mostradas na Tabela 30. Foram definidas quatro tipos de perdas sendo: L-1 associada à vida e integridade física de tripulação, passageiros e pessoas em solo, L-2 associada à colisão de aeronave com obstáculo no ar ou em solo ou ainda por mal tempo, L-3 associada a insatisfação da viagem por passageiros decorridos por traumas, sustos ou aborrecimentos na experiência da viagem e L-4 associada a prejuízos econômicos em virtude de reparos materiais, indenizações judiciais, causas trabalhistas e perdas de oportunidades de novos negócios.

Tabela 30 – STPA-Sec: Perdas de Missão para o Sistema TWR

ID	**Perdas (ou *Losses*)**	**Descrição**
L-1	Óbitos e Feridos	Perda associada as possibilidades de ocorrência de óbitos e feridos seja por colisão de aeronaves ou por qualquer ação resultante de falha de missão do Sistema TWR.
L-2	Colisão de Aeronave	Perda associada a possibilidade de ocorrência de colisão de aeronaves com obstáculos no ar ou no solo.
L-3	Insatisfação de Passageiros	Perda associada a possibilidade de passageiros sofrerem uma experiência traumática em decorrência de sustos.
L-4	Prejuízos Econômicos e Danos à Reputação	Perda associa da possibilidade de perdas financeiras significativas e impacto na credibilidade das organizações responsáveis.

Em seguida são definidas as situações de perigos, mostrado na Tabela 31, que podem impactar nas perdas. Foram levantados o perigo H-1, Aeronave decolando ou pousando sem autorização do ATCO, que representa uma situação que pode estar

atrelada a uma falsa autorização partindo de um ataque cibernético tipo *spoofing*. O H-1 está relacionado com as perdas L-1, L-2 e L-4. A L-3 foi excluída pela premissa de que não chegaria no conhecimento da tripulação e dos passageiros. Próximo Perigo, H-2 Aeronave em operação com manutenção vencida, situação que pode resultar em falhas mecânicas e de outros sistemas da aeronave e está associada às perdas L-1, L-2 e L-4. Para o perigo H-3, Obstáculo em distância da aeronave inferior ao limite mínimo, representa uma autorização concedida sem a verificação correta dos requisitos e está associada a L-1, L-2 e L-4. No perigo H-4, piloto executa manobra brusca ao desviar de obstáculo. É considerado que o piloto conseguiu desviar de um obstáculo ao custo de prover uma experiência traumática aos passageiros, associada as L-3 e L-4. Por fim, para o perigo H-5, Aeronave com atrasos significativos no pouso ou decolagem, é levantado a possibilidade do controlador não estar recebendo corretamente os dados necessários para prover uma autorização segura. Como consequência, provoca atrasos nos voos e aborrecimento dos passageiros.

Tabela 31 – STPA-Sec: Situações de Perigo do Sistema TWR

ID	**Perigos ou ameaças (*Hazards* e *Threats*)**	**Perdas associadas**	**Descrição**
H-1	Aeronave decolando ou pousando sem autorização do ATCO	L-1, L-2 e L-4	Situação que pode estar associada a emissão de falsa autorização
H-2	Aeronave em operação com manutenção vencida	L-1, L-2 e L-4	Situação associada a falha mecânica na aeronave.
H-3	Obstáculo em distância da aeronave inferior ao limite mínimo.	L-1, L-2 e L-4	Representa uma autorização concedida equivocadamente pelo controlador.
H-4	Piloto executa manobra brusca ao desviar de obstáculo	L-3 e L-4	Situação em que o piloto percebe um obstáculo e consegue desviar a aeronave.
H-5	Aeronaves com atrasos no pouso ou decolagem	L-3 e L-4	Situação que o controlador não recebe os feedbacks para autorizar operações.

Aproveitando a capacidade do STPA-Sec as perdas e perigos incialmente levantados correspondem às questões de segurança crítica e de utilidade funcional da aplicação. Além das possibilidades de ameaças cibernéticas.

Avançando com a análise, são elaboradas restrições do sistema, ou *System Constraints* (SC) que é mostrado na Tabela 32 e estão relacionados com Perigos ou Ameaças. As restrições são levantadas para mitigar as ativações dos perigos. Todos os perigos devem ter pelo menos uma SC.

Tabela 32 – STPA-Sec: Restrições do Sistema TWR

ID	**Restrições do Sistema (*System Constraints*)**	**Perigos ou Ameaças associadas**
SC-1	As autorizações para pouso e decolagem devem ser verificadas via dados e via voz entre controlador e piloto.	H-1, H-3, H-4
SC-2	As aeronaves devem ter suas manutenções verificadas pelo piloto.	H-2
SC-3	O controlador deve consultar os sistemas para verificar obstáculos no espaço aéreo, no pátio, mal tempo e conceder somente uma única autorização por vez.	H3, H4
SC-4	O radar, o sistema meteorológico, a comunicação com o pátio e o sistema da torre devem ter redundância para casos de indisponibilidade.	H-1, H-3, H-5

Avançando para o processo 2 do STPA-Sec, a estrutura de controle para atender as SC identificadas para o Sistema TWR é ilustrada na Figura 42. O controlador ATCO tem interações com o Piloto, com o sistema TWR e com o Fiscal do Pátio. O Piloto tem interações com os sistemas da cabine de comando para configurar os equipamentos da aeronave para modo pouso ou decolagem. Os sistemas da cabine interagem com a aeronave por meio de atuadores e sensores. A interação do piloto direto com a aeronave representa as ações de verificação de manutenção. O sistema da torre interage com o sistema meteorológico e sistema radar para fornecer os *feedbacks* ao controlador da torre. O sistema radar foi representado com interação com o equipamento radar por meio de atuador e sensor.

Ainda na Figura 42, cada interação representada pelas setas em laranja corresponde a uma ação de controle partindo de um controlador de nível hierárquico superior para um de nível abaixo. As setas em azul representam os *feedbacks* dos controladores inferiores e ou de processos controlados. A estrutura de controle pode ser modificada se no avanço da análise forem observadas ações não cobertas.

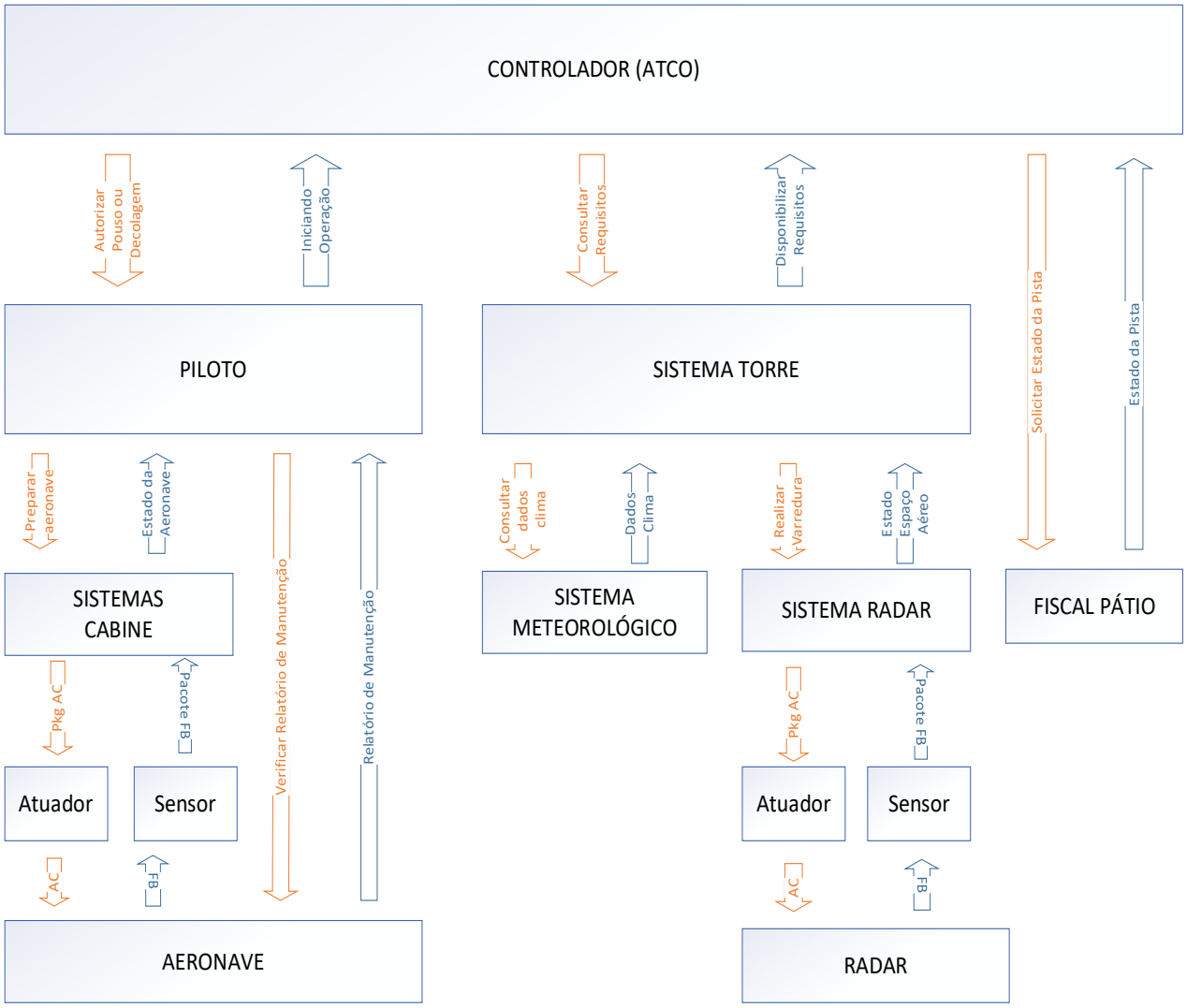

Figura 42 – STPA-Sec: Estrutura de Controle do Sistema TWR

Ainda no processo 2 do STPA-Sec são definidas as responsabilidades dos principais controladores e estão mostradas na Tabela 33 .

Tabela 33 – STPA-Sec: Responsabilidade do Controlador TWR

ID	Controlador	Responsabilidade	Restrição
R-1	Controlador ATCO	Autorizar ou negar pouso ou decolagem conforme consulta aos requisitos	SC-1
R-2	Piloto	Operar aeronaves somente com inspeções válidas.	SC-2
R-3		Quando autorizado, preparar a aeronave para o pouso ou decolagem.	SC-2
R-4	Sistema Torre	Disponibilizar a visualização dos requisitos de forma íntegra.	SC-3, SC-4

Avançando para o processo 3, são definidas algumas ações de controle inseguras (UCA) que são mostrada na Tabela 34. Foram levantadas somente para as ações de controle do ATCO e do Piloto.

Tabela 34 – STPA-Sec: Ações de Controle Inseguras TWR

Ação de Controle	**Tipo UCA**	**Descrição UCA**	**Perigo**
Autorizar pouso ou decolagem	Não executada quando requisitada	UCA-1: ATCO não visualiza dados sobre espaço aéreo.	H-3, H-5
Autorizar pouso ou decolagem	Executada incorretamente	UCA-2: ATCO visualiza dados incorretos sobre o espaço aéreo.	H-3
Autorizar pouso ou decolagem	Executada quando não disparada pelo ATCO	UCA-3: Piloto recebe autorização que não partiu do ATCO	H-1
Verificar Manutenção da Aeronave	Não executada quando requisitada	UCA-4: Aeronave voando sem manutenção.	H-2
Preparar Aeronave	Executada incorretamente	UCA-5: Aeronave em modo incorreto para pouso ou decolagem.	H-2

A partir das ações de controle inseguras são montados os cenários de perda, *Loss Scenarios* (LS) mostrado na Tabela 35. Foram elabora um LS para cada UCA.

Tabela 35 – STPA-Sec: Cenários de Perda TWR

UCA	**Fatores para a UCA**	**Cenários de perda para a UCA-1**
UCA-1	Erros de comunicação, ataques *denial of service*,	LS-1: Dispositivos conectados na rede aumentam o nível de tráfego de dados congestionando as comunicações entre o sistema da torre e os sistemas meteorológico e radar.
UCA-2	Parâmetros incorretos, ataque *tampering*,	LS-2: os valores disponibilizados no console do sistema torre não correspondem aos detectados pelo sistema meteorológico e radar.
UCA-3	Ataque *Spoofing*	LS-3: Dispositivo não legítimo reproduz pacotes de autorização.
UCA-4	Falha Procedimento	LS-4: Aeronave com defeito no sistema mecânicos e de comando e controle.
UCA-5	Feedback incorreto, ataque *tampering*,	LS-5: Aeronave repassa estado incorreto para o piloto.

Com os cenários de perda levantados, novos requisitos podem ser incluídos. Para o exemplo da LS-1, um requisito possa ser a necessidade de dimensionar a rede corretamente e dispor de caminhos alternativos. Além de outro requisito de *security*, que poderiam ser inclusão de filtros de endereços lógicos e endereços físicos para evitar que qualquer dispositivo se conecte a rede, reduzindo assim as chances de ataques de negação de serviços.

Para o cenário de perda LS-2, a descrição indica que há uma perda de integridade dos dados, que levanta a preocupação dos dados sofrerem uma alteração intencional do caminho. Como requisito, pode ser sugerido a inclusão de um caminho redundante.

O cenário de perda LS-3 representa a possibilidade de um ataque *spoofing* e como requisito pode-se adicionar a necessidade de confirmação em outro canal (VHF) entre o piloto e o controlador. Este cenário induz a inclusão de medidas de criptografia e autenticação. Porém, conflita com o requisito de transmissões de mensagens de autorização para todas as aeronaves da área.

Para o cenário de perda LS-4, que representa a possibilidade de uma aeronave operar sem estar com a manutenção corretamente verificada, pode-se incluir alertas obrigatórios na cabine de comando.

Para o cenário de perda LS-5, que representa a possibilidade do sistema da aeronave informar um estado incorreto dos equipamentos como posição de *flaps* e reversos. Está sendo considerado a possibilidade do sistema estar infectado com um tipo de vírus que induz o piloto ao erro, como acreditar que configurou a aeronave corretamente. Pode-se sugerir a inclusão de verificações de integridade como varredura com antivírus de forma periódica.

7 RESULTADOS E DISCUSSÕES

Nesta seção são apresentadas as comparações das características das metodologias, dos resultados obtidos nos estudos de caso e associada com propriedades desejáveis para uma metodologia ideal. Além disso, é proposto uma estratégia para uma condução de análise integrada de segurança crítica e segurança cibernética em sistemas ciber físicos.

7.1 COMPARATIVO DAS CARATERÍSTICAS DAS METODOLOGIAS

As metodologias BDMP, S-Cube, CHASSIS, GTST-MLD e STPA-Sec são comparadas com informações extraídas de suas descrições na revisão bibliográfica. Os atributos referentes as entradas, ferramentas de simulação, saídas e tipo e são mostradas na Tabela 36.

Tabela 36 – Quadro comparativo das características das metodologias

Metodologia	**Entrada**	**Ferramentas**	**Saídas**	**Tipo**
BDMP	Árvore com gatilhos e parâmetros quantitativos das folhas.	Risk Spectrum, BDMPathfinder	Conjuntos de Corte e probabilidade de ocorrência do evento topo.	Qualitativo e Quantitativo
S-CUBE	Arquitetura do sistema e parâmetros quantitativos.	Risk Spectrum, VisualFigaro	Cenários de Possíveis Ataques e Falhas	Qualitativo e Quantitativo
CHASSIS	Diagramas UML e Tabelas.	Qualquer ferramenta UML	Requisitos de segurança crítica e de segurança cibernética.	Qualitativo
GTST-MLD	Árvores de Funções e Componentes.	Qualquer ferramenta para diagramas lógicos.	Conexões de Funções e Componentes. Probabilidade da função falhar.	Qualitativo e Quantitativo
STPA-Sec	Perdas, Perigos, Restrições de Alto Nível e Estruturas de Controle.	Visual Pro, STAMP Workbench, SafetyHAT e outras.	Cenários de Perdas, Requisitos de Controle e Contramedidas.	Qualitativo

O tipo de entrada varia em cada metodologia. Para uma análise em BDMP, deve-se construir árvore de falhas que permitem a inclusão dos gatilhos e elementos que representam proteções e ameaças cibernéticas. As árvores BDMP também aceitam parâmetros quantitativos como taxa de falhas em cada folha. Utiliza-se uma ferramenta

dedicada Risk Spectrum (RISK SPECTRUM AB, 2023) para as simulações. Outra ferramenta BDMPathfinder também está disponível (CZEKSTER; MORISSET, 2021). Como saídas qualitativas, obtém se os conjuntos de cortes que levam a ocorrência do evento topo e, para saídas quantitativas, a probabilidade de ocorrência do evento topo em um determinado valor tempo.

Para uma análise S-cube, desenha-se a arquitetura do sistema e atribui-se as probabilidades de falhas e ataques em cada elemento. No S-cube não é possível representar elementos externos oriundo do ambiente como probabilidades de ondas marítimas ou de mal tempo. O escopo da análise se restringe a arquitetura do sistema e ainda é necessário todos elementos do sistema tenham sido representados por uma classe no S-Cube. Por exemplo, o S-Cube não possui classes que representam uma redundância em *standby*, mas possui uma classe que representa um elemento votador com duas ou três entradas. Utiliza-se uma ferramenta dedicada Risk Spectrum (RISK SPECTRUM AB, 2023) para as simulações e pode-se customizar a base de conhecimento com qualquer compilador de linguagem Figaro (KHAN et al., 2021). Como saídas qualitativas, o S-Cube fornece possíveis cenários sequencias para ocorrências de ataques e falhas que levem ao comprometimento da função do sistema. Além de fornecer as probabilidades dos cenários.

Na metodologia CHASSIS, elabora-se casos de uso representando-os em diagramas UML e em formulários com perguntas chaves. Pode ser usado em qualquer ferramenta UML. Com a realização de todas as etapas do CHASSIS, os resultados de saída, que são apenas qualitativos, correspondem aos requisitos de segurança crítica e segurança cibernética de forma refinada. No CHASSIS considera-se que o atendimento a todos os requisitos levantados é suficiente para garantir a segurança do sistema.

Na condução do GTST-MLD, são construídas a árvore com as funções objetivos e a árvore com os componentes, ambas são conectadas. Os componentes digitais ainda devem ser associados aos tipos de ameaças cibernéticas que podem estar suscetíveis. Pode ser usado qualquer ferramenta visual que permita conexões e portas lógicas. Como resultado qualitativo, a construção de ambas as árvores permite uma visualização de um caminho de um ameaça cibernética até a função objetivo. Além disso, como nenhuma função pode ficar sem estar associada a um componente, é possível identificar ausência de sistemas na árvore ST. Para um resultado quantitativo, deve-se inserir os

valores de probabilidades do componente falhar ou ser atacado, com base em suas conexões, aplicando os cálculos de probabilidade de união ou interseção de eventos sucessivamente até chegar na função topo. Esses cálculos podem ser realizados manualmente ou utilizando ferramentas como o MATLAB.

No STPA-Sec são definidas como entradas as perdas, perigos e restrições do sistema. Em seguida, elaboram-se as estruturas de controle. Há diversas ferramentas disponíveis entre comerciais como VisualPro e gratuitas como STAMP Workbench (MIT PSASS GROUP, 2023). No desenvolvimento da análise também são elaboradas responsabilidades e restrições dos controladores e ações de controle inseguras. As saídas correspondem a cenários de perdas que podem indicar novos requisitos e contramedidas. A metodologia fornece apenas resultados qualitativos.

Com relação ao modo de condução, na Tabela 37 são mostrados o raciocínio de construção do modelo em cada metodologia. No BMDP, o analista é guiado pelas falhas possíveis ao sistema, similar ao método FTA, porém com a possibilidade de incluir os gatilhos de uma folha para qualquer outra na árvore. No S-cube, o foco é construção correta da arquitetura do sistema interligando os módulos. No CHASSIS, inicia-se com o processo de casos de uso, onde se define os atores do sistema, suas funções e extensões. Em seguida, insere-se as ameaças e possíveis contramedidas. Para o GTST-MLD estabelece as funções do sistema, que podem ser decompostas em subfunções na árvore GT. Em seguida, para cada função deve-se pensar “como” pode ser atendida e este questionamento traz repostas para montar a árvore ST. Por último, no STPA-Sec devem ser levantadas as perdas, perigos, restrições do sistema, e desenhadas as estruturas de controle que devem conter as ações de controle e receber os *feedbacks*.

Tabela 37 – Quadro comparativo do raciocínio de análise

Metodologia	**Enfoque**	**Raciocínio de construção**
BDMP	Falha	Similar a construção de árvore de falhas, porém, permite a possibilidade de adicionar os recursos de gatilhos e de elementos que representam a possibilidade de defesa e ataques cibernéticos.
S-CUBE	Arquitetura	Elabora-se a arquitetura do sistema com os elementos selecionados de acordo com a classe correspondente no S-Cube.
CHASSIS	Casos de Uso	A partir da elaboração dos casos de uso, são feitos os levantamentos de

		ameaças, erros de uso, e contramedidas, requisitos.
GTST-MLD	Objetivos	São estabelecidas funções do sistema que podem ser decompostas em subfunções. Cada função ou subfunção deve estar conectada aos sistemas e componentes necessários para seu atendimento.
STPA-Sec	Controle	São levantadas as perdas, os perigos e as restrições do sistema. Em seguida desenhados as estruturas de controle.

7.2 RESULTADOS NO ESTUDO DE CASO 1 - SISTEMA AH

São discutidos os resultados obtidos pela aplicação de cada metodologia ao estudo de caso 1, que se refere ao Sistema *Anti-Heeling* detalhado no capítulo 5. As metodologias são avaliadas em sua eficiência em identificar os problemas de segurança cibernética *jamming* e *spoofing* e o conflito da relação entre implantações de criptografia e autenticação com a redução de latência.

A Tabela 38 mostra as saídas obtidas com cada metodologia para a identificação da possibilidade de ocorrência de *jamming*. Com exceção do S-Cube, para uma correta identificação é necessário que já se tenha conhecimento da possibilidade desse ataque. As metodologias BDMP, CHASSIS, GTST-MLD e STPA-Sec são dependentes da experiência da equipe de análise em segurança cibernética. Já o S-Cube foi capaz de apontar o cenário de ataque sem ter sido previamente modelado na entrada.

Tabela 38 – Saídas das análises de ataque *jamming* no sistema AH

Metodologia	**Tipo de Saídas**	**Deteção do Jamming**
BDMP	Grupos de Corte	Folhas “Anti-Jamming”, “Jamming” e “Variação de Carga Acima dos Limites”
S-CUBE	Cenários de Ataques e Falhas	Cenário com a etapa “jamming_attack (Rede_Comunicacao)” implica em indisponibilidade do Sistema AH.
CHASSIS	Requisitos	Prover proteções contra ataques *Jamming*
GTST-MLD	Conexões de Funções e Componentes	Mapeamento dos dispositivos suscetíveis a um ataque *Jamming* e adição da função “Evitar exposição a variação de carga”.
STPA-Sec	Cenários de Perda e Fatores Casuais	- Sistema AH não ajusta tanque de lastro quando o navio está inclinado. - Excesso de tráfego na rede impede o recebimento e envio de pacotes

Continuando na Tabela 38, no BDMP foi obtido o grupo de corte com as folhas “Anti-Jamming”, “Jamming” e “Variação de Carga acima dos limites” que seriam necessários para a ocorrência do evento topo, “Navio Tombado”. A folha “Anti-Jamming” representa uma proteção contra o ataque e sua ocorrência está ligada a ineficiência da proteção ou uma não implementação da mesma. A folha “Jamming” representa a ocorrência de um ataque bem sucedido que deixaria o sistema AH indisponível. Por último, ainda seria necessário ocorrer uma “Variação de Carga Acima dos Limites” para que o navio tombe. Este evento está relacionado com a possibilidade de ocorrer uma variação do centro de massa do navio seja por onda marítima, por deslocamentos de cargas internas ou até mesmo por manipulação.

O resultado pelo S-cube apontou que um cenário com apenas uma etapa de ataque *jamming* na rede de comunicação deixaria o Sistema AH indisponível. Esse cenário foi obtido porque todos os componentes do sistema estão conectados em uma única rede. O elemento que representa a rede de comunicação é da classe “network_zone” que possui susceptibilidade a este tipo de ataque.

No CHASSIS, a identificação do ataque *jamming* inicia no momento em que se identifica no diagrama de erros de uso (MUSD), na Figura 27, a ameaça “Inundar a rede com pacotes” que interfere no caso de uso “Ler Pacote Enviado pelo Sensor” do ator “Controlador”. Em seguida, na Tabela 13, onde foram levantadas as causas, aponta-se a possibilidade de *jamming* e inclui-se a contramedida “Incluir Medidas Anti-Jamming”.

Os resultados obtidos no GTST-MLD permitem a visualização de que os componentes sensor, atuador e controlador estão suscetíveis a ataques “*Denial of Service*”, que corresponde ao tipo de ataque *jamming* no modelo STRIDE (Figura 31). Os componentes afetam a função “Controlar nível de água nos tanques de lastro” (Figura 30). Ainda foi adicionado na árvore GT a função “Evitar exposição a variação de carga”, que obriga a definição de novos componentes como “Alerta Meteorológico”.

No STPA-Sec, a identificação da possibilidade de ataque *jamming* iniciou no levantamento do perigo H-1 da Tabela 17, “Sistema AH Indisponível”. A definição do perigo H-1 levou às definições SC-1, “... operar continuamente ...”, e SC-3, “...implementar proteções cibernética”, mostrados na Tabela 18. Em seguida, é levantada a ação de controle insegura UCA-1, “Sistema AH não ajusta tanque de lastro

quando o navio está inclinado", mostrada na Tabela 20. Por fim, é identificado a possibilidade do cenário de perda S-4, "excesso de tráfego na rede impede o recebimento e envio de pacotes", mostrado na Tabela 22.

Em todas as metodologias aplicadas ao sistema *Anti-Heeling* foram apontadas as possibilidades de ataques tipo *jamming.* Ainda que as metodologias BDMP, CHASSIS, GTST-MLD e STPA-Sec sejam fortemente dependentes da experiência dos analistas na definição das entradas, houve identificações guiadas intuitivamente. No S-Cube a identificação foi automática conforme proposta do modelo. Quanto a capacidade de cobertura e rastreabilidade o STPA-Sec forneceu mais detalhes sobre possíveis causas dos cenários de perda e forneceu uma melhor visualização na relação entre um problema de baixo nível até o alto nível. Quanto a capacidade de análise quantitativa, o BDMP apresenta ser a melhor alternativa por permitir a entrada de parâmetro quantitativos e simular dentro da própria ferramenta, porém, tais parâmetros de entradas dependem de técnicas de estimativas como as mencionadas na seção 2.5.

Para a identificação do ataque *Spoofing* as saídas são mostradas na Tabela 39. No S-cube os passos do cenário de ataque foram gerados automaticamente na simulação. Nas demais metodologias BDMP, CHASSIS, GTST-MLD e STPA-Sec foram necessários que em algum momento, o analista informasse as ameaças.

Tabela 39 – Saídas das análises de ataque *spoofing* no sistema AH

Metodologia	**Tipo de Saídas**	**Deteção do Spoofing**
BDMP	Grupos de Corte	Folhas "Autenticação", "Criptografia", "Spoofing" e "Opr Maliciosa nos Tanques"
S-CUBE	Cenários de Ataques e Falhas	Etapas para um cenário de ataque *spoofing*: 1 - access (Rede_Comunicacao) 2 - attacker_scan_network (Rede_Comunicacao) 3 - bypass_autentification (Rede_Comunicacao) 4 - attacker_stablish_connection (Atuador_SW) 5 - send_false_instructions_to_actuator ()
CHASSIS	Requisitos	O sistema deve ter protocolos de autenticação. O sistema deve implementar criptografia.
GTST-MLD	Conexões de Funções e Componentes	Mapeamento dos dispositivos suscetíveis a um ataque *Spoofing* e relacionados a função "Evitar manipulação no controle dos tanques de lastro".
STPA-Sec	Cenários de Perda e Fatores Casuais	O Sistema AH altera nível do tanque de lastro mesmo com o navio balanceado O controlador calcula novo nível para os tanques equivocadamente

No BDMP, um spoofing ocorrido no sensor, atuador ou controlador ativa um gatilho para a folha “Opr Maliciosa Tanques”, que ativa outro gatilho, que incide 100% de chances de ocorrência do evento “Variação de Cargas Acima dos limites”. Foram representadas as medidas contra spoofing que são autenticação e criptografia. Portanto, para ocorrer o evento topo, “Navio tombado”, seriam necessários uma ineficiência ou não implementação de ambas as medidas de Autenticação e Criptografia. Em seguida, um ataque spoofing com sucesso manipule os níveis de água dos tanques de lastro (Tabela 39).

O cenário de ataque do S-Cube na Tabela 39 representa os passos possíveis de um atacante para executar uma ação *spoofing* bem sucedida. O ataque inicia com um acesso a rede de comunicação, em seguida, uma operação de varredura, continuando com uma autenticação burlada. Nas últimas etapas há um estabelecimento de conexão com um dos componentes, no caso, com o atuador, e há o envio de uma instrução falsa partindo do dispositivo impostor. As etapas do cenário auxiliam na definição de barreiras para evitar a progressão de um ataque.

No CHASSIS, a identificação do ataque *spoofing* inicia no momento em que se identifica no diagrama de erros de uso (MUSD), na Figura 27, a ameaça “Enviar Falsa Ação de Controle” que interfere no caso de uso “Operar Tanques de Lastro” do ator “Atuador”. Em seguida, na Tabela 13, foi levantada a causa, a consequência e adicionada a recomendação “Implementar autenticação”.

No GTST-MLD permitem a visualização de que os componentes sensor, atuador e controlador estão suscetíveis a ataques “*Spoofing*” (Figura 31). Os componentes afetam a função “Evitar manipulação no controle dos tanques de lastro” (Figura 30). Pela conexão nota-se que o ataque é suficiente para comprometer a função topo “Prevenir Tombamento do Navio”.

No STPA-Sec, a identificação da possibilidade de ataque *spoofing* iniciou no levantamento do perigo H-3 da Tabela 17, “Sistema AH Manipulado”. A definição do perigo H-3 levou a definição do SC-3, “...implementar proteções cibernética”, mostrados na Tabela 18. Em seguida, é levantada a ação de controle insegura UCA-2, “Sistema AH altera nível do tanque de lastro mesmo com o navio balanceado”, mostrada na Tabela 20. Por fim, é identificado a possibilidade do cenário de perda, “o controlador

calcula novo nível para os tanques equivocadamente” que pode estar relacionado ao fator casual de recebimento de *feedback* incorreto.

Todas as metodologias identificaram as possibilidades de ataques tipo *spoofing.* O S-Cube gerou um cenário muito útil e direto com algumas etapas do ataque, mas o STPA-Sec ainda forneceu uma cobertura mais completa das causas.

Quanto a capacidade de identificar o conflito de criptografia e autenticação com a redução da latência, o STPA-Sec e o BDMP apresentaram-se como melhores. O S-Cube não foi possível fornecer esse resultado por não ter outras representações. No BDMP, o conflito foi apenas possível de se identificar na análise quantitativa, conforme Tabela 7 e Tabela 8. No CHASSIS foi levantado o requisito “sistema deve implementar criptografia desde que não consuma largura de banda excessiva” que pode representar uma compreensão de que a criptografia aumenta a latência e contribui para as chances de indisponibilidade das comunicações. Na análise pelo GTST-MLD não foi possível de identificar essa relação de conflito de forma direta, porém, as saídas do modelo indicam necessidade de incluir criptografia e autenticação e prover meios alternativos para caso de indisponibilidade da rede. Para o STPA-Sec são levantadas as restrições do controlador “O computador de bordo deve certificar que está recebendo o valor de inclinação do sensor legítimo” e “O computador de bordo deve operar em rede de comunicação de baixa latência”, ambas mostradas na Tabela 21. Essas restrições indicam uma reformulação da arquitetura para que, além da implementação da autenticação, haja um redimensionamento da rede de comunicação para suportar a largura de banda.

7.3 RESULTADOS NO ESTUDO DE CASO 2 - SISTEMA TWR

São comparados os resultados obtidos pela aplicação de cada metodologia ao estudo de caso 2, Sistema de Torre de Controle de Aeródromo (TWR), detalhado no capítulo 6. As metodologias são avaliadas em sua eficiência em identificar os problemas de emissão de autorização indevida. Este problema pode ser causado por erros de detecções ou por transmissão de falsa autorização.

A Tabela 40 mostra as saídas obtidas pela aplicação de cada metodologia na identificação de problemas com emissão de autorização indevida para pouso ou decolagem decorrentes de intervenções maliciosas. Os ataques cibernéticos possíveis seriam uma reprodução de uma falsa autorização ou uma manipulação de *feedbacks* que induz o controlador ao erro.

Tabela 40 – Saídas das análises de autorização indevida no TWR

Metodologia	**Tipo de Saídas**	**Resultados**
BDMP	Grupos de Corte	Reprodução do pacote: “Spoofing” seguido por “Objeto na rota de colisão” não desviado. Manipulação do feedback: Access Control List” violado, seguido por “Intrusion Prevent System”, e seguido por “Tampering” com autorização indevida e seguido por “Objeto na rota de colisão” não desviado.
S-CUBE	Cenários de Ataques e Falhas	Reprodução do pacote: 1. Access (Data_Link_Aeronave) 2. attacker_stablish_connection (Console_Aeronave) 3. send_false_data (Console_Aeronave) Falsificação de dados do Sistema TWR: 1. access_network (Rede_Sistema_Torre) 2. attacker_scan_network (Rede_Sistema_Torre) 3. bypass_autentification (Servidor_TWR) 4. attacker_stablish_connection (Servidor_TWR) 5. exploit_Vuln_Priv_Escalator (Servidor_TWR) 6 Exploit_Vuln_Integrity_loss (Servidor_TWR) Falsificação de dados do Radar: 1. access (Rede_Interna_Radar) 2. attacker_scan_network (Rede_Interna_Radar) 3. attacker_stablish_connection (CTRL_RADAR_SW) 4. send_false_feedback (CTRL_RADAR_SW)
CHASSIS	Requisitos	- A operação de autorização deve confirmada também via Rádio VHF; - O sistema TWR deve incluir redundância na visualização, via outra console obtendo dados por canais alternativos. - O sistema TWR deve adotar protocolos de autenticação aos demais sistemas consultados.
GTST-MLD	Conexões de Funções e Componentes	Mapeamento dos componentes necessários para a função “Gerenciar autorizações” e quais estão suscetíveis ao ataque cibernético de *Tampering* (adulteração): - Sistema Meteorológico -Radar -Console -Rede TWR -VHF -Enlace de dados com a aeronave
STPA-Sec	Cenários de Perda e fatores casuais	-Dispositivo não legítimo reproduz pacotes de autorização. Um fator casual poderia ser a leitura de pacotes em claro por um atacante e montagem de pacote falso.

		- Valores disponibilizados no console do sistema torre não correspondem aos detectados. Um fator casual poderia ser a captura e bloqueio do pacote legítimo em um roteador no caminho entre o sistema Radar e a Torre. Remontagem do pacote com dado adulterado.

Analisando as saídas do BDMP na Tabela 40, temos como um primeiro grupo de corte a ocorrência de ataque "*spoofing"* seguido do evento "Objeto na rota de colisão" não desviado. Esse grupo de corte representa a possibilidade de reprodução de mensagens de autorização de pouso ou decolagem. O segundo grupo de corte corresponde aos eventos "Access Control List" violado, seguido por "Intrusion Prevent System", em seguida "Tampering". Com isso se teria o *feedback* manipulado, que induz a uma autorização indevida. Por último o evento "Objeto na rota de colisão" não desviado. Este segundo grupo representa a possibilidade do controlador de tráfego aéreo emitir uma autorização com base em dados adulterados no radar ou no sistema meteorológico.

Os resultados fornecidos pelo S-cube, mostrados na Tabela 40, representam a reprodução maliciosa de uma mensagem de autorização. As etapas de "Access (Data_Link_Aeronave)", "attacker_stablish_connection (Console_Aeronave)" e "send_false_data (Console_Aeronave)" indicam que se um atacante obtiver acesso ao enlace de dados com a aeronave e estabeleça uma conexão, então estará apto a enviar qualquer tipo de mensagem para a aeronave. Outros cenários com sequências de ataques correspondem às possibilidades de falsificação de dados no sistema da Torre e no Radar, ou seja, uma manipulação nos *feedbacks* com o objetivo de induzir o controlador a conceder uma autorização com base em dados incoerentes.

Continuando na Tabela 40, no CHASSIS, a identificação da possibilidade de uma autorização indevida se inicia no momento em que são levantadas as ameaças de "Alterar Dados de Feedback" e "Reproduzir MSG de Autorização" ilustradas na Figura 38. Com os desdobramentos dos processos chega-se à especificação dos requisitos de inclusão de autenticação, criptografia e redundância de visualização como contramedidas a possibilidade de alteração de *feedbacks*. Outro requisito levantando refere-se à utilização de confirmação da operação por outro canal, no caso, via Rádio VHF.

No GTST-MLD são visualizadas as conexões dos componentes necessários para realizar a função “Gerenciar autorizações”. São levantados quais componentes estão suscetíveis aos ataques cibernéticos. Foram apontados que os componentes na árvore ST, “Sistema Meteorológico”, “Radar”, “Console”, “Rede TWR” e “Enlace de dados com a aeronave”, mostrados na Figura 40 e Figura 41, estão sujeitos a ataques de *spoofing* e *tampering*. Pode ser visualizado ainda que o “VHF” em conjunto com “Datalink Aeronave” provê uma redundância na comunicação entre controlador de tráfego e piloto. Essa redundância pode ser uma contramedida para confirmação de uma autorização de pouso ou decolagem.

No STPA-Sec, a identificação da possibilidade de ocorrência de autorização para pouso ou decolagem indevida se inicia com o levantamento dos perigos H-1,” Aeronave decolando ou pousando sem autorização do ATCO”, e H-3, “Obstáculo em distância da aeronave inferior ao limite mínimo” mostrados na Tabela 31. Em seguida, são levantadas as restrições de sistemas SC-1, “As autorizações para pouso e decolagem devem ser verificadas via dados e via voz entre controlador e piloto”, e SC-3, “O controlador deve consultar os sistemas para verificar obstáculos no espaço aéreo, no pátio, mal tempo e conceder somente uma única autorização por vez”, mostrados na Tabela 32. As restrições SC-1 e SC-3 levam à construção da estrutura de controle mostrada na Figura 42. A estrutura de controle se desdobra no levantamento das ações de controle inseguras UCA-2, “ATCO visualiza dados incorretos sobre o espaço aéreo”, e UCA-3, “Piloto recebe autorização que não partiu do ATCO” (Tabela 34). Ao final da análise é identificado um cenário de perda referente aos valores disponibilizados no console do sistema torre, que não corresponderam aos detectados pelo sistema meteorológico. Para este cenário, um possível fator casual poderia ser, a captura e bloqueio de transmissão de um pacote legítimo em um roteador no caminho entre o sistema Radar e a Torre. Outro cenário de perda trata-se da possibilidade de existir um dispositivo reproduzindo mensagens de autorização. Este último, pode estar relacionado ao fator casual de um atacante ler os pacotes transmitidos sem criptografia. A partir dessa captura, o atacante consegue montar outro pacote e realizar uma transmissão falsa.

Em todas as metodologias foi possível representar problemas referentes às possibilidades de envio de mensagens de falsa autorização para pouso ou decolagem e

de alteração dos *feedbacks*. As abordagens STPA-Sec e S-Cube apresentaram-se como melhores alternativas. O STPA-Sec por abranger mais possibilidades de causas para o problema, direcionando o analista a implantar barreiras. O S-Cube por fornecer cenários detalhados com passos de possíveis ataques.

7.4 DESEMPENHO EM CRITÉRIOS DESEJÁVEIS

As metodologias BDMP, S-Cube, CHASSIS, GTST-MLD e STPA-Sec são comparadas com critérios levantados (Tabela 3) para que uma metodologia seja considerada ideal nas análises de segurança crítica e segurança cibernética em sistemas ciber físicos.

O primeiro critério desejável é que a metodologia seja "**integrada"**, que significa ser capaz de representar elementos de segurança crítica, segurança cibernética e as relações entre ambas as áreas. Essas relações foram mostradas na seção 2.2 que correspondem a dependência condicional, reforço mútuo, antagonismo e independência. Todos as metodologias selecionadas satisfazem este critério. Porém, a análise quantitativa pelo BDMP na seção 5.4 apresentou de forma mais objetiva uma identificação de antagonismo entre as áreas. Para análise qualitativa o STPA-Sec mostra-se mais completo ao permitir visualização dos requisitos conflitantes como baixa latência e implementação de criptografia e autenticação mostrados na seção 5.8.

O segundo critério desejável é que a metodologia forneça resultados "**qualitativos e quantitativos"**. As metodologias BDMP, S-Cube e GTST-MLD atendem a estes critérios. O BDMP mostrou-se mais eficiente por gerar ambos tipos de resultados dentro da mesma ferramenta conforme apresentado na seção 5.4.

O terceiro critério é produzir **"resultados automáticos",** ou seja, ser capaz de identificar problemas de segurança crítica e segurança cibernética mesmo sem ter sido previamente definidos nas entradas. O objetivo é que se tenha menor dependência da experiência dos analistas. Para este critério, o S-Cube é o mais eficiente conforme resultados apresentados nas seções 5.5 e 6.3. Este justamente é o propósito da criação do S-Cube, mas ainda existem muitas limitações, como por exemplo, não contempla representações de elementos em modo de redundância *standby*.

O quarto critério é ser **“gerenciável e legível”,** que significa prover melhores meios para a organização do modelo e facilite a compreensão dos analistas e outras pessoas envolvidas. Neste critério o STPA-Sec mostrou-se mais eficiente para gerenciamento e com bom nível de legibilidade. A compreensão do BDMP não é trivial para não especialistas. O S-Cube não contempla documentação. O CHASSIS embora forneça uma documentação bem legível mostrando diagramas, tabelas e requisitos, não se apresenta como uma alternativa gerenciável devido ao alto volume de informações geradas em seus processos de forma manual. O GTST-MLD torna-se difícil a compreensão quando se aumenta a quantidade de elementos em um sistema.

O quinto critério, “**ampla aplicação**”, refere-se à capacidade da metodologia ser utilizada em todas as fases de um sistema (concepção, desenvolvimento e operação). Também se refere a capacidade de ser aplicada em diversas áreas como marítima, aviação, ferroviária, nuclear e outras. Todas as metodologias podem ser aplicadas em qualquer área que envolvam sistemas ciber físicos conforme observados nos estudos de caso nos capítulos 5 e 6. Porém, quanto à fase do sistema, na análise do BDMP e GTST-MLD já se exige um conhecimento prévio da arquitetura do sistema, sendo mais voltada para as fases de operação. O CHASSIS concentra-se na fase de concepção. O S-Cube concentra na fase de desenvolvimento e operação. O STPA-Sec abrange todas as fases.

O sexto critério agrupa as capacidades da metodologia ser “**flexível, modular e escalável**”. Flexível por permitir uma reanálise facilitada quando sofrer alterações no projeto. Modular por permitir a divisão em unidades menores. Escalável por permitir um crescimento do projeto sem invalidar completamente análises anteriores ou ter que refazer toda análise dispendiosamente. Para este grupo de critérios o STPA-Sec mostra-se mais eficiente pela possibilidade de construir estruturas de controles hierárquicas e com interações em paralelos conforme mostrado na seção 6.6. O S-Cube atenderia somente ao critério de ser flexível, uma vez que a alteração da arquitetura não exigiria uma reanálise onerosa.

O sétimo critério é permitir uma análise “**rastreável**”, que possibilita conectar problemas oriundos de elementos de baixo nível com o alto nível do sistema. O STPA-Sec mostra-se como melhor alternativa ao fornecer uma fácil visualização das

interligações de fatores casuais (baixo nível) com as perdas, perigos e restrições (alto nível). Estas constatações podem ser observadas nas seções 5.8 e 6.6.

O oitavo e último critério levantado, trata-se das disponibilidades de "**Ferramentas dedicadas**" para realização das análises pelas metodologias. Neste critério, o STPA-Sec mostra-se amplamente melhor por existir diversas alternativas de ambientes de simulação comerciais e gratuitas. As ferramentas para cada metodologia foram mostradas na Tabela 36. O CHASSIS e o GTST-MLD não possuem ferramentas próprias. O GTST-MLD necessita que as análises quantitativas sejam realizadas em outros ambientes matemáticos como o MATLAB. O BDMP e o S-Cube são dependentes de um único ambiente, o Risk Spectrum AB que é comercial.

A Tabela 41 resume qual a metodologia melhor atende cada critério levantado de uma metodologia ideal. A metodologia STPA-Sec se sobressaiu em seis dos oitos critérios levantados. O BDMP surge como melhor alternativa em dois critérios. O S-Cube ganha no critério de resultados automáticos. As metodologias CHASSIS e GTST-MLD não foram apontadas como melhor em nenhum critério.

Tabela 41 – Metodologia com melhor desempenho por critério

	Critérios Desejáveis	**Melhor metodologia**	**Justificativas**
1	Integração *safety* e *security*	BDMP e STPA-Sec	Identificação de conflitos nas seções 5.4 e 5.8
2	Qualitativos e quantitativos	BDMP	Produção de ambos resultados em uma única simulação (seção 5.4).
3	Resultados automáticos	S-Cube	Cenários de ataques e falhas mostrados nas seções 5.5 e 6.3.
4	Gerenciável e Legível	STPA-Sec	Organização modular e construção intuitiva.
5	Ampla aplicação	STPA-Sec	Abrange as fases de concepção, desenvolvimento e operação do sistema.
6	Flexível, Escalável e Modular	STPA-Sec	Construção pautada em unidades de estruturas de controles (seção 6.6).
7	Rastreável	STPA-Sec	Interliga problemas de baixo nível ao alto nível do sistema (seções 5.8 e 6.6).
8	Ferramentas dedicadas	STPA-Sec	Disponibilidade de diversas ferramentas (Tabela 36).

7.5 PROPOSTA DE ESTRATÉGIA DE ANÁLISE

Com o intuito de auxiliar na condução de uma análise integrada de segurança crítica e segurança cibernética em sistemas ciber físicos, é apresentado uma proposta de estratégia de seleção de metodologia. A proposta foi inspirada na técnica de identificação de perigos de Hardy (2012) e fundamentada nos resultados obtidos neste estudo. A parte inicial está mostrada na Figura 43.

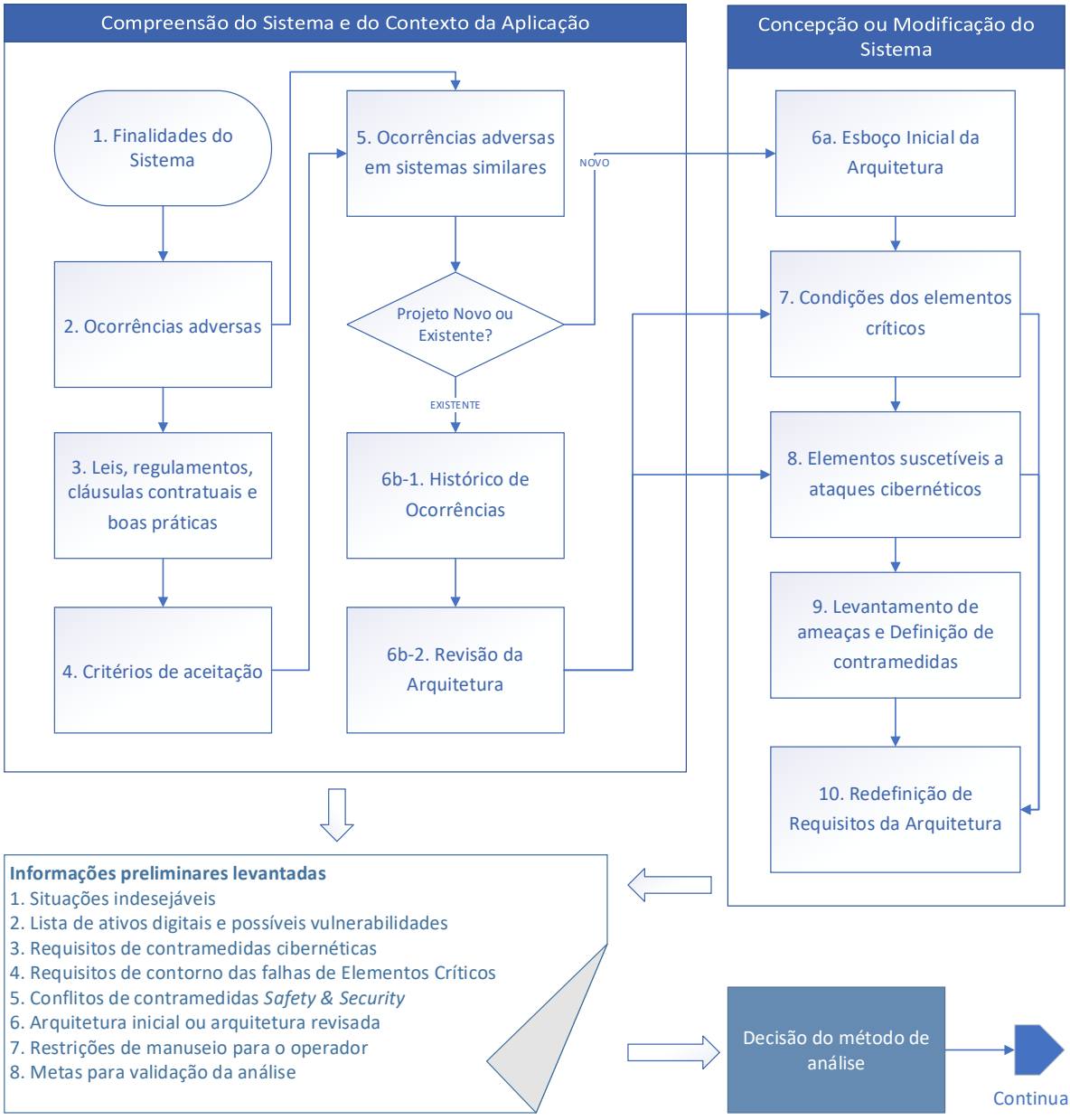

Figura 43 – Estratégia de Condução de Análise Parte 1

As atividades são inicialmente agrupadas em “Compreensão do Sistema e Contexto da aplicação”, onde se deve estudar o sistema, e “Concepção ou Modificação do Sistema”, onde já se elabora ou modifica uma arquitetura para o sistema.

A atividade 1, “**Finalidade do Sistema”**, consiste em entender os objetivos da aplicação, incluindo suas limitações de operação, requisitos iniciais e interfaces com sistemas externos.

A atividade 2, “**Ocorrências Adversas”**, consiste em levantar perigos e ameaças iniciais associadas ao sistema. Devem ser considerados eventos externos e erros de operação. Portanto, nesta atividade o enfoque são as falhas e situações indesejáveis.

A Atividade 3, “**Leis, regulamentos, cláusulas contratuais e boas práticas**”, consiste em extrair requisitos, obrigatórios e opcionais, para relacioná-los ao projeto e análise e adotar melhores práticas de gestão. Algumas normas e guias de boas práticas, mesmo que não tenham força de lei ou contrato, foram elaborados a partir de problemas de acidentes no passado. Portanto, pode ser muito útil adoção de alguns desses documentos.

A Atividade 4, “**Critérios de aceitação**”, com resultados da etapa anterior, são levantados e estabelecidos as métricas na avaliação de segurança crítica e segurança cibernética necessárias para permitir o comissionamento da plataforma crítica.

A atividade 5, “**Ocorrências adversas em sistemas similares**”, consiste em analisar acidentes e incidentes que aconteceram em plataformas críticas similares. Por exemplo, avaliar o ataque STUXNET (LANGNER, 2011) que ocorreu em usina nuclear e verificar se o novo sistema pode estar suscetível ao mesmo ataque.

A Atividades 6, para um sistema novo, elabora-se um esboço da arquitetura, “**6-a. Esboço Inicial da Arquitetura**”. Caso a análise seja em um sistema já existente e em operação, realiza-se uma análise no histórico de incidentes da aplicação, “**6-b-1. Histórico de Ocorrências**”. Em seguida, realiza-se uma revisão da arquitetura com inclusão de modificações, ““**6-b-2. Revisão da Arquitetura**””.

A atividade 7, “**Condições dos Elementos Críticos**”, consiste em analisar previamente recursos potencialmente problemáticos. Como por exemplo, a verificação

se o fornecimento de energia elétrica ao sistema é dependente exclusivamente de terceiros ou se existe redundância com gerador de energia, ou baterias com carga suficiente para alimentar o sistema e os periféricos. Outro elemento é a comunicação. Deve-se verificar se existem caminhos alternativos em caso de indisponibilidade de uma rede de comunicação.

Na atividade 8, “**Elementos Suscetíveis a ataques cibernéticos**”, deve ser feito o levantamento dos dispositivos digitais já previamente definidos na elaboração ou revisão da arquitetura.

Na atividade 9, “**Levantamento de perigos, ameaças, contramedidas e conflitos**”, são listadas problemas referentes a falha de componentes, vulnerabilidades cibernéticas e proteções para esses problemas já visualizados. Deve-se incluir também os conflitos entre medidas de segurança crítica e segurança cibernética já previamente conhecidos. Por exemplo, implantação de criptografia *versus* redução de latência, *broadcast* para amplo conhecimento *versus* proteger de interceptadores.

A atividade 10, “**Redefinição dos requisitos da arquitetura**”, consiste em mais uma oportunidade prévia para se definir a estrutura do sistema. Nesta fase já é desejável que os problemas previamente identificados já tenham recebido um tratamento.

Todas as atividades desta estratégia proposta, mostrada na Figura 43, são importantes para definições consistentes das entradas submetidas em qualquer metodologia. Foi constatado neste estudo que uma análise eficiente tende a depender mais do conhecimento e experiência dos analistas do que da metodologia escolhida. Ao final dessas atividades, espera-se obter uma “**relação de informações preliminares levantadas**” (Figura 43), que servirão de subsídios para decisão de escolha de uma ou mais metodologias na análise do sistema alvo.

A continuação da proposta está mostrada na Figura 44. São estabelecidos alguns questionamentos extraídos das características, exigência da análise e critérios desejáveis. Destaca-se que para uma análise quantitativa, a estimação dos parâmetros deve ser realizada de forma externa.

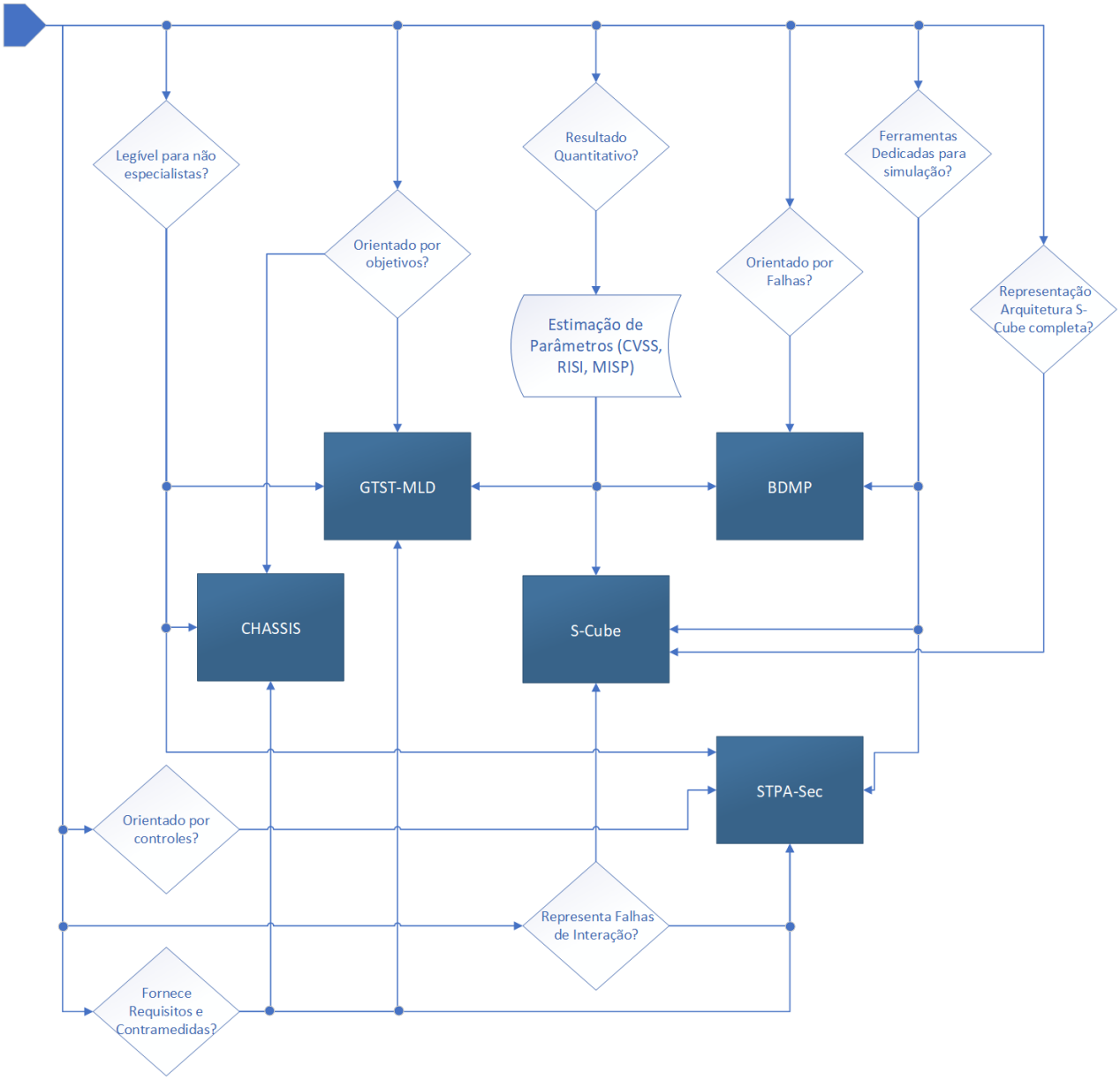

Figura 44 – Estratégia de Condução de Análise Parte 2

Para a escolha da metodologia, não se deve limitar aos questionamentos na Figura 44. As discussões apresentadas em todo o estudo, especialmente na Tabela 36, na Tabela 37 e na Tabela 41, podem ser consideradas na seleção da metodologia. Ou ainda, pode se aplicar mais uma metodologia ao mesmo sistema. Ressalta-se que por se tratar de sistema cujas falhas podem impactar vidas humanas, torna-se prudente adotar elevado rigor nas análises de segurança.

.

8 CONSIDERAÇÕES FINAIS

São apresentadas as conclusões referentes a aplicação das metodologias abordadas nos estudos de caso e sugeridas oportunidades de trabalhos futuros para ampliação das questões discutidas.

8.1 CONCLUSÕES DO ESTUDO

Neste estudo foram apresentadas as comparações das principais metodologias de análise integrada de aspectos de segurança crítica e de segurança cibernética. No desenvolvimento foi evidenciada a relevância da integração de ambos tipos de análises. Essas análises conjuntas justificaram-se pela existência de relações antagônicas entre ambas as áreas, onde uma contramedida aplicada para proteger um lado pode comprometer o outro. Outra razão para a realização de análises integradas também está relacionada à ampliação de ativos de tecnologia da informação nas operações de plataformas críticas, o que aumenta a possibilidade de ataques cibernéticos afetarem a segurança crítica. Para chegar a essas conclusões, foram selecionadas as metodologias de análise integrada BDMP, S-Cube, CHASSIS, GTST-MLD e STPA-Sec. Estas metodologias satisfizeram critérios de seleção como possuírem mais citações e afirmações de eficácia. Em seguida, foram elaborados estudos de casos de sistemas ciber físicos. O primeiro estudo de caso consistiu de um sistema *anti-heeling* de navios, e o segundo consistiu de um sistema de torre de controle de aeródromo. Ambos os estudos de caso foram avaliados pela ótica de cada metodologia. Os resultados obtidos nas simulações foram, de maneira geral, capazes de identificar problemas de segurança crítica, de segurança cibernética e conflitos entre as áreas. Os conflitos identificados poderiam ter passados despercebidos em análises separadas. Nos resultados e discussões, as metodologias foram comparadas entre si com base nas características descritas na revisão da literatura. Também foram comparadas com os resultados obtidos nas análises dos estudos de caso e foram comparadas com o desempenho em critérios desejáveis para uma metodologia ideal. Uma outra conclusão do estudo é que uma análise eficaz ainda depende significativamente da experiência do analista, principalmente para definições de entradas consistentes. Para mitigar essa lacuna, os resultados das comparações serviram de subsídios para a elaboração de uma proposta de estratégia de condução de análise. A proposta está estruturada em sequência de

atividades iniciais que visa coletar informações preliminares que serão usadas como entrada nas metodologias selecionadas.

8.2 PERSPECTIVAS DE TRABALHOS FUTUROS

Uma das grandes limitações para as análises integrada de segurança crítica e segurança cibernética é a capacidade de fornecer resultados quantitativos. As técnicas e fontes de dados para levantamento de parâmetros referentes a probabilidades de ataques cibernéticos são escassas. Muitos dos estudos que usaram parâmetros quantitativos são questionáveis por não considerar, por exemplo, a influência de uma proteção cibernética. Como sugestão de contorno para trabalho futuro, expõe-se a necessidade de desenvolvimento de bases de dados para registros de ocorrências de ataques cibernética que contemple detalhes.

A metodologia S-Cube apresentou um grande atrativo pelo fato de gerar cenários automáticos de ataques e falhas a partir da entrada da arquitetura do sistema. Essa capacidade reduz a dependência da experiência do analista de segurança crítica. Porém, o S-Cube é um projeto embrionário com muitas limitações na representação de elementos. Sugere-se uma ampliação do modelo para cobrir mais aspectos de segurança crítica e cibernética.

As metodologias CHASSIS e GTST-MLD não possuem ferramentas dedicadas. Para representar os modelos, podem ser usadas qualquer ferramenta UML ou visual de diagramas lógicos. O GTST-MLD necessita que a simulação quantitativa seja realizada em outra ferramenta de cálculo. Estes fatos, dificultam o gerenciamento dos modelos. Como oportunidade para trabalhos futuros, são sugeridas o desenvolvimento de ferramentas dedicadas para tais modelos.

Como última sugestão para trabalhos futuros, são estimuladas as pesquisas para concepção de novas metodologias, além da ampliação das já existentes para o atendimento aos critérios desejáveis de uma metodologia ideal. As novas pesquisas podem fazer experimentos com uso de recursos de inteligência artificial.

REFERÊNCIAS

ABUEMERA, E. A.; ELZOUKA, H. A.; SAAD, A. A. Security Framework for Identifying threats in Smart Manufacturing Systems Using STRIDE Approach. **2022 2nd International Conference on Consumer Electronics and Computer Engineering, ICCECE 2022**, p. 605–612, 2022.

AKTOUCHE, S. R. et al. **Towards Reconciling Safety and Security Risk Analysis Processes in Railway Remote Driving**. 2021 5th International Conference on System Reliability and Safety, ICSRS 2021. **Anais**...Institute of Electrical and Electronics Engineers Inc., 2021.

ALMEIDA J. R. JR.; CAMARGO J. B. JR.; CUGNASCA P. S. Safety and Security in Critical Applications and in Information Systems – a Comparative Study. **Ieee Latin America Transactions**, v. 11, n. 4, 2013.

ALOTAIBI, F. M.; VASSILAKIS, V. G. SDN-Based Detection of Self-Propagating Ransomware: The Case of BadRabbit. **IEEE Access**, v. 9, p. 28039–28058, 2021.

ASHRAF, I. et al. A Survey on Cyber Security Threats in IoT-Enabled Maritime Industry. **IEEE Transactions on Intelligent Transportation Systems**, v. 24, n. 2, p. 2677–2690, 1 fev. 2023.

ASPLUND, F. et al. Rapid Integration of CPS Security and Safety. **IEEE Embedded Systems Letters**, v. 11, n. 4, p. 111–114, 1 dez. 2019.

ATLAS MARINE. **AMS alarm system, Solution for Anti Heeling system**. Disponível em: <http://atlasmarine.sg/alarm-system/,>. Acesso em: 17 abr. 2023.

AVIŽIENIS, A. et al. Basic concepts and taxonomy of dependable and secure computing. **IEEE Transactions on Dependable and Secure Computing**, v. 1, n. 1, p. 11–33, 2004.

BHOLE, M.; KASTNER, W.; SAUTER, T. **A Model Based Framework for Testing Safety and Security in Operational Technology Environments**. IEEE International Conference on Emerging Technologies and Factory Automation, ETFA. **Anais**...Institute of Electrical and Electronics Engineers Inc., 2022.

BOUISSOU, M. **BDMP (Boolean logic Driven Markov Processes) ® as an alternative to Event Trees Benchmark on dependability of complex dynamic systems**. Proceedings of the 17nd European Safety and Reliability Conference (ESREL 2008). **Anais**...2008.

BRABAND, J. **What ' s Security Level got to do with Safety Integrity Level ?** 8th European Congress on Embedded Real Time Software and Systems (ERTS 2016). **Anais**...2016. Disponível em: <https://hal.science/hal-01289437>

BUSQUIM E SILVA, R. A. et al. Cybersecurity Assessment Framework for Digital Interface Between Safety and Security at Nuclear Power Plants. **International Journal of Critical Infrastructure Protection**, v. 34, 1 set. 2021.

BUTTGEREIT, S. et al. Demo: Leveraging SDN in Critical Infrastructures. **2021 24th Conference on Innovation in Clouds, Internet and Networks and Workshops, ICIN 2021**, p. 86–88, 1 mar. 2021.

BYRES, ERIC.; LOWE, JUSTIN.; LEVERSAGE, DAVID. **Repository of Industrial Security Incidents (RISI)**. Disponível em: <http://www.risidata.com/Database>.

CCDCOE. **NATO Cooperative Cyber Defence Centre of Excellence**. Disponível em: <https://ccdcoe.org/>. Acesso em: 9 out. 2023.

CEN. **CEN/CLC/JTC 13 - Cybersecurity and Data Protection**. Disponível em: <https://standards.cen.eu/>. Acesso em: 2 jun. 2023.

CZEKSTER, R. M.; MORISSET, C. **BDMPathfinder: A tool for exploring attack paths in models defined by Boolean logic Driven Markov Processes: Short Paper**. Proceedings - 2021 17th European Dependable Computing Conference, EDCC 2021. **Anais**...Institute of Electrical and Electronics Engineers Inc., 2021.

DE SOUZA, N. P. et al. Extending STPA with STRIDE to identify cybersecurity loss scenarios. **Journal of Information Security and Applications**, v. 55, p. 102620, 2020.

DECEA. **TWR - Torre de Controle de Aeródromo**. . Acesso em: 13 set. 2023.

DI MAIO, F.; MASCHERONA, R.; ZIO, E. Risk Analysis of Cyber-Physical Systems by GTST-MLD. **IEEE Systems Journal**, v. 14, n. 1, p. 1333–1340, mar. 2020.

EL-KADY, A. H. et al. Analysis of safety and security challenges and opportunities related to cyber-physical systems. **Process Safety and Environmental Protection**, v. 173, p. 384–413, 1 maio 2023.

EUROCAE. **WG-72 Aeronautical Systems Security**. Disponível em: <https://www.eurocae.net/about-us/working-groups/>. Acesso em: 2 jun. 2023.

FALLIERE, N.; MURCHU, L. O.; CHIEN, E. **W32.Stuxnet Dossier**. [s.l: s.n.]. Disponível em: <https://nsarchive.gwu.edu/document/21440-document-44>. Acesso em: 4 fev. 2024.

FAN, S.; YANG, Z. Safety and security co-analysis in transport systems: Current state and regulatory development. **Transportation Research Part A: Policy and Practice**, v. 166, p. 369–388, 1 dez. 2022.

GEORGE, P. G.; RENJITH, V. R. **Evolution of Safety and Security Risk Assessment methodologies towards the use of Bayesian Networks in Process Industries**. **Process Safety and Environmental Protection**Institution of Chemical Engineers, , 1 maio 2021.

GUZMAN, N.; KOZINE, I.; LUNDTEIGEN, M. An integrated safety and security analysis for cyber-physical harm scenarios. **Safety Science**, v. 144, 1 dez. 2021.

HARDY, T. L. **Software and system safety : Accidents, incidents, and lessons learned**. Second ed. [s.l.] Great Circle Analytics, 2012.

HOLLERER, S.; KASTNER, W.; SAUTER, T. **Towards a Threat Modeling Approach Addressing Security and Safety in OT Environments**. IEEE International Workshop on Factory Communication Systems - Proceedings, WFCS. **Anais**...Institute of Electrical and Electronics Engineers Inc., 9 jun. 2021.

HOLLERER, S.; SAUTER, T.; KASTNER, W. **Risk Assessments Considering Safety, Security, and Their Interdependencies in OT Environments**. ACM International Conference Proceeding Series. **Anais**...Association for Computing Machinery, 23 ago. 2022.

IAEA. **Computer and information security at nuclear facilities | IAEA**. Disponível em: <https://www.iaea.org/topics/computer-and-information-security>. Acesso em: 2 jun. 2023.

IMO. **Maritime cyber risk**. Disponível em: <https://www.imo.org/en/OurWork/Security/Pages/Cyber-security.aspx>. Acesso em: 5 jun. 2023.

INTERNATIONAL ELECTROTECHNICAL COMMISSION. **IEC 62443: Security for industrial automation and control systems**. , 2020. Disponível em: <https://www.iec.ch/blog/understanding-iec-62443>

INTERNATIONAL ELECTROTECHNICAL COMMISSION. **TC 65 Industrial-process measurement, control and automation**. Disponível em:

<https://www.iec.ch/dyn/www/f?p=103:7:617133078323724::::FSP_ORG_ID,FSP_LANG_ID:1250,25>. Acesso em: 2 jun. 2023.

INTERNATIONAL SOCIETY OF AUTOMATION. **ISA99, Industrial Automation&Control Sys Security- ISA**. Disponível em: <https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa99>. Acesso em: 2 jun. 2023.

JAPS, S. **Security safety by model-based requirements engineering**. Proceedings of the IEEE International Conference on Requirements Engineering. **Anais**...IEEE Computer Society, 1 ago. 2020.

KHAN, S. et al. **Model Checking the Multi-Formalism Language FIGARO**. Proceedings - 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks, DSN 2021. **Anais**...Institute of Electrical and Electronics Engineers Inc., 1 jun. 2021.

KHAN, S.; KATOEN, J.-P.; BOUISSOU, M. A Compositional Semantics of Boolean-Logic Driven Markov Processes. **IEEE Transactions on Dependable and Secure Computing**, p. 1–15, 2023.

KRIAA, S. **Joint safety and security modeling for risk assessment in cyber physical systems**. Paris: L'Universite Paris-Saclay, 2016.

KRIAA, S.; BOUISSOU, M.; LAAROUCHI, Y. **A model based approach for SCADA safety and security joint modelling: S-cube**. IET Conference Publications. **Anais**...Institution of Engineering and Technology, 2015.

KRIAA, S.; BOUISSOU, M.; LAAROUCHI, Y. A new safety and security risk analysis framework for industrial control systems. **Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability**, v. 233, n. 2, p. 151–174, 1 abr. 2019.

KRIAA, S.; BOUISSOU, M.; PIÈTRE-CAMBACÉDÈS, L. **Modeling the Stuxnet attack with BDMP: Towards more formal risk assessments**. 7th International Conference on Risks and Security of Internet and Systems, CRiSIS 2012. **Anais**...2012.

LANGNER, R. Stuxnet: Dissecting a cyberwarfare weapon. **IEEE Security and Privacy**, v. 9, n. 3, p. 49–51, maio 2011.

LEVESON, N. G. Using STAMP to develop leading indicators. **Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft fur Informatik (GI)**, v. P-232, p. 597–600, 2014.

LEVESON, N.; THOMAS, J. **STPA Handbook (MIT-STAMP-001)**. [s.l: s.n.].

LISOVA, E.; ŠLJIVO, I.; ČAUŠEVIĆ, A. Safety and Security Co-Analyses: A Systematic Literature Review. **IEEE Systems Journal**, v. 13, n. 3, p. 2189–2200, 1 set. 2019.

LYU, X.; DING, Y.; YANG, S. H. **Safety and security risk assessment in cyberphysical systems**. **IET Cyber-Physical Systems: Theory and Applications**Institution of Engineering and Technology, , 1 set. 2019.

MISP. **MISP Open Source Threat Intelligence Platform. Open Standards For Threat Information Sharing**. Disponível em: <https://www.misp-project.org/>. Acesso em: 11 jan. 2024.

MIT PSASS GROUP. **MIT Partnership for Systems Approaches to Safety and Security (PSASS) | STAMP Tools**. Disponível em: <http://psas.scripts.mit.edu/home/research-2/>. Acesso em: 19 nov. 2023.

MOS, M. A.; CHOWDHURY, M. M. The Growing Influence of Ransomware. **IEEE International Conference on Electro Information Technology**, v. 2020- July, p. 643–647, 2020.

NAI FOVINO, I.; MASERA, M.; DE CIAN, A. Integrating cyber attacks within fault trees. **Reliability Engineering and System Safety**, v. 94, n. 9, p. 1394–1402, 2009.

NICOLETTI, S. M. et al. Model-based joint analysis of safety and security: survey and identification of gaps. **Computer Science Review**, v. 50, p. 100597, 1 nov. 2023.

NIST. **Common Vulnerability Scoring System Calculator**. Disponível em: <https://nvd.nist.gov/vuln-metrics/cvss/v3-calculator>. Acesso em: 22 nov. 2023.

ORG, F. **Common Vulnerability Scoring System version 4.0 Draft Specification Document**. Disponível em: <https://www.first.org/cvss/v4-0/cvss-v40-specification.pdf>. Acesso em: 6 dez. 2023.

OUEIDAT, T.; FLAUS, J.-M.; MASSÉ, F. A New Way to Generate Automatically the Attacks Scenarios and Combine them with Safety Risks. p. 327–334, 2023.

PEFFERS, K. et al. A Design Science Research Methodology for Information Systems Research. **https://doi.org/10.2753/MIS0742-1222240302**, v. 24, n. 3, p. 45–77, dez. 2014.

PEKARIC, I. et al. A systematic review on security and safety of self-adaptive systems. **Journal of Systems and Software**, v. 203, 1 set. 2023.

PEREIRA, D. P.; HIRATA, C.; NADJM-TEHRANI, S. A STAMP-based ontology approach to support safety and security analyses. **Journal of Information Security and Applications**, v. 47, p. 302–319, 1 ago. 2019.

PIETRE-CAMBACEDES, L.; BOUISSOU, M. **Modeling safety and security interdependencies with BDMP (Boolean logic Driven Markov Processes)**. Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics. **Anais**...2010.

RAMOS, R. B.; JÚNIOR, J. B. C. **Safety and Security Integrated Analysis Approaches Considering New Updates for Maritime Systems**. Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023). **Anais**...2023.

RASPOTNIG, C.; KARPATI, P.; OPDAHL, A. L. Combined assessment of software safety and security requirements: An industrial evaluation of the CHASSIS method. **Journal of Cases on Information Technology**, v. 20, n. 1, p. 46–69, 1 jan. 2018.

RASPOTNIG, C.; PETER KARPATI; VIKASH KATTA. A Combined Process for Elicitation and Analysis of Safety and Security Requirements. **EMMSAD 2012: Enterprise, Business-Process and Information Systems Modeling**, v. 1, p. 347–361, 2012.

RISK SPECTRUM AB. **RiskSpectrum® | Risk & Reliability Software**. Disponível em: <https://www.riskspectrum.com/>. Acesso em: 6 jun. 2023.

RTCA. **Airport Security Access Control Systems - RTCA**. Disponível em: <https://www.rtca.org/sc-224-airport-security-access-control-systems-11/>. Acesso em: 2 jun. 2023.

SANTOS, E. et al. Análise de Perda de Pacotes em Sistema de Controle em Rede sem Fio com Aplicação de Filtros Kalman. n. September, 2017.

SBTA. **Brazilian Society of Air Transportation Research**. Disponível em: <http://sitraer.cic.unb.br/index.php/sbta/>. Acesso em: 21 fev. 2024.

SHIPUNOV, I. S. et al. **Investigation of Computer Incidents as an Important Component in the Security of Maritime Transportation**. Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, ElConRus 2021. **Anais**...Institute of Electrical and Electronics Engineers Inc., 26 jan. 2021.

SUN, X. **A Safety-Security Integrated Analysis Approach**. 2nd International Conference on Mechanical, Electronic, Control and Automation Engineering (MECAE 2018). **Anais**...Atlantis Press, 1 mar. 2018. Disponível em: <https://www.atlantis-press.com/proceedings/mecae-18/25893716>. Acesso em: 18 abr. 2023

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.

YANG, X. et al. Physical Security and Safety of IoT Equipment: A Survey of Recent Advances and Opportunities. **IEEE Transactions on Industrial Informatics**, v. 18, n. 7, p. 4319–4330, 1 jul. 2022.

YAQOOB, T.; ABBAS, H.; SHAFQAT, N. Integrated Security, Safety, and Privacy Risk Assessment Framework for Medical Devices. **IEEE Journal of Biomedical and Health Informatics**, v. 24, n. 6, p. 1752–1761, 1 jun. 2020.

YARZA, I. et al. Safety and security collaborative analysis framework for high-performance embedded computing devices. **Microprocessors and Microsystems**, v. 93, 1 set. 2022.

YOUNG, W.; LEVESON, N. G. Inside risks an integrated approach to safety and security based on systems theory: Applying a more powerful new safety methodology to security risks. **Communications of the ACM**, v. 57, n. 2, p. 31–35, 1 fev. 2014.

ZHOU, C. et al. Risk-Based Scheduling of Security Tasks in Industrial Control Systems with Consideration of Safety. **IEEE Transactions on Industrial Informatics**, v. 16, n. 5, p. 3112–3123, 1 maio 2020.

GLOSSÁRIO

Ameaça	*Threats*, problema com origem em *security*
Cut Sets	Grupos de corte, conjunto que leva a ocorrência do evento topo
Cybersecurity	Segurança Cibernética, Security, relacionado a ataques intencionais
Evento topo	Folha mais alta de uma árvore, módulo do último nível
Grupo de Corte	*Cut Sets,* conjunto que levam a ocorrência do evento topo
Hazard	*Perigo, problema com origem em Safety*
Jamming	Ataque de Negação de Serviços
Perigo	*Hazard, problema com origem em Safety*
Qualitativo	Relacionados a combinações lógicas, valores discretos e sentenças textuais.
Quantitativo	Relacionados a valores matemáticos, probabilidades
Safety	Segurança crítica, relacionado a acidentes não intencionais
Security	Segurança Cibernética, *Cybersecurity*, relacionado a ataques intencionais
Segurança Cibernética	*Security or Cybersecurity,* relacionado a ataques intencionais
Segurança Crítica	*Safety, relacionado a acidentes não intencionais*
Sistema *Anti-Heeling*	Sistema que visa evitar inclinação de navios
Sistemas Ciber Físicos	Sistemas com interligações digitais e mecânicas
Spoofing	Ataque que consiste em fazer um dispositivo assumir a identidade de outro
Tampering	Ataque de adulteração de dados

APÊNDICE A – CÓDIGO EM MATLAB PARA CADEIAS DE MARKOV

```
% Universidade de São Paulo, Escola Politécnica
% Autor: Rogério Brito Ramos, rogerio.ramos@usp.br
% Sistema Anti-Heeling

% Inicia com a tela limpa
clf; clear;
% Declaração da Unidade de tempo variando de 0 até maximo incrementado em deltaT para
ajuste % do gráfico
deltaT=1/6; maximo = 1000; t=0:deltaT:maximo;

lambda1 = 0.001; % Taxa de falha para M1: Proteção contra Spoofing
lambda2 = 0.0005; % Taxa de Falha para M2: Sistema AH
lambda3 = 0.001; % Taxa de Falha para M3: Proteção contra Ataques Jamming
lambda4 = 0.005; % Taxa de Falha para M4: Evitar exposição a variação de carga

% Declaração das funções de Probabilidade dos estados
syms P1(t) P2(t) Pf(t);
% Equações Diferenciais dos estados
edP1 = diff(P1,t) == -(lambda1+lambda2+lambda3)*P1;
edP2 = diff(P2,t) == (lambda2+lambda3)*P1-(lambda1+lambda4)*P2;
edPf = diff(Pf,t) == lambda1*P1+(lambda1+lambda4)*P2;

% Vetor das equações diferenciais
edos = [edP1; edP2; edPf];
% Vetor das condições iniciais
cond = [P1(0)==1;  P2(0)==0; Pf(0)==0];
% Vetor com a solução das equações diferenciais
[P1sol(t), P2sol(t), Pfsol(t)]= dsolve (edos,cond);

% Representa a soma das probabilidades dos estados seguros P1sol e P2Sol
R=P1sol+P2sol;

% Coleta a Probabilidade do sistema está operando no tempo 1000
fprintf('R(1000) = %.4f \n', R(1000));
% Graphical
hold on;
% Desenha em vermelho o gráfico da probabilidade de estar em estado de falha
fplot(@(t) Pfsol(t), [0 maximo],'-r');
% Desenha em verde o gráfico da probabilidade de estar em um estado seguro
fplot(@(t) R(t), [0 maximo],'-g');

% Ajustes nos gráficos para facilitar a visualização
xlabel('Time (d) in days');
ylabel('Probability');
title('Anti Heeling System')
legend('Toppled', 'Safe');
hold off; grid on;
% Fim do script
```