



ENGENHARIA MILITAR

**Mobilidade,
Contramobilidade e
Proteção**

- Redes de comunicações militares: desafios tecnológicos e propostas para atendimento dos requisitos operacionais do Exército Brasileiro – Pág 05
- Dimensionamento do reforço ao cisalhamento com compósitos de resina e fibras em vigas-parede de concreto – Pág 26
- Avaliação da corrosão de ligas de magnésio por ensaios de imersão e evolução de hidrogênio – Pág 38
- Análise de comportamento de malware utilizando redes neurais recorrentes - uma abordagem por intermédio da previsão de opcodes – Pág 45
- Otimização da trajetória de robôs móveis diferenciais de tração usando algoritmo de otimização por enxame de partículas – Pág 57
- Rede IoT assistida por sistema de aeronaves remotamente pilotadas para apoio em operações de recuperação de desastres – Pág 63
- A Influência do Teor de Austenita Revertida no Comportamento Mecânico do Aço Maraging 250 – Pág 75



Biblioteca do Exército

Tradição e qualidade em publicações

 Biblioteca do Exército (BIBLIE) - Casa do Barão de Loreto é uma centenária instituição cultural do Exército Brasileiro que contribui para o provimento, a edição e a difusão de meios bibliográficos necessários ao desenvolvimento e aperfeiçoamento da cultura profissional-militar e geral.

**SEJA NOSSO
ASSINANTE**

e receba nossos livros
editados em sua residência.



Tel.: (21) 2519-5707

Praça Duque de Caxias, nº 25
Palácio Duque de Caxias
Ala Marcílio Dias - 3º Andar
Centro - CEP 20.221-260
Rio de Janeiro – RJ



Acesse:

www.bibliex.eb.mil.br





REVISTA MILITAR DE CIÊNCIA E TECNOLOGIA

COMANDANTE DO EXÉRCITO:

Gen Ex Edson Leal Pujol

DEPARTAMENTO DE CIÊNCIA E TECNOLOGIA:

Gen Ex Décio Luís Schons

DEPARTAMENTO DE EDUCAÇÃO E CULTURA DO EXÉRCITO:

Gen Ex Tomás Miguel Miné Ribeiro Paiva

EDITOR E DIRETOR DA BIBLIEEx:

Cel Cav Marco André Leite Ferreira

CORPO REDATORIAL:

- Presidente: Gen Bda Armando Morado Ferreira • Editor geral: Marcelo de Miranda Reis • Editor adjunto: Paulo Henrique Coelho Maranhão – IME (SE/4) • Engenharia Civil, Transportes e Ciências Ambientais: Antônio Carlos Rodrigues Guimarães – IME (SE/2)
- Engenharia Nuclear: Sérgio de Oliveira Vellozo – IME (SE/7) • Engenharia Elétrica: Antônio Eduardo Carrilho da Cunha – IME (SE/3) • Engenharia Mecânica: André Luiz Tenório Rezende – IME (SE/4) • Ciência dos Materiais: André Ben-Hur da Silva Figueiredo – IME (SE/8) • Ciência da Computação: Paulo Cesar Salgado Vidal – IME (SE/9) • Engenharia de Defesa: Antônio Eduardo Carrilho da Cunha – IME (SE/10) • Engenharia Química e Química: Kátia Regina de Souza – IME (SE/5) • Engenharia Cartográfica: Heloísa Alves Silva Marques – IME (SE/6) • Outras áreas: Aderson Campos Passos – IME (SE/9).

EDITORES ASSOCIADOS EXTERNOS:

- Dr. André Fenili – Universidade Federal do ABC, Santo André, SP • Dr. Artur Ziviani – Lab. Nacional de Computação Científica (LNCC), Petrópolis, RJ • Dr. Fernando Fachini Filho – Instituto Nacional de Pesquisas Espaciais, SP • Dr. José Carlos Costa da Silva Pinto – Universidade Federal do Rio de Janeiro, RJ • Dr. José Carlos Maldonado – Universidade de São Paulo, São Carlos, SP • Drª. Júlia Célia Mercedes Strauch – Escola Nacional de Ciências Estatísticas, RJ • Dr. Luiz Pereira Calóba – Universidade Federal do Rio de Janeiro, RJ • Dr. Otto Corrêa Rotunno Filho – COPPE/Universidade Federal do Rio de Janeiro, RJ • Dr. Richard Magdalena Stephan – COPPE/Universidade Federal do Rio de Janeiro, RJ • Dr. Webe João Mansur – COPPE/Universidade Federal do Rio de Janeiro, RJ • Dr. Carlos Alberto Nunes Consenza – COPPE/Universidade Federal do Rio de Janeiro, RJ

COMPOSIÇÃO:

INSTITUTO MILITAR DE ENGENHARIA - IME
Praça General Tibúrcio, 80 - Praia Vermelha - Rio de Janeiro-RJ - CEP 22.290-270 - Tel.: (21) 2546-7115

Website: <http://rmct.ime.eb.br> - Email: rmct@ime.eb.br

REVISÃO, PROGRAMAÇÃO E DIAGRAMAÇÃO:

Coordenador: Rubenildo Pithon de Barros • Revisores: Giana Elvira de Oliveira Marquês; Suzana Marly da Costa Magalhães • Programador Web: Rubenildo Pithon de Barros • Diagramador: Luiz Tadeu Carqueja Mota

PROJETO GRÁFICO:

BIBLIEEx

IMPRESSÃO:

Tavares & Tavares Empreendimentos Comerciais LTDA



REVISTA MILITAR de
CIÊNCIA e TECNOLOGIA

REVISTA MILITAR DE CIÊNCIA E

TECNOLOGIA - Volume XXXVII

3º Trimestre de 2020

Direção, edição e distribuição:

BIBLIOTECA DO EXÉRCITO (BIBLIEEx)

Palácio Duque de Caxias - Praça D. de Caxias, 25

3º andar - Ala Marcílio Dias

Centro - Rio de Janeiro-RJ

CEP 20.221-260

Tel.: (21) 2519-5707

ACESSE NOSSAS REVISTAS DIGITAIS



Nossa capa:

Homenagem ao Corpo de Engenheiros

Projeto Gráfico: BIBLIEEx



EDITORIAL



REVISTA MILITAR de
CIÊNCIA e TECNOLOGIA

Prezados membros da comunidade que busca a pesquisa e o desenvolvimento científico-tecnológico de interesse do Exército e do País, no tocante à Estratégia Nacional de Defesa, é com muita satisfação que me dirijo a vocês leitores desta tradicional revista.

Na edição do último trimestre pode ser observado o novo projeto gráfico de nossa revista, mais harmonioso e integrado ao design das demais revistas do Exército Brasileiro. Alguns artigos destacaram o trabalho conduzido por pesquisadores do Instituto Militar de Engenharia na busca de soluções para o enfrentamento da pandemia do COVID-19.

Nosso desafio para o próximo ano será integrar a revista a iniciativa multiuniversitária que promove acesso aberto e aperfeiçoamento da publicação acadêmica, através do uso do “Open Journal System”, que é um software livre para gerenciamento e publicação de periódicos científicos.

Nesta edição, visando atender aos requisitos operacionais básicos do sistema de comando e controle das operações militares, foram pesquisados os desafios tecnológicos para as redes de comunicações militares. Na área de defesa cibernética e segurança da informação, foi feita uma análise de comportamento de malware utilizando redes neurais recorrentes. Para otimizar as eventuais ações da Mão Amiga do Exército Brasileiro, foi estudado o uso de aeronaves remotamente pilotadas, em uma rede IoT, para apoio em operações de recuperação de desastres. No campo da mecatrônica e inteligência artificial, foi debatida a otimização da trajetória de robôs móveis usando algoritmo de otimização por enxame de partículas. A pesquisa no escopo da Engenharia de Fortificação e Construção passou pelo dimensionamento do reforço ao cisalhamento com compósitos de resina e fibras em vigas-parede. Além disso, o uso de materiais estratégicos para emprego militar foi o tema dos artigos sobre: avaliação da corrosão de ligas de magnésio e influência do teor de austenita no comportamento mecânico do aço Maraging 250.

O hábito da leitura, além de aumentar o conhecimento, aprimora o vocabulário e ajuda na construção textual. Aproveitem esta edição!

Marcelo de Miranda Reis

SUMÁRIO

05

- Redes de comunicações militares: desafios tecnológicos e propostas para atendimento dos requisitos operacionais do Exército Brasileiro

Marcelo José Camilo, David Fernandes Cruz Moura e Ronaldo Moreira Salles

26

- Dimensionamento do reforço ao cisalhamento com compósitos de resina e fibras em vigas-parede de concreto

Fábio G. Ávila, Ana M. A. J. Abreu, Luiz A. V. Carneiro*

38

- Avaliação da corrosão de ligas de magnésio por ensaios de imersão e evolução de hidrogênio

Ana Caroline Crema de Almeida Fontes, Carlos Nelson Elias

45

- Análise de comportamento de malware utilizando redes neurais recorrentes - uma abordagem por intermédio da previsão de opcodes

Davi Gomes de Albuquerque, Lincoln de Queiroz Vieira, Ricardo Sant'Ana e Julio Cesar Duarte*

57

- Otimização da trajetória de robôs móveis diferenciais de tração usando algoritmo de otimização por enxame de partículas

Gustavo S. Rodrigues, Carlos L. M. de Souza Juniorb, Elias D. R. Lopes, Marcelo A. Acuña*

63

- Rede IoT assistida por sistema de aeronaves remotamente pilotadas para apoio em operações de recuperação de desastres

Erick M. Moreira, Fábio Luiz Jr, Leandro Moreira, Jauvane C. de Oliveira e Paulo F. F. Rosa

75

- A Influência do Teor de Austenita Revertida no Comportamento Mecânico do Aço Maraging 250

Daniel Fonseca da Cunha, Alaelson Viera Gomes, Achille Arantes Bassi, Luiz Paulo de Mendonça Brandão*

Redes de comunicações militares: desafios tecnológicos e propostas para atendimento dos requisitos operacionais do Exército Brasileiro

Marcelo José Camilo^{*a}, David Fernandes Cruz Moura^b e Ronaldo Moreira Salles^a

^a Instituto Militar de Engenharia (IME) – Praça Gen. Tibúrcio, 80 – Urca, Rio de Janeiro – RJ, 22290-270

^b Universidade Estadual de Campinas (UNICAMP) – Cidade Universitária Zeferino Vaz – Barão Geraldo, Campinas – SP, 13083-970

* camilo.marcelo@eb.mil.br

RESUMO: As operações militares possuem atividades ininterruptas e, na maioria das vezes, ocorrem em cenários desconhecidos e sempre hostis. Uma condição necessária para o sucesso de uma operação militar é que todos os elementos integrantes desta precisam, durante todo o desenrolar da operação, manter as comunicações com seus escalões superiores e subordinadas. Atualmente, estão disponíveis para emprego várias tecnologias de redes de comunicações. Entretanto, nenhuma infraestrutura de comunicações pronta para uso reúne condições de suportar os Princípios de Emprego das Comunicações do EB bem como os Requisitos Operacionais Básicos do Sistema de Comando e Controle desta instituição simultaneamente, deixando desafios, tais como a integração automática das camadas de enlace, segurança na camada física e desempenho em cenários de espectro rádio carregado, para o atendimento destes Princípios e Requisitos. Neste trabalho propomos as tecnologias de rádios cognitivos e de transmissão full-duplex em banda como candidatas para enfrentar estes desafios.

PALAVRAS-CHAVE: Comunicações Táticas. Rádios Cognitivos. Transmissão Full-Duplex em Banda.

ABSTRACT: Military operations have uninterrupted activities and mostly occur in unknown and always hostile scenarios. A necessary condition for a successful military operation is that its deployed elements communicate with its superior and subordinate levels. A lot of networking technologies are now available for employment. However, no ready-to-use communications infrastructure is able to simultaneously support envisioned Brazilian Army (in Portuguese, EB) Communications Operational Principles, as well as EB Command and Control System's Basic Operational Requirements. This scenario brings several challenges to fulfill the aforementioned principles and requirements, such as the automatic integration of link layers, security in the physical layer, and performance in overloaded radio spectrum scenarios. In this work, we propose cognitive radios and in-band full-duplex operation as candidate technologies to overcome these challenges.

KEY WORDS: Tactical communications. Cognitive Radios. In-Band Full-Duplex operation.

1. Introdução

O Exército Brasileiro deve ser capaz de atuar em operações de guerra e de não guerra, requerendo, normalmente, combinações simultâneas destas duas modalidades. Assim, este dois tipos de operações podem coexistir em diferentes proporções, a depender da situação enfrentada e dos objetivos estratégicos.

As operações militares ocorrem de maneira

ininterrupta, ou seja, possuem atividades 24 horas por dia, 7 dias por semana. Outra característica das operações militares é que, na maioria das vezes, ocorrem em cenários desconhecidos e sempre hostis. Uma condição necessária para o sucesso de uma operação militar é que todos os elementos integrantes de uma operação militar precisam, durante todo o desenrolar da operação, manter as comunicações com seus escalões superiores e subordinadas.

Os manuais e portarias do EB listam os Princípios de Emprego das Comunicações do EB bem como os Requisitos Operacionais Básicos do Sistema de Comando e Controle desta instituição, um dos principais sistemas utilizados pelo EB tanto na execução das operações quanto no planejamento destas atividades. Os citados princípios e requisitos tem por objetivo estabelecer diretrizes para as ligações entre os diversos níveis hierárquicos, visando atender a condição necessária para o sucesso da operação citada no parágrafo anterior.

Atualmente, estão disponíveis para emprego várias tecnologias de redes de comunicações, tais como, rede rádio de combate com transmissão digital, rede pacote rádio, redes de arquitetura estação-base, redes *ad hoc*, redes de acesso múltiplo por divisão de tempo e redes rádio de banda ultra larga. Além dessas, há algumas tecnologias em desenvolvimento das quais podemos citar as de rádios cognitivos e a de transmissão *full-duplex* em banda.

Entretanto, conforme exposto em [1] estudo, nenhuma infraestrutura de comunicações pronta para uso reúne condições de suportar os Princípios de Emprego das Comunicações do EB bem como os Requisitos Operacionais Básicos do Sistema de Comando e Controle desta instituição simultaneamente. Outro aspecto abordado em um estudo [1] é que três requisitos são estreitamente relacionados e constituem um forte compromisso no sistema de comunicações militar: alcance, capacidade e mobilidade. O forte relacionamento existente se dá, em geral, pelo fato que ao variar um dos fatores os outros dois também variam. Na verdade, ao se maximizar dois fatores o terceiro acaba sendo minimizado.

Sendo assim, um sistema de comunicações militar deve ser uma única estrutura lógica integrada por múltiplas tecnologias de enlace. A integração destas tecnologias devem ser feitas de maneira automática e sem a necessidade de intervenção humana visando a agilidade e a flexibilidade. As tecnologias prontas para uso não possibilitam esta integração automática deixando desafios para o atendimento aos Princípios de Emprego das Comunicações e ao Requisitos do Sistema de Comando e Controle do EB.

O principal desafio das redes atuais é prover a

possibilidade de realizar, em tempo de operação, sem perda de capacidade, alcance e/ou mobilidade, a troca de tecnologias utilizadas. Outro desafio das redes táticas atuais é aumentar a segurança na camada física das redes de comunicações táticas sem fio. Além destes, temos o desafio de aumentar o desempenho da redes rádios, especialmente em cenários de alto tráfego ou de interferência eletromagnética.

Duas tecnologias que figuram como candidatas promissoras para enfrentar os desafios citados acima são a tecnologia de rádios cognitivos e a de transmissão *full-duplex* em banda.

O objetivo deste trabalho é analisar as redes de comunicações à luz dos Princípios de Emprego das Comunicações do Exército Brasileiro e dos Requisitos Operacionais Básicos do Sistema de Comando e Controle desta instituição, levantando seus desafios e propondo tecnologias candidatas para enfrentar estes desafios.

A estrutura deste texto é a seguinte: na Seção 2, descrevemos as operações do Exército Brasileiro. Na Seção 3, apresentamos as Redes de Comunicações do Exército Brasileiro. Na Seção 4, analisamos as redes de comunicações do EB à luz dos Princípios de Emprego das Comunicações do EB e nos Requisitos Operacionais Básicos do Sistema de Comando e Controle desta instituição. Na Seção 5, estudamos as tecnologias existentes para o emprego nas redes de comunicações do Exército Brasileiro. Na Seção 6, levantamos os desafios tecnológicos para que as redes de comunicações militares atendam aos requisitos operacionais do Exército Brasileiro bem como propomos as tecnologias de rádios cognitivos e a de transmissão *full-duplex* em banda como candidatas promissoras para enfrentar os desafios levantados. Por fim, na Seção 7, fazemos nossas conclusões.

2.Operações do Exército Brasileiro

O Exército Brasileiro, junto com a Marinha do Brasil e a Força Aérea Brasileira, constituem as FFAA (Forças Armadas) brasileiras. Estas instituições

estão subordinadas ao MD (Ministério da Defesa) e destinam-se à defesa da pátria, à garantia dos poderes constitucionais e, por iniciativa de qualquer destes, da lei e da ordem [2].

O EB (Exército Brasileiro) executa suas missões através de operações militares que são ações coordenadas de militares, em respostas a uma situação específica. O conjunto de princípios que servem de base para que o EB desempenhe suas missões é descrito em manuais desta instituição.

O EB20-MC-10.202 - Manual de Campanha – FORÇA TERRESTRE COMPONENTE [3], cita que a integração e a sincronização das operações do EB com as outras forças armadas e outras organizações públicas, tais como Polícia Militar e Defesa Civil, ou privadas, tais como concessionárias de energia elétrica e de telefonia, participantes de uma operação é de responsabilidade da FTC (Força Terrestre Componente). A FTC, assim como a FNC (Força Naval Componente), a FAC (Força Aérea Componente), entre outros elementos operacionais, são subordinados ao Cmt Op (Comando Operacional) Conjunto.

O citado manual também enfatiza que a FTC não possui uma organização fixa, devendo ser estruturada para atender às demandas do planejamento operacional. Os meios que integram a FTC são adjudicados pelo Cmt Op Conjunto, em consonância com os objetivos do MD, levadas em consideração as necessidades levantadas no planejamento operacional e as disponibilidades do citado Ministério. Caso, no decorrer dos planejamentos ou da operação, o Comando da FTC identifique a necessidade de outros meios, estes poderão ser solicitados ao Cmt Op Conjunto.

Dependendo do vulto da operação, a FTC pode ser composta de uma ou mais brigadas. Uma brigada corresponde à grande unidade de menor escalão do EB capaz de atuar de maneira independente. A brigada é, geralmente, formada por 3 batalhões ou regimentos, que constituem seus elementos operacionais. Além disso, uma brigada possui elementos de apoio às operações tais como a Companhia de Engenharia de Combate, responsável pelas atividades de mobilidade, contramobilidade e proteção da operação, e, de especial

interesse para este artigo, a Cia Com (Companhia de Comunicações), responsável por instalar, explorar e manter os sistemas de comunicações da Brigada.

O EB20-MF-10.103 - Manual de Fundamentos - OPERAÇÕES [4], cita que, em função dos princípios e procedimentos utilizados, as operações do EB são classificadas em *Operações de Guerra* e *Operações de Não Guerra*. Nas subseções seguintes faremos uma breve descrição destas operações com foco nos aspectos que influenciam a instalação, a utilização e a manutenção das redes de comunicações utilizadas nestas ações. Nosso objetivo nesta descrição é entender os objetivos e necessidades que levam a utilização das redes de comunicações pelo EB.

2.1 Operações de guerra

As *Operações de Guerra* são as que utilizam o poder militar, explorando a plenitude de suas características de emprego da força, ou seja, a violência militar em sua maior expressão. Nelas empregam-se todas as capacidades das organizações operativas do EB, ou ameaça fazê-lo, aplicando os princípios e procedimentos de combate.

O emprego real da EB em uma *Operação de Guerra*, teria por cenário áreas geograficamente afastadas das instalações físicas das OMs (Organizações Militares). As operações simuladas para adestramento da tropa ocorrem no mesmo cenário.

Neste tipo de operação, o ambiente operacional é constituído de uma pequena estrutura centralizada de acompanhamento a ser utilizada pelo comando da FTC para acompanhar seus subordinados. Esta estrutura também é utilizada pelos órgãos de controle e de apoio à operação, tais como o COTer (Comando de Operações Terrestres), OM responsável por orientar e coordenar o preparo e o emprego do EB, em conformidade com as políticas e diretrizes estratégicas desta instituição e do MD. Os elementos subordinados à FTC, na maior parte do tempo, realizam suas missões em áreas afastadas desta estrutura. Estas áreas de atuação variam de acordo com as necessidades operacionais e são, geralmente, ambientes desconhecidos e hostis.

2.2 Redes de comunicações nas operações de guerra

Conforme já citado na Subseção anterior, o emprego real do EB em uma *Operação de Guerra*, teria por cenário áreas geograficamente afastadas das instalações físicas das OMs.

Neste tipo de operação, a rede de comunicações da citada infraestrutura é geralmente uma rede cabeada que é interligada à EBNet, rede de dados que interliga todas as OMs do EB. Os comandantes dos elementos, tais como brigadas e batalhões, da FTC desdobrados são conectados, através de enlaces de HCLOS (*High-Capacity Line-Of-Sight* ou Linha de Visada de Alta Capacidade) providos pelo MTO (Módulo de Telemática Operacional) do EB, com a rede do comando da operação, o que provê conectividade uns com os outros e com a estrutura de acompanhamento. O MTO possibilita dotar as comunicações militares via rádio de pleno acesso à rede pública de telefonia fixa ou celular e integrar-se a qualquer cenário remoto através de sistemas de comunicação via satélite. Este equipamento permite a comunicação de dados, voz e imagens no campo de batalha.

Os comandantes das brigadas ou dos batalhões quando se encontram embarcados no PCT (Posto de Comando Tático) e os demais elementos desdobrados se comunicam através de rede de dados hierárquica, que utilizam as faixas de frequência UHF (*Ultra High Frequency* ou Frequência Ultra Alta) e VHF (*Very High Frequency* ou Frequência Muito Alta). Todos os equipamentos utilizados nestas redes tem capacidade de comunicações utilizando protocolo IP (Internet Protocol ou Protocolo Internet). A **figura 1** ilustra este cenário.

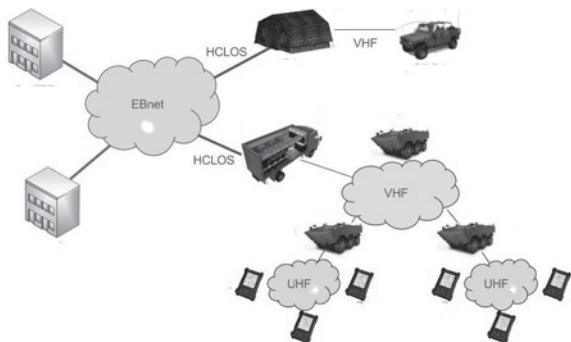


Fig. 1 – Redes de Comunicações nas Operações de Guerra.

2.3 Operações de não guerra

Nas *Operações de Não Guerra*, as FFAA, embora fazendo uso do poder militar, são empregadas em tarefas que não envolvam o combate propriamente dito, exceto em circunstâncias especiais, em que esse poder é usado de forma limitada. Podem ocorrer, inclusive, casos nos quais as ações militares não exerçam necessariamente o papel principal.

As *Operações de Não Guerra*, como as de GLO (Garantia da Lei e da Ordem), ocorridas durante os grandes eventos sediados pelo Brasil, como a Copa do Mundo FIFA, em 2014, os Jogos Olímpicos do Rio de Janeiro, em 2016, entre outros, são comuns em áreas urbanas e, dada a grande presença do Exército no território brasileiro, geralmente ocorrem próximas às instalações de OMs do EB.

Nesse tipo de operação, há emprego maior de postos de comando fixos em OMs do EB. Entretanto, a maioria dos militares circulam por uma área urbana pré determinada. Eventualmente existe a necessidade de deslocar militares em posições afastadas para defender uma estrutura estratégica, como, por exemplo, uma subestação de energia elétrica.

2.4 Redes de comunicações nas operações de não guerra

Nas *Operações de Não Guerra*, descritas na Subseção anterior, há emprego maior de postos de comando fixos (aquartelados). Os militares que ficam em posições fixas utilizam amplamente redes infraestruturadas de alta capacidade, como a EBNet ou a ROD (Rede Operacional de Defesa), rede segregada, estabelecida pela MD, que proporciona grande segurança para o fluxo de informações necessário à condução de operações conjuntas e propicia interoperabilidade às FFAA brasileiras [5]. Estas redes são os pontos de conectividade com os militares que circulam por uma área urbana pré determinada os quais trafegam suas informações através de rede celular.

Os militares em posições afastadas, para defender uma estrutura estratégica, são conectados à estrutura principal do sistema através de enlace de HCLOS. Quando este deslocamento acontecer, esta parte da operação se assemelha, do ponto de vista das redes de

comunicações, à uma *Operação de Guerra*. A **figura 2** ilustra este cenário.

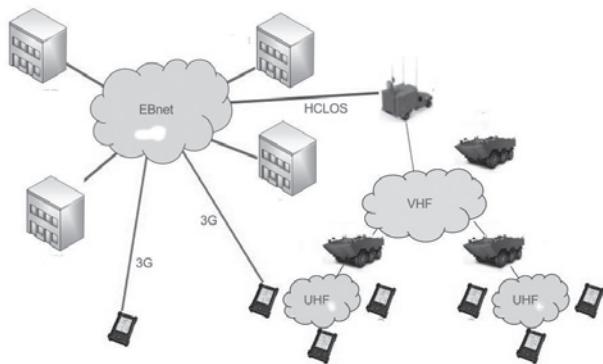


Fig. 2 – Redes de Comunicações nas Operações de Não Guerra.

2.5 Abrangência das operações do Exército Brasileiro

As operações do EB se desenvolvem em todo espectro dos conflitos, que varia desde a prevenção de ameaças à solução dos conflitos armados, passando pelo gerenciamento de crises.

Em situações de paz (estável ou instável) ou de crises, empregam-se, entre outras medidas, as de caráter militar, mediante o uso de forças militares com a aplicação de parte de suas capacidades, para evitar a escalada da crise ou anular a possibilidade de realização de operações militares de guerra de vulto. Realizam-se também operações militares em apoio às autoridades governamentais nacionais ou internacionais.

O EB20-MF-10.103 - Manual de Fundamentos - OPERAÇÕES [4] define que o EB deve ser capaz de atuar nos dois tipos de operações citados no início desta Seção, requerendo, normalmente, combinações simultâneas destas duas modalidades. Assim, estes dois tipos de operações podem coexistir em diferentes proporções, a depender da situação enfrentada e dos objetivos estratégicos.

Uma importante característica das operações militares, que determinam a estrutura das redes de comunicações a serem utilizadas, é que estas ocorrem de maneira ininterrupta, ou seja, possuem atividades 24

horas por dia, 7 dias por semana. Outra característica das operações militares é que, na maioria das vezes, ocorrem em cenários desconhecidos e sempre hostis.

Uma condição necessária para o sucesso de uma operação militar é que, todos os elementos integrantes desta precisam, durante todo o desenrolar da operação, manter as comunicações com seus superiores e subordinadas com os seguintes objetivos:

Objetivo 1

Fornecer informações que auxiliem a tomada de decisão pelos comandantes militares;

Objetivo 2

Possibilitar ao escalão superior a emissão de ordens aos elementos subordinados bem como o acompanhamento da execução destas; e

Objetivo 3

Permitir ao escalão subordinado enviar aos superiores informações sobre os acontecimentos ocorridos durante o desempenho de suas missões.

As características e objetivos das operações militares descritas nesta seção norteiam a instalação, a operação e a manutenção das redes de comunicações do EB. Estas redes serão descritas na próxima Seção.

3. Redes de comunicações do Exército Brasileiro

O trabalho de [6] cita que redes de comunicações táticas sem fios têm diferentes requisitos, expectativas, necessidades, restrições de garantia de informações etc. Esta característica se estende a todas as redes de comunicações militares.

Um primeiro aspecto importante que molda as redes de comunicações militares é a organização hierárquica de toda força militar. Salles [1] toma como exemplo um exército padrão, composto por divisões, cada uma sendo formada por brigadas, que por sua vez são formadas por batalhões e suas companhias, as quais possuem pelotões e seus grupos de combate. Em uma situação de emprego de força militar, o exército padrão é fracionado em suas unidades e subunidades de modo a garantir mobilidade e agilidade.

Os objetivos 2 e 3, citados na seção anterior, enfatizam

a organização hierárquica do EB. Essa característica faz com que os sistemas, computacionais, de transmissão de voz etc., que utilizam as redes de comunicações militares, também possuam organização hierárquica. Assim, as redes de comunicação militares devem ser modulares, permitindo interoperabilidade entre as frações em qualquer situação e independência quando separadas, sem interromper as comunicações em nenhum momento.

O exposto nos dois parágrafos anteriores mostra que as redes de comunicação militares possuem requisitos específicos, tais como alta mobilidade, grande agilidade, modularidade, interoperabilidade e independência. Este requisitos distinguem as rede de comunicações militares das demais redes.

Nesta seção, iniciamos nosso estudo das redes de comunicações do EB com a descrição dos sistemas de comunicações desta instituição. Em seguida, abordaremos os Princípios de Emprego das Comunicações do EB. Por fim, mostramos os Requisitos Operacionais do Sistema de Comando e Controle do EB.

3.1 Sistema de Comunicações do Exército Brasileiro

OC11-61-Manual de Campanha-COMUNICAÇÕES NA DIVISÃO DE EXÉRCITO [7] estrutura o SICOMEx (Sistema de Comunicações do Exército) em SEC (Sistema Estratégico de Comunicações) e SISTAC (Sistema Tático de Comunicações).

3.2 Sistema Estratégico de Comunicações

O SEC tem por objetivo o estabelecimento das ligações de longa distância, dentro do território nacional, para o atendimento das necessidades correntes e estratégicas do Exército, podendo, excepcionalmente, prestar apoio de comunicações a qualquer escalão do EB no exterior.

Para tanto, este sistema dispõe de meios de comunicações de grande versatilidade para o estabelecimento das ligações, como a EBNet (Rede Corporativa Privativa do Exército), as RRSEC (Redes Rádio do Sistema Estratégico de Comunicações) e a RITEEx (Rede Integrada de Telecomunicações do

Exército). A **figura 3** mostra o SEC e seus componentes.



Fig. 3 – Sistema Estratégico de Comunicações.

A página do CITEEx (Centro Integrado de Telemática do Exército), órgão de apoio setorial diretamente subordinado ao DCT (Departamento de Ciência e Tecnologia), descreve missão deste Centro como “estabelecer, manter e operar os sistemas de informática e comunicações de interesse do SC2Ex (Sistema de Comando e Controle do Exército) no seu nível mais elevado” [8]. No contexto desta missão, o CITEEx e sua OMs subordinadas são os responsáveis pela instalação, manutenção e gerenciamento do SEC e de seus componentes.

Os subsistemas do SISTAC são descritos a seguir.

- **EBNet**

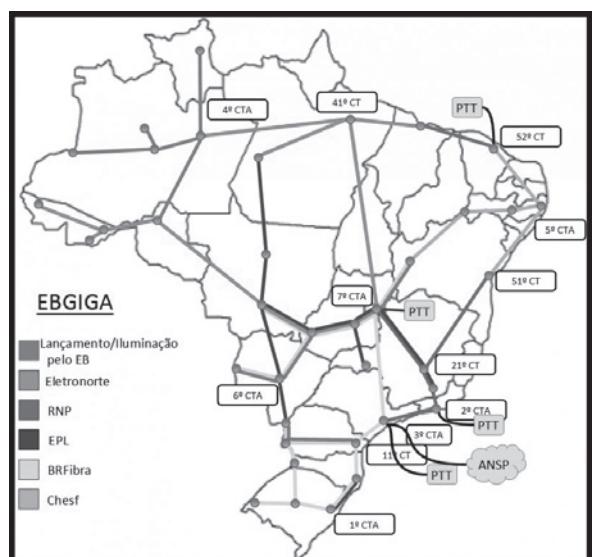


Fig. 4 – Backbone Principal da EBNet.

A EBNet é uma rede de dados que interliga todas as OMs do EB. Esta tem por finalidade o tráfego de dados

dos sistemas administrativos, tais como os sistemas de pessoal, de material, logístico etc., desta instituição. Esta rede possui tanto enlaces instalados e gerenciados por profissionais do EB quanto por enlaces contratados. A **figura 4** mostra o *backbone* principal da EBNet.

A **figura 4** cita os CTAs (Centros de Telemática de Área) e os CTs (Centros de Telemática). Os CTs e CTAs são OMs subordinadas ao CITEEx, e, portanto, responsáveis pela instalação, manutenção e gerenciamento da infraestrutura de comunicações do EB. Nesta mesma figura, os termos PTT (Ponto de Troca de Tráfego) e ANSP (*Academic Network at São Paulo* ou *Rede Acadêmica em São Paulo*) representam pontos de ligação da EBNet com outras redes de comunicações.



Fig. 5 – Backbone Principal da RRSEC.

A RRSEC é um conjunto de rádios fixos cuja finalidade é transmissão de dados em caso de indisponibilidade da EBNet. Toda a infraestrutura da RRSEC é instalada, manutenida e gerenciada por profissionais do EB. A **figura 5** mostra o *backbone* principal da RRSEC.

A **figura 5** cita, além dos CTAs e CTs, as RMs (Regiões Militares). As RMs são as OMs do EB responsáveis, entre outras atribuições, pelos aspectos administrativos em sua área de atuação. Estas OMs são grandes utilizadores tanto da EBNet quanto da RRSEC.

• RITEEx

A RITEEx é rede que possibilita a comunicação telefônica entre as OMs do EB. Também é instalada, manutenida e gerenciada por profissionais do EB e utiliza como infraestrutura o mesmo *backbone* da EBNet.

O SEC permite a integração com sistemas de comunicações do MD e de outros ministérios, como o SISCOMIS (Sistema de Comunicações Militares por Satélite) e o SIVAM (Sistema de Vigilância da Amazônia). Este sistema é utilizado em tempo de paz. Ele provê todas as funcionalidades necessárias para o correto funcionamento de uma força armada em seu cotidiano. De modo geral, possui enlaces com maior capacidade e menor mobilidade [1]. O SEC é interligado à Internet.

3.3 Sistema Tático de Comunicações

O SISTAC é o conjunto de meios de comunicações e informática pertencentes as unidades operacionais do EB, destinado ao preparo e emprego de tropas. Sua utilização, portanto, se dá em missões de adestramento ou em operações. Pode-se dizer que possui enlaces com menor capacidade e maior mobilidade [1]. A **figura 6** mostra o SISTAC e seus subsistemas.



Fig. 6 – Sistema Tático de Comunicações.

Os subsistemas do SISTAC são descritos a seguir.

• Subsistema de Rádio de Combate

É utilizado para a transmissão de voz em operações de guerra, ou seja, em áreas geograficamente afastadas das instalações físicas das OMs que consequentemente, devido ao ambiente ser desconhecido e hostil, não permite a instalação de infraestrutura como antenas de

retransmissão. Possui baixo alcance, alta mobilidade e grande facilidade de instalação. Devido ao ambiente hostil, é utilizado com restrição e, portanto, possui baixo tráfego de dados.

- **Subsistema de Rádio Troncalizado**

É utilizado para a transmissão de voz em operações de não guerra, ou seja, em áreas urbanas e, geralmente, próximas às instalações de OMs do EB. Necessita da utilização de infraestrutura, como antenas de retransmissão, do próprio EB ou de outras órgãos governamentais ou privados que apoiam a operação. Possui grande alcance, alta mobilidade e grande volume de tráfego.

- **Subsistema de Dados**

O Subsistema de Dados não é um subsistema físico e sim um subsistema lógico e utiliza a infraestrutura física do Subsistema Rádio de Combate ou do Subsistema Rádio Troncalizado, dependendo do tipo da operação. É utilizado para a transmissão de dados em operações tanto de guerra quanto de não guerra. Este subsistema visa dar suporte ao acompanhamento das operações em tempo real, preconizado pelas atividades de comando e controle. A maioria do seus requisitos variam em funções do tipo de operação no qual está sendo utilizado.

Em operações de guerra, possui alcance, mobilidade e facilidade de instalação semelhantes ao Subsistema Rádio de Combate. Em operações de não guerra, estes requisitos são semelhantes ao Sistema de Rádio Troncalizado. Entretanto, em virtude deste subsistema dar suporte as atividade de comando e controle, possui nos dois tipos de operações grande volume de dados.

A infraestrutura do SISTAC também varia em função da operação, principalmente da área geográfica desta. Os três subsistemas do SISTAC geralmente, possuem a mesma arquitetura. A figura 7 mostra o *backbone* principal do SISTAC em uma operação. Nesta figura, os retângulos representam as unidades militares, a linhas retas, os *links* de comunicações e os círculos, os elementos de comutação da rede.

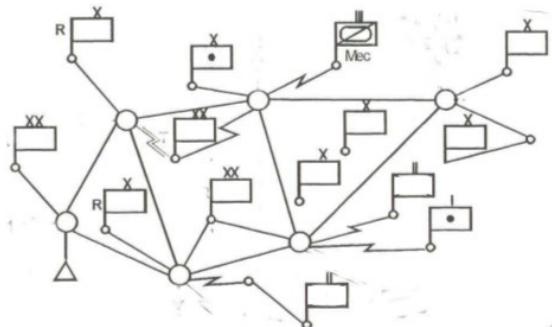


Fig. 7 – *Backbone* Principal do SISTAC em uma Operação.

3.4 Sistema de comunicações do Exército Brasileiro em operações

Numa operação de guerra, os componentes do SEC são expandidos para alcançar a estrutura centralizada de acompanhamento, citada na Subseção 2.1. Os elementos que realizam suas missões em áreas afastadas desta estrutura utilizam o SISTAC para manter as comunicações com a citada estrutura.

Numa operação de não guerra, como por exemplo os Jogos Olímpicos do Rio de Janeiro, realizados em 2016, o SEC funciona como o *backbone* central da estrutura de comunicações e os diversos Sistemas Táticos são interligados ao primeiro. O Sistema Estratégico também é interligado à outras redes de comunicações que possuem informações de interesse para os decisores militares que comandam a operação, por exemplo, redes dos órgãos de segurança pública, redes que possuem acesso às câmeras de vídeo dos órgãos responsáveis pelo gerenciamento do trânsito na cidade etc. Os militares em posições afastadas utilizam o SISTAC para manter as comunicações como o citado *backbone*.

A **tabela 1** faz um resumo da utilização do SEC e do SISTAC nas Operações de Guerra e de Não Guerra.

Tab. 1 – Resumo da utilização do SEC e do SISTAC nas operações do EB.

Sistema	Operação de Guerra	Operação de Não Guerra
SEC	Expandido para alcançar a estrutura centralizada de acompanhamento	Sistema principal
SISTAC	Sistema principal	Utilizado pelos militares em posições afastadas

Dois conjuntos de requisitos a serem considerados na construção das redes de comunicações militares, tanto como infraestrutura do SEC quanto do SISTAC, são os Princípios de Emprego das Comunicações no Exército Brasileiro e os Requisitos Operacionais do Sistema de Comando e Controle desta instituição. Estes requisitos serão explicados nas duas próximas subseções.

3.5 Princípios de emprego das comunicações no Exército Brasileiro

As operações militares, tanto de guerra como de não guerra, se valem dos sistemas de comunicações estratégico e tático listados na seção anterior. Independentemente do tipo de operação ou da modalidade do sistema empregado, é natural que tais soluções devam atender a requisitos de comunicações diversos e desafiadores a um só tempo. O EB apresenta uma coletânea com estes requisitos no O C11-1 - Manual de Campanha - EMPREGO DAS COMUNICAÇÕES [9], documento que define os “Princípios de Emprego das Comunicações Militares”, listados na **tabela 2**.

Tab. 2 – Princípios de Emprego das Comunicações Militares [9].

Princípio	Descrição
Tempo integral	Operar 24 horas por dia, todos os dias da operação.
Rapidez	Colocar a rede em operação em um curto período de tempo com o objetivo de proporcionar, o mais breve possível, o estabelecimento de contato entre os integrantes da operação.
Amplitude de desdobramento	Estar operacional em toda a área da operação.
Integração	Operar junto com os sistemas dos escalões superior e subordinado.
Flexibilidade	Adequar-se rapidamente às mudanças, principalmente ao deslocamento geográfico das organizações militares durante a operação.
Apoio em profundidade	O escalão superior, mais recuado, deve apoiar os escalões subordinados, mais avançados.
Continuidade	Manter as comunicações, a qualquer custo, durante todo o transcurso da operação.
Confiabilidade	Estar sempre disponível, estabelecendo caminhos alternativos para a transmissão das mensagens.
Emprego centralizado	Concentrar meios em centros e eixos de comunicações, permitindo melhor aproveitamento dos mesmos.

Princípio	Descrição
Apoio cerrado	Encurtar as distâncias sempre que possível para facilitar as comunicações.
Segurança	Impedir, ou pelo menos dificultar, a obtenção da informação pelo inimigo.
Prioridade	Estabelecer comunicações e transmitir mensagens de acordo com a prioridade preestabelecida.

Os Princípios de Emprego das Comunicações Militares são um conjunto de diretrizes para que as redes de comunicações militares possam estabelecer as ligações entre os diversos níveis hierárquicos, visando que todos os elementos integrantes de uma operação militar possam, durante todo o desenrolar da operação, manter as comunicações com seus superiores e subordinadas de maneira eficiente, eficaz e segura, condição determinante para o sucesso de uma operação.

3.6 Requisitos operacionais do sistema de comando e controle do Exército Brasileiro

Um dos principais sistemas utilizado pelo EB, tanto na execução das operações quanto no planejamento destas atividades, é o **SC2Ex**. Portanto, as redes de comunicações militares devem atender aos requisitos deste sistema que visa automatizar as tarefas necessárias para atingir os Objetivos 1, 2 e 3, citados na Seção 2.

O **SC2Ex** é uma ferramenta de integração de informações entre os componentes da operação. Este sistema visa incrementar a capacidade de compartilhamento de informações e, também, permitir o adequado fluxo de ordens.

O EB definiu os ROBs (Requisitos Operacionais Básicos) do SC2Ex para os níveis Brigada e Divisão na PORTARIA Nº 032 - EME, DE 19 DE MAIO DE 2003 [10]. Estes requisitos são divididos em três categorias: 1) Absolutos, devem, obrigatoriamente, ser atendidos; 2) Desejáveis, podem ser atendidos mas não obrigatoriamente; e 3) Complementares, podem não ser atendidos. Ao todo são 21 requisitos Absolutos, 29 requisitos Desejáveis e 3 requisitos Complementares.

Devido a sua extensão, a lista completa dos ROBs foi

omitida deste documento, mas encontrada na Portaria Nº 032 - EME, de 2003.

Agora que já conhecemos as redes de comunicações, os Princípios de Emprego das Comunicações e os Requisitos Operacionais do Sistema de Comando e Controle do Exército Brasileiro, faremos uma análise da utilização destas redes nas operações do EB visando atender os Princípios de Emprego das Comunicações e os Requisitos Operacionais do Sistema de Comando e Controle desta instituição.

4. Análise das redes de comunicações atuais do Exército Brasileiro

Alguns trabalhos na literatura tratam do assunto de redes de comunicações táticas. O livro de G. F. Elmasry, intitulado *Tactical wireless communications and networks: design concepts and challenges* [6], é uma obra de referência no tema. Neste texto, o autor fornece uma descrição completa das comunicações militares táticas e tecnologia de redes modernas. Este livro sistematicamente compara as técnicas de comunicações militares táticas com suas equivalentes comerciais, apontando semelhanças e diferenças. Em particular, ele examina cada camada da pilha de protocolos e mostra como requisitos táticos e de segurança específicos resultam em alterações da abordagem comercial. O autor sistematicamente conduz os leitores por esse tópico complexo, primeiramente fornecendo informações sobre a abordagem arquitetônica na qual a análise será baseada, e depois entrando em detalhes sobre as tecnologias e técnicas de comunicação sem fio e de redes táticas.

Outra referência neste assunto é o livro de M. J. Ryan e M. R. Frater, intitulado *Tactical communications for the digitized battlefield* [11]. Neste trabalho é explicado que os sistemas tradicionais de comunicações táticas consistem em vários subsistemas separados, compouca integração entre eles e com os sensores externos e sistemas de armas. Os autores citam que a rede rádio de combate fornece as comunicações de alta mobilidade exigidas pelas tropas no ambiente operacional, enquanto os sistemas de comunicações troncalizados forneceram comunicações de alta capacidade, entre

as tropas em ambientes com infraestrutura instalada, em detrimento da mobilidade. O foco deste livro está nas tecnologias da era da informação que prometem oferecer integração contínua de dados em tempo real, criando uma arquitetura de rede lógica única para facilitar a movimentação de dados por todo o espaço de batalha. Como a estrutura dessa rede é limitada pelo compromisso fundamental entre alcance, mobilidade e capacidade que se aplica a todos os sistemas de comunicações, é improvável que essa rede seja baseada em uma única tecnologia de rede. Este livro apresenta uma arquitetura para essa rede e mostra como seus subsistemas podem ser integrados para formar uma única rede lógica.

O artigo de K. Wongthavarawat e A. Ganz, intitulado *IEEE 802.16 based last mile broadband wireless military networks with quality of service support* [12] cita que os futuros conceitos de guerra do DoD (*Department of Defense* ou Departamento de Defesa) dos Estados Unidos alavancam a superioridade das informações e exigirão grandes melhorias na transferência de informações em termos de maior largura de banda, suporte a QoS (*Quality of Service* ou Qualidade de Serviço) e conexão a um *backbone* de alta velocidade.

Outro artigo que merece ser citado é o de J. C. Juarez, A. Dwivedi, A. R. Hammons, S. D. Jones, V. Weerackody e R. A. Nichols, intitulado *Free-space optical communications for next-generation military networks* [13]. Neste texto, o autor cita a necessidade contínua de maior capacidade para aplicações militares, especialmente em conceitos operacionais centrados em rede que promovem o uso de informações como fundamental para obter superioridade no campo de batalha. Como exemplo, é citado que o acesso e a distribuição de dados de sensores é um princípio importante da guerra centrada na rede e, no entanto, os links de radiofrequência terão dificuldades para fornecer a capacidade necessária. Neste artigo, ainda é enfatizado que, apesar das comunicações ópticas de espaço livre terem o potencial de atender a essas necessidades militares emergentes, oferecendo

aumentos drásticos de capacidade, existem muitos desafios técnicos em várias camadas da pilha de protocolos de comunicação para a exploração deste potencial.

Um artigo que trata especificamente das redes de comunicações do EB é de R. M. Salles, D. F. C. Moura, J. M. A. Carvalho e M. R. Silva, intitulado *Novas perspectivas tecnológicas para o emprego das comunicações no exército brasileiro* [1]. Este texto mostra uma visão geral sobre a estrutura das comunicações militares no EB e como técnicas atuais de redes de dados podem contribuir em sua evolução. Devido ao seu foco deste trabalho nas redes de comunicações do EB, abordando os Princípios de Emprego das Comunicações e os Requisitos do Sistema de Comando e Controle desta instituição, o utilizamos como referência para nossa análise.

Esta análise é explicada nesta seção.

4.1 Análise dos princípios gerais das redes de comunicações do Exército Brasileiro

Os autores de [1] primeiro definem cinco outros termos utilizados comumente na concepção de redes de comunicações, os Princípios Gerais, conforme mostrado na **tabela 3**, para fazer um estudo dos Princípios de Emprego das Comunicações e os Requisitos do Sistema de Comando e Controle do EB. Esses Princípios Gerais dão uma visão dos requisitos técnicos de uma rede de comunicações.

Tab. 3 – DESCRIÇÃO DOS PRINCÍPIOS GERAIS [1].

Princípio Geral	Descrição
Escalabilidade	Capacidade de um sistema expandir-se sem degradar-se.
Desempenho	Existem várias aspectos associados ao desempenho. Na Tabela 4 são tratados os de maior relevância para a maioria das aplicações.
Segurança	Capacidade de proteger os dados da rede de usuários não autorizados.
Gerenciabilidade	Capacidade de levantar-se os parâmetros da rede para suas utilização como base para ações preventivas.
Usabilidade	Facilidade de utilização e configuração dos elementos da rede.

Tab. 4 – Principais Aspectos Associados ao Desempenho [1].

Aspecto	Descrição
Capacidade	Tráfego, em bps, que a rede é capaz de disponibilizar.
Utilização	Percentual médio da capacidade usada.
Utilização Máxima	Percentual da utilização em que a rede é considerada saturada.
Vazão	Quantidade de dados úteis transferidos sem erro por segundo.
Acurácia	Quantidade de tráfego útil corretamente transmitido, relativo ao tráfego total.
Eficiência	Quantidade de dados úteis transmitidos, descontados os dados utilizados para controle das diversas funções da rede.
Atraso ou Latência	Tempo médio entre o momento em que uma mensagem está pronta para ser transmitida e sua recepção no destino.
Variação de Atraso	Percentual de variação no atraso médio.
Tempo de Resposta	Tempo entre um pedido de serviço e a recepção de uma resposta.
Disponibilidade	Razão entre o tempo disponível e o tempo de funcionamento.
Redundância	Capacidade da rede de funcionar plenamente mesmo que alguns de seus recursos estejam indisponíveis.
Prioridade	Capacidade da rede de atender mensagens prioritárias.

Em seguida, o trabalho de [1] mapeia os princípios citados na **tabela 2** nos Princípios Gerais. Esse mapeamento é mostrado na **tabela 5**. O mapeamento dos Princípios de Emprego das Comunicações nos Princípios Gerais simplifica a análise e permite uma avaliação mais direta das tecnologias disponíveis quanto às suas aplicações nos subsistemas de comunicação do Exército.

Tab. 5 – mapeamento dos Princípios de Emprego das Comunicações em Princípios Gerais [1].

Princípio Geral	Princípio(s) de Emprego das Comunicações
Escalabilidade	Amplitude de desdobramento e Integração.
Desempenho	Tempo integral, Rapidez, Confabilidade, Continuidade e Prioridade.
Segurança	Segurança..
Gerenciabilidade	Apoio em profundidade, Emprego centralizado e Apoio cerrado.
Usabilidade	Flexibilidade.

Observa-se, pela **tabela 5**, que todos os Princípios de Emprego das Comunicações do EB podem ser mapeados para um Princípio Geral. Nota-se também que Desempenho e Escalabilidade são os Princípios Gerais que tem mais Princípios de Emprego mapeados e que todos os Princípios Gerais possuem, pelo menos, um Princípio de Emprego das Comunicações mapeado.

4.2 Análise dos requisitos do Sistema de Comando e Controle do Exército Brasileiro

Numa continuação da análise, o texto de [1] faz uma associação dos ROBs do Sistema de Comando e Controle do EB, com os Princípios Gerais vistos anteriormente na **tabela 3**. É importante ressaltar que os ROBs mais intimamente ligados a requisitos de software foram associados ao desempenho (já que a rede deve dar suporte à funcionalidade) e/ou à usabilidade.

Foram dados pesos para cada atributo em virtude de sua categoria: atributos da categoria Absolutos têm peso 10, desejáveis, peso 6 e Complementares, 3.

Dos 21 requisitos Absolutos (peso 10) definidos pelo EB, o citado trabalho classificou 12 quanto ao princípio da Usabilidade, 9 quanto ao Desempenho, 3 quanto à Escalabilidade, 2 quanto à Segurança e 1 quanto à Gerenciabilidade. Observou-se que um único requisito pode atender a mais de um princípio. Dos 29 requisitos Desejáveis (peso 6), classificou 14 quanto ao Desempenho, 11 quanto à Usabilidade, 7 quanto à Escalabilidade, 4 quanto à Gerenciabilidade e 1 quanto à Segurança. Dos três requisitos Complementares (peso 3), classificou um quanto ao Desempenho, um quanto à Usabilidade, um quanto à Escalabilidade e um quanto à Segurança.

Diante da classificação, o trabalho [1] que avaliou o grau de importância de cada Princípio Geral, bastando para isso somar os pesos obtidos nos três tipos de requisitos. Os pesos e percentual dos pesos para os princípios utilizados comumente na concepção de redes de comunicações estão listados na **tabela 6**.

Tab 6 – Peso e Peso Percentual dos Princípios Gerais nas Redes de Comunicações do EB [1].

Princípio Geral	Peso	Peso Percentual
Escalabilidade	189	37,4%
Desempenho	177	35,1%
Segurança	75	14,9 %
Gerenciabilidade	34	6,7 %
Usabilidade	29	5,8 %

Estes pesos dão uma visão da importância de cada Princípio Geral para as redes de comunicações militares. Pelo resultado da **tabela 6**, podemos notar que escalabilidade e desempenho são os Princípios Gerais com maiores pesos. Entretanto, os cinco Princípios Gerais têm peso considerável (maior que 5%) nas redes de comunicações militares.

De posse dos resultados da importância de cada Princípio Geral para as redes de comunicações militares, faremos um estudo do SISTAC com o objetivo de levantar os desafios tecnológicos para que as redes de comunicações militares atendam aos requisitos operacionais do EB.

Esta opção pelo SISTAC, em detrimento do SEC, foi motivada pelos desafios inerentes às redes táticas. Dentre estes, destacamos 1) interconexão de tecnologias de enlace, 2) aumento da segurança na camada física e 3) aumento do desempenho em cenários de alto tráfego ou de interferência eletromagnética. Explicamos estes desafios na Seção 6 na qual também fazemos propostas para enfrentá-los.

Na próxima Seção, faremos um estudo de algumas tecnologias quanto ao possível emprego no SISTAC.

5. Tecnologias existentes para o emprego nas redes de comunicações do Exército Brasileiro

Para efeito deste trabalho, dividimos as tecnologias de comunicações em dois grupos: 1) Tecnologias prontas

para uso e 2) Tecnologias em Desenvolvimento. As tecnologias detes grupos serão descritas resumidamente nas próximas duas subseções.

5.1 Tecnologias prontas para uso

Segundo Salles [1], algumas tecnologias candidatas a um possível emprego no SISTAC são as seguintes:

- **Rede Rádio de Combate com Transmissão Digital**

É um sistema de radiodifusão, que se distingue pela emissão de um sinal digital. Utiliza a rede rádio legada para transmissão de dados. Pode trabalhar em HF (*High Frequency* ou Frequência Alta), com taxas aproximadamente 2,4 Kbps, ou em UHF/VHF, com taxas de 16 Kbps, utilizando-se de modems específicos. A principal vantagem do rádio digital está na melhoria da qualidade da transmissão. A digitalização do rádio e a integração com outras tecnologias de transmissão de dados oferecem também uso mais eficiente do espectro. Fabricantes como a *Harris Corporation*® e a *Thales Communications*® possuem produtos disponíveis no mercado que utilizam esta tecnologia.

- **Rede Pacote Rádio**

É uma tecnologia que aumenta as taxas de transferência de dados nas redes rádios existentes. Esta permite o transporte de dados por pacotes utilizando comutação por pacotes. Sendo assim, este sistema oferece uma taxa de transferência de dados muito mais elevada que as taxas de transferência que a tecnologia de Rede Rádio de Combate com Transmissão Digital. Este sistema, em situações ideais, pode ultrapassar a marca dos 170 kbps. No entanto na prática, esta taxa está em torno dos 40 kbps.

Este tecnologia utiliza a mesma estrutura convencional das Redes Rádio de Combate, interligando diversas subredes através de máquinas *bridges*, proporcionando *re-broadcast*. Produtos dos fabricantes que utilizam a tecnologia de rede rádio de combate com transmissão digital também disponibilizam esta

tecnologia de rede pacote rádio.

- **Redes de Arquitetura Estação-Base**

Fazem uso de uma estação base para intermediar a comunicação entre quaisquer duas estações. Estas estações-base permitem um maior alcance destas redes. Entretanto, a instalação destas estações é uma atividade que requer tempo e recursos. Portanto, são de difícil utilização em locais desconhecidos e/ou hostis. Entre as redes mais comuns que se utilizam desta arquitetura estão: telefones celulares e rádios *half-duplex* de dupla frequência. Fabricantes como a *Ericson GE*® e a *Motorola*® possuem produtos disponíveis no mercado que utilizam esta tecnologia.

- **Redes Ad Hoc**

Redes *ad hoc* são um tipo de rede que não possuem um nó ou terminal especial, geralmente designado como ponto de acesso, para o qual todas as comunicações convergem e que as encaminha para os respectivos destinos. Assim, uma rede de computadores *ad hoc* é aquela na qual todos os terminais funcionam como roteadores, encaminhando de forma comunitária as comunicações advindas dos terminais vizinhos. A maleabilidade de sua topologia e a ausência de necessidade de infraestrutura prévia permitem um rápido desenvolvimento de uma rede *ad hoc*. Nenhuma preparação da área de cobertura é preciso porque itens como torres ou linhas de transmissão não são necessários. Estes atributos das redes *ad hoc* as diferenciam das redes do tipo estação-base, nas quais a conectividade é conhecida *a priori*.

Por estes motivos, elas são adequadas às comunicações próprias dos campos de batalha, onde, na maioria das vezes, nenhuma infraestrutura prévia está disponível. Os pontos chaves dessas redes são o controle de acesso ao meio, utilizado para compartilhar os recursos de canal entre os nós, e o roteamento, que consiste em encontrar uma rota entre emissor e receptor através de um número desconhecido de nós intermediários. As taxas variam com a tecnologia sem fio empregada. As empresas *rede-livre Freifunk*®, da Alemanha, e a *openwireless*®, da Suiça, desenvolveram o protocolo B.A.T.M.A.N. (Better Approach To Mobile

Ad-hoc Networking ou Melhor Abordagem para Redes Ad-hoc Móveis) de roteamento para redes WLAN de computadores abertas redes sem fios do tipo Ad-hoc.

• TDMA Repetido

O TDMA (*Time Division Multiple Access* ou Acesso Múltiplo por Divisão de Tempo) é um sistema rádio digital que funciona dividindo um canal de frequência em até oito intervalos de tempo distintos. Cada usuário ocupa um *slot* de tempo específico na transmissão, o que impede problemas de interferência. A temporização é determinada por uma estação de controle. Alcançam taxas de aproximadamente 500 Kbps. Os sistemas celulares de segunda geração, como o GSM (*Global System for Mobile Communications* ou Sistema Global para Comunicações Móveis), utilizam o TDMA na sua interface com a estação móvel.

• Redes Rádio UWB

O UWB (Ultrawideband ou Banda Ultra Larga) opera de forma diferente das demais tecnologias rádio. Além de ter um espectro de atuação amplo, o UWB transmite por rajadas de sinais (centenas por segundo). A combinação do “gatilho rápido” com a ampla cobertura de banda permite que o UWB consuma menos energia e consiga taxas de transmissão mais altas.

Opera na faixa de 3,1 GHz a 10,6 GHz. O padrão de transmissão sem fio UWB utiliza sinais de rádio de baixa energia na forma de pulsos curtos (0,1 a 1,5 ns). Com a geração de milhões de pulsos por segundo, o UWB pode atingir taxas de transmissão de até 100 Mbps, entretanto seu alcance é, no máximo, 10 m. Aparelhos sem fios próximo ao usuário, como por exemplo, impressora, mouse, teclado ou MP3 Player utilizam esta tecnologia de comunicações.

Uma descrição mais detalhada das tecnologias acima citadas pode ser encontrada no livros de A. S. Tanenbaum, intitulado *Redes de Computadores* [14] e de A. F. Molisch, intitulado *Wireless Communications* [15].

Duas tecnologias promissoras para sua utilização em redes táticas e, consequentemente, para utilização

no SISTAC, são as de rádios cognitivos e a de transmissão *full-duplex* em Banda. Essas tecnologias são descritas na duas próximas subseções.

5.2 Tecnologias em desenvolvimento

• Rádios Cognitivos

Um RC (Rádio Cognitivo), termo cunhado pelo Prof. Dr. Joseph Mitola em 1999 [16], de maneira simples, é um rádio inteligente [17]. Conforme definição de Haykin [18], é um sistema de comunicações sem fio inteligente, capaz de utilizar metodologias de inteligência artificial, tais como *reinforcement learning* [19], algoritmos genéticos [20], entre outras, para adaptar seus parâmetros operacionais à dinâmica do ambiente ao seu redor. Por consequência, é um tipo de rádio que aprende com experiências anteriores e lida com situações que não foram planejadas em sua concepção.

Assim, um RC [21] é um rádio que percebe e está ciente de seu ambiente operacional e pode, dinâmica e automaticamente, ajustar seus parâmetros operacionais de acordo com esse ambiente.

Um RC [22] é definido formalmente como um rádio que pode mudar seus parâmetros de transmissão com base na interação com o ambiente no qual opera. Estes parâmetros podem ser das camadas física (frequência, potência do sinal e estratégias de modulação e de codificação utilizadas), de enlace (técnicas de acesso ao meio e de retransmissão), de rede (algoritmos de roteamento), ou das camadas superiores (estratégia de autenticação e algoritmo de criptografia).

Para que um RC [17] opere de maneira adequada, mais do que simplesmente informações do espectro eletromagnético, ele deve utilizar quatro tipos de informações, todas fornecidas pelo ambiente de operação no qual está imerso: 1) do ambiente rádio; 2) dos requisitos de QoS da aplicação; 3) dos recursos disponíveis para o dispositivo; e 4) da política regulatória do uso do espectro.

O mesmo autor [17] menciona que aplicações militares, tais como sistemas logísticos, acionamento de armas, navegação, geolocalização, radares e redes de sensores são grandes utilizadoras do espectro e de

diferentes tipos de sistemas sem fio. Estas aplicações necessitam de acesso eficiente e seguro ao espectro. Em um cenário militar [17], tipicamente, há um grande número de dispositivos e sistemas de comunicações heterogêneos que precisam ser interconectados. Por exemplo: diferentes sistemas táticos (manobra, guerra eletrônica, defesa antiaérea etc.) de uma mesma força militar, diferentes forças armadas de um país ou mesmo de diferentes países em uma aliança internacional. Pode haver operações em terra, mar e ar.

Ainda pode ocorrer uma mistura de redes centralizadas e descentralizadas. Muitos destes sistemas precisam ser instalados rapidamente em locais desconhecidos e muitas vezes hostis. Pode haver conexões cujo emprego seja de duração curta, por exemplo, conexões entre aeronaves, ou de longa, como sistemas de comunicações entre os presidentes dos países de uma aliança multinacional e o general comandante da operação militar.

A coexistência e a interoperabilidade de redes também é de vital importância no caso de alianças militares e operações conjuntas. Como um RC pode receber um sinal (forma de onda) no padrão de uma rede *A* e transformá-lo no padrão de uma rede *B* para reencaminhamento, tem-se que, neste cenário, um RC pode funcionar como *bridge*, permitindo uma rápida integração entre forças distintas em coalizão. Em casos onde sistemas legados não cognitivos existam, as RC podem evitar problemas, tais como os de interferência, escolhendo faixas de frequência que não estão sendo utilizadas pelos sistemas legados.

• Transmissão Full-Duplex em Banda

Transmissão sem fio FD (Full-Duplex) em banda [23] significa que um dispositivo rádio tem a capacidade de receber e transmitir informações ao mesmo tempo na mesma faixa de frequência. Esta capacidade é habilitada pelas técnicas de SIC (*Self-Interference Cancellation* ou Cancelamento de Auto-Interferência). Nenhuma das tecnologias de redes atuais possuem esta característica. Deve-se notar que nem a TDD (*Time-Division Duplexing* ou Duplexação por Divisão do Tempo) nem a FDD (*Frequency-Division Duplexing* ou

Duplexação por Divisão de Frequência) é considerado operação FD real nos termos modernos, apesar de elas permitirem conversas bidirecionais simultâneas, porque a perspectiva da comunicação FD é deslocada para o uso do espectro na camada física [24].

Agora que temos a importância de cada Princípio Geral para as redes de comunicações militares e do estudo de algumas tecnologias de comunicações, verificaremos as lacunas e desafios tecnológicos para que as redes de comunicações militares atendam aos Princípios de Emprego das Comunicações e aos Requisitos do Sistema de Comando e Controle do EB.

6. Desafios tecnológicos para que as redes de comunicações militares atendam aos requisitos operacionais do Exército Brasileiro

Autores [1] fazem uma avaliação do atendimento aos Princípios Gerais pelas tecnologias prontas para uso descritas na Seção 5. Esta avaliação é mostrada na **tabela 7**.

Tab. 7 – Atendimento dos Princípios Gerais pelas Tecnologias Prontas para Uso Descritas.

Tecnologia	Requisitos Atendidos	Requisitos Não Atendidos
Rede Rádio de Combate com Transmissão Digital	1, 4 e 5	2 e 3
Rede Pacote Rádio	1, 3, 4 e 5	2
Redes de Arquitetura Estação-Base	3, 4 e 5	1 e 2
Redes Ad Hoc	3, 4 e 5	1 e 2
TDMA Repetido	3, 4 e 5	1 e 2
Redes Rádio UWB	2, 3, 4 e 5	1

1-Escalabilidade, 2-Desempenho, 3-Segurança, 4-Gerenciabilidade e 5-Usabilidade.

Os mesmos autores [1] concluem que nenhuma infraestrutura de comunicação reúne condições de suportar os Princípios de Emprego das Comunicações e os Requisitos do Sistema de Comando e Controle do EB simultaneamente. Sendo assim, um sistema de comunicações militar deve ser uma única estrutura

lógica integrada por múltiplas tecnologias de enlace. A integração destas tecnologias devem ser feitas de maneira automática e sem a necessidade de intervenção humana visando a agilidade e a flexibilidade.

Outro aspecto abordado pelos mesmos autores [1] é que três requisitos são estritamente relacionados e constituem um forte compromisso no sistema de comunicações militar: alcance, capacidade e mobilidade. Por exemplo: um sistema que possibilita altas mobilidade e alcance não será dotado de alta taxa de transmissão. O forte relacionamento existente se dá, em geral, pelo fato que ao variar um dos fatores os outros dois também variam. Na verdade, ao se maximizar dois fatores o terceiro acaba sendo minimizado. A **figura 8** exemplifica tal situação.

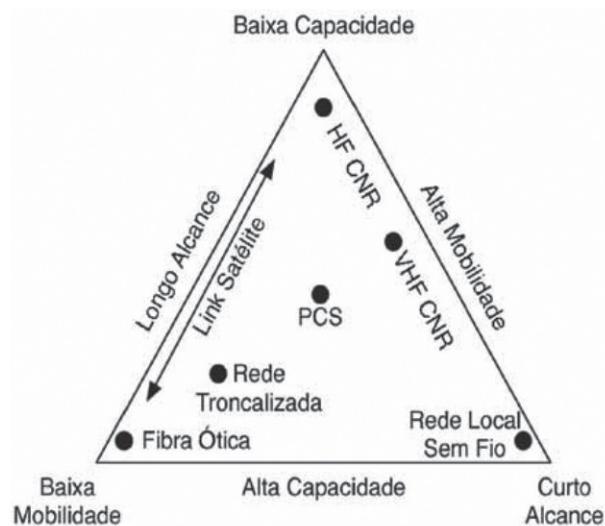


Fig. 8 – Triângulo de Compensação [11].

O Triângulo de Compensação se constitui em uma forma rápida e simples de avaliação da pertinência de uma determinada tecnologia de comunicação quanto ao emprego em uma operação.

Pelo exposto nesta seção, percebemos que as tecnologias de redes de comunicações prontas para uso deixam desafios para o atendimento aos Princípios de Emprego das Comunicações do EB bem como para os Requisitos Operacionais Básicos do Sistema de Comando

e Controle desta instituição. Destacaremos neste artigo, três desafios para que as redes de comunicações militares atendam ao requisitos operacionais do EB. Estes desafios são descritos a seguir.

- **Desafio 1 – Interconexão de tecnologias de enlace**

O principal desafio das redes atuais é prover a possibilidade de realizar, em tempo de operação, sem perda de capacidade, alcance e/ou mobilidade, a interconexão de tecnologias de enlaces utilizadas.

O desafio citado no parágrafo anterior surge pois nenhuma infraestrutura de comunicação reúne condições de suportar os Princípios de Emprego das Comunicações e os Requisitos do Sistema de Comando e Controle do EB simultaneamente. Consequentemente, um sistema de comunicações militar deve ser uma única estrutura lógica integrada por múltiplas tecnologias de enlace. Entretanto, as redes de comunicações que utilizam as tecnologias prontas para uso não tem a possibilidade de integração automática, sem intervenção humana, destas tecnologias.

Como exemplo da impossibilidade citada no parágrafo anterior, imagine a situação em que se dispõe das tecnologias de Redes Rádio UWB e de Rede Rádio de Combate com Transmissão Digital. A primeira tecnologia possui maior taxa de transmissão e menor consumo de energia que a segunda. Entretanto, as Redes Rádios UWB possuem menor alcance. Uma rede lógica ideal seria aquela que, sempre que a distância entre o transmissor e o receptor fosse menor que o alcance das Redes Rádios UWB, utilizaria esta tecnologia e, quando a distância fosse maior que o citado alcance, utilizaria a Rede Rádio de Combate com Transmissão Digital. Esta mudança de tecnologia precisaria ser feita dinâmica e automaticamente.

O mesmo raciocínio do parágrafo anterior aplica-se quanto possuímos as tecnologias de Rede de Arquitetura Estação-Base e, novamente, a Rede Rádio de Combate com Transmissão Digital. A primeira tecnologia possui maior segurança pois a transmissão é feita por linha de visada do rádio com sua estação-base porém requer a instalação das estações-base. Numa operação

em ambiente urbano, a Rede de Arquitetura Estação-Base é a mais utilizada pois o cenário é conhecido e, na maioria das cidades, a infraestrutura necessária já se encontra instalada. Entretanto, esta torna-se inadequada para as operações onde a atuação se dá em ambientes desconhecidos e, muitas vezes, hostis, ou mesmo quando a atuação precisa ser tempestiva com pouquíssimo tempo para a preparação. Nestas operações a Rede Rádio de Combate com Transmissão Digital é a mais adequada.

No exemplo do parágrafo anterior, o problema das redes atuais surge quando, por exemplo, tropas do EB estão em uma operação em ambiente urbano e, de maneira imediata, parte desta tropa precisa ser deslocar para um local desconhecido. Geralmente este problema é resolvido utilizando dois equipamentos rádios. Numa rede lógica ideal teríamos um único equipamento capaz de, quando detectasse a estação-base, faria a transmissão por linha de visada, e quando isto não acontecesse, utilizaria a tecnologia de Rede Rádio de Combate Digital.

- **Desafio 2 – Aumentar a segurança na camada física**

Outro desafio das redes táticas atuais é aumentar a segurança na camada física das redes de comunicações táticas sem fio.

O desafio citado no parágrafo anterior surge pois nas tecnologias prontas para uso, descritas na Seção 5, tradicionalmente, a segurança é abordada nas camadas superiores da pilha de protocolos usando criptografia. Estas abordagem que utiliza criptografia com base em chaves secretas, tem sua garantia de sigilo fundamentada na suposição que os interceptadores tem recursos computacionais limitados. Entretanto os usuários e, consequentemente os interceptadores, possuem grande poder computacional hoje em dia e este poder aumenta de forma continua. Outro desafio desta abordagem é garantir que as chaves não sejam interceptadas por nós não autorizados durante a distribuição destas pela rede.

- **Desafio 3 - Aumentar o desempenho em cenários de alto tráfego ou de interferência eletromagnética.**

Outro desafio das redes atuais é aumentar seu desempenho, especialmente em cenários de alto tráfego ou de interferência eletromagnética.

O desafio citado no parágrafo anterior surge pois, com exceção das redes de rádio UWB, as tecnologias prontas para uso não possuem desempenhos satisfatórios. Esta falta de desempenho, em parte, é causada pelo estratégia de uso do espectro das redes rádio atuais.

As redes rádios atuais utilizam o espectro de maneira fixa, ou seja, utilizam uma ou mais faixa de frequência de maneira pré determinada ou configurada por software. Esta característica obriga que um rádio utilize um ou mais canais ainda que estes estejam sobrecarregados e existam outras canais livres. Outro problema desta abordagem é a facilidade de um ataque de interferência, especialmente quando o atacante possui a tecnologia de rádios cognitivos [25].

6.1 Sugestões de soluções para os desafios tecnológicos das redes de comunicações militares

Duas tecnologias que figuram como candidatas promissoras para enfrentar os desafios citados acima são a tecnologia de rádios cognitivos e a de transmissão *full-duplex* em banda.

Conforme citado [17], um RC pode receber um sinal no padrão de uma *rede A* e transformá-lo no padrão de uma *rede B* para reencaminhamento. Tem-se que, neste cenário, uma RC pode funcionar como *bridge*, permitindo uma rápida integração entre múltiplas tecnologias de enlace. Outro autor [18] cita que, com a utilização de técnicas de *reinforcement learning* [19], algoritmos genéticos [20], um RC pode aprender, de forma automática e sem intervenção humana, a melhor forma de fazer a integração das tecnologias de enlace, dados os requisitos das camadas superiores da rede. Esta capacidade candidata a tecnologia de RC a enfrentar o Desafio 1.

Com o desenvolvimento da tecnologia de transmissão FD em banda [24], surge a capacidade de transmissão e recepção simultâneas. Autores [24] citam que rádios militares com a tecnologia FD poderão:

- a) produzir guerra eletrônica, via ataque de

interferência, contra as comunicações inimigas, ao mesmo tempo que recebem dados, b) produzir guerra eletrônica, via ataque de interferência, contra a interceptação de dados pelo inimigo, ao mesmo tempo que recebem dados, c) interceptar as comunicações inimigas ao mesmo tempo que transmitem dados e d) interceptar as comunicações inimigas e interferir nas comunicações simultaneamente. A capacidade b) descrita acima permite a aumentar a segurança na camada física.

Um RC pode mitigar os efeitos de um ataque de interferência. Dois exemplos de trabalhos, [25] e [26], propõem esta aplicação.

Com a utilização conjuntas das tecnologias de RC e FD, este aumento de segurança pode ser feito sem desconsiderar outros requisitos das redes táticas tais como: 1) restrição de uso de energia; 2) desconhecimento tanto do canal entre o rádio espião inimigo e o transmissor quanto do canal entre o espião e o receptor; e 3) necessidade do aumento da dificuldade de detecção da posição dos rádios aliados pelo inimigo. A junção das duas tecnologias citadas possibilita o aumento da segurança nas transmissões em redes rádios em cenários militares. Estas características colocam as tecnologias de RC e FD como candidatas a enfrentar o Desafio 2.

A tecnologia de rádios cognitivos permite, entre outras funcionalidades, que um rádio: 1) determine quais partes do espectro estão disponíveis (detecção espectral), 2) selecione o melhor canal disponível (decisão espectral), 3) coordene o acesso a este canal com outros usuários (compartilhamento espectral) e 4) desocupe o canal quando perceberem que este se tornou sobrecarregado (mobilidade espectral) [22]. Assim esta tecnologia possibilita a agilidade espectral em cenários de alto tráfego ou de interferência eletromagnética, aumentando o desempenho da rede rádio.

RC pode prover agilidade espectral em cenários de alto tráfego ou interferência eletromagnética devido a sua capacidade de escolher, entre outros parâmetros de transmissão, a melhor faixa de frequência para transmitir. Em utilização junto com a tecnologia de FD, RC pode aumentar significativamente a capacidade de transmissão. Assim, RC é um candidato a enfrentar

o Desafio 3.

Um aspecto que deve ser considerado ao enfrentar estes desafios é necessidade de existência de uma arquitetura peculiar às redes de comunicações militares de modo a solucionar os desafios citados acima e prover os fundamentos para o correto e claro fluxo de informações e a interoperabilidade entre os diversos sistemas.

Consequentemente, a definição de uma arquitetura de redes cognitivas para redes de comunicações militares é um passo importante na utilização tanto da tecnologia de RC quanto de FD para estas redes. Devido a complexidade desta arquitetura, uma estratégia mais segura é propor arquiteturas parciais para as camadas física, de enlace de dados, de redes, de transporte e superiores. Outra abordagem é dividir a arquitetura por requisito ou conjunto de requisitos, por exemplo: desempenho, segurança etc.

Fizemos uma proposta de arquitetura enfatizando o requisito segurança e utilizando a divisão por camadas. Esta proposta é mostrada na **figura 9** - Arquitetura de Rede Cognitiva pra Redes Táticas sem Fios.

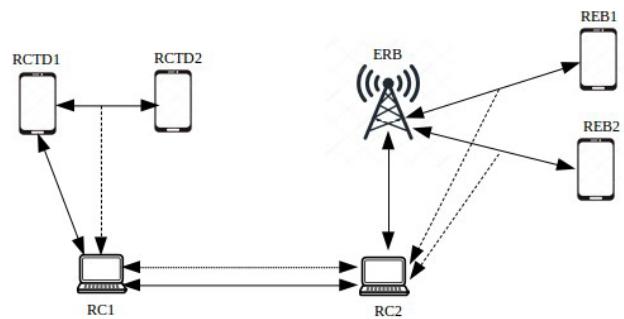


Fig. 9 – Arquitetura de Rede Cognitiva pra Redes Táticas sem Fios.

Nesta proposta, temos duas infraestruturas de redes prontas para uso: uma rede rádio de combate com transmissão digital, à esquerda, formada pelos nós RCTD1 e RCTD2, e uma rede de arquitetura estação-base, à direita, formada pelos nós REB1, REB2 e ERB. Além destas, temos dois nós cognitivos: RC1 e RC2, na parte inferior da figura, formando uma rede cognitiva. Estas redes formam uma única estrutura lógica integrada por múltiplas tecnologias de enlace.

As possíveis transmissões de mensagens entre os nós, na camada de enlace, das redes estão representadas por linha contínuas. Os sensoriamentos do espectro executados pelos nós cognitivos estão indicados pelas linhas tracejadas e as trocas de informações gerenciais entre os nós cognitivos estão representadas pelas linhas pontilhadas.

Uma propriedade importante desta proposta de arquitetura é que ela possibilita o funcionamento das redes prontas para uso sem a necessidade de mudanças nestas, possibilitando assim o aproveitamento das diversas infraestruturas legadas existentes.

A primeira função dos nós cognitivos é realizar, em tempo de operação, sem perda de capacidade, alcance e/ou mobilidade, a troca de tecnologias utilizadas. Como exemplo desta função temos o caso no qual uma aplicação do RCTD1 desejar se comunicar com sua aplicação correspondente no REB2. Nesta hipótese, a rede cognitiva fará o roteamento das mensagens.

Outra função dos nós cognitivos é aumentar a segurança na camada física das redes de comunicações táticas sem fio. Assim, caso o nó RC1 possua a capacidade FD, este produzirá, ao mesmo tempo que recebe dados, guerra eletrônica, via ataque de interferência, como uma medida contra uma possível interceptação de dados pelo inimigo.

Os nós cognitivos também tem a função de aumentar o desempenho da redes rádios, especialmente em cenários de alto tráfego ou de interferência eletromagnética. Para executar esta, os nós cognitivos, graças a sua capacidade de sensoriamento do espectro, determinarão quais partes do espectro estão disponíveis e selecionarão o melhor canal para transmitir, assim como outros parâmetros de transmissão para um desempenho ótimo.

Para realizar as funções citadas acima, os nós cognitivos precisarão de uma arquitetura que permita a comunicação das diversas camadas da rede com uma Máquina Cognitiva. Um exemplo desta arquitetura é mostrado na figura 10.

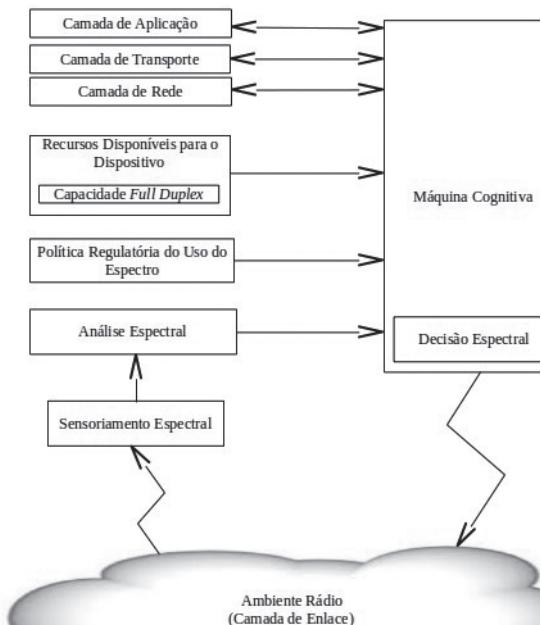


Fig. 10 – Arquitetura para Nós Cognitivos em Redes Táticas Sem Fios.

A Máquina Cognitiva é o elemento principal desta arquitetura. Esta é responsável pela Decisão Espectral, na qual escolhe os parâmetros mais adequados para a transmissão, dados pelos quatro tipos de informação [17]: 1) do ambiente rádio; 2) dos requisitos de QoS da aplicação; 3) dos recursos disponíveis para o dispositivo; e 4) da política regulatória do uso do espectro.

Nossa proposta de arquitetura encontra-se em fase de estudo de viabilidade. Estamos aplicando-a em nossos estudos sobre mecanismos que aumentem a segurança na camada física das redes de comunicações táticas sem fio e, paralelamente, atendam aos requisitos e características destas redes.

7. Conclusões e proposta de trabalhos futuros

Conforme exposto [1], nenhuma das tecnologias de redes de comunicações prontas para uso reúne condições de suportar todos os Princípios de Emprego das Comunicações do EB bem como os Requisitos Operacionais Básicos do Sistema de Comando e Controle desta instituição. Por

exemplo, um sistema que possibilita altas mobilidades e alcances não será dotado de altas taxas de transmissão.

Outro aspecto abordado [1] é que três requisitos são estreitamente relacionados e constituem um forte compromisso no sistema de comunicações militar: alcance, capacidade e mobilidade. O forte relacionamento existente se dá, em geral, pelo fato que ao variar um dos fatores os outros dois também variam. Na verdade, ao se maximizar dois fatores o terceiro acaba sendo minimizado. Este aspecto ilustra uma lacuna das redes não cognitivas: a necessidade de automaticamente utilizar duas ou mais tecnologias para que possa maximizar, numa mesma situação, o alcance a capacidade e a mobilidade.

As tecnologias de rádios cognitivos e a de transmissão *full-duplex* em banda são propostas para enfrentar os desafios citados neste trabalho visando maximizar, ao mesmo tempo, o alcance, a capacidade e a mobilidade das redes de comunicações militares.

Um aspecto que deve ser considerado ao enfrentar estes desafios é necessidade de existência de uma arquitetura

peculiar às redes de comunicações militares de modo a solucionar os desafios citados acima e prover os fundamentos para o correto e claro fluxo de informações e a interoperabilidade entre os diversos sistemas.

Consequentemente, a definição de uma arquitetura de redes cognitivas para redes de comunicações militares é um passo importante na utilização tanto da tecnologia de RC quanto de FD para estas redes. Devido a complexidade desta arquitetura, uma estratégia mais segura é propor arquiteturas parciais para as camadas física, de enlace de dados, de redes, de transporte e superiores. Outra abordagem é dividir a arquitetura por requisito ou conjunto de requisitos, por exemplo: desempenho, segurança etc.

Outro aspecto a ser abordado é que o enlace, estudado neste trabalho, é apenas um dos componentes dos sistemas de comunicações. Aspectos relacionados às outras camadas de um sistema de comunicações devem ser investigados para que as redes de comunicações militares possam efetivamente atender aos seus objetivos, conforme preconizado pelo Exército.

Referências bibliográficas

- [1] Salles, R. M.; Moura, D. F. C.; Carvalho, J. M. A.; Silva, M. R.; Novas perspectivas tecnológicas para o emprego das comunicações no exército brasileiro; Revista Militar de Ciência e Tecnologia 2008, XXV, 68.
- [2] Brasil; Constituição da República Federativa do Brasil, Senado Federal, Brasília, 1988.
- [3] Brasil; Estado Maior do Exército; EB20-MC-10.202 - Manual de Campanha - FORÇA TERRESTRE COMPONENTE, Estado Maior do Exército, Brasília, 2014.
- [4] Brasil; Estado Maior do Exército; EB20-MF-10.103 - Manual de Fundamentos - OPERAÇÕES, Estado Maior do Exército, Brasília, 2015.
- [5] Brasil; Estado Maior do Exército; EB20-MF-10.1205 - Manual de Campanha - COMANDO E CONTROLE, Estado Maior do Exército, Brasília, 2015.
- [6] Elmasry, G. F.; *Tactical wireless communications and networks: design concepts and challenges*, John Wiley & Sons, West Sussex, 2012.
- [7] Brasil; Estado Maior do Exército; C-11.61 - Manual de Campanha – COMUNICAÇÕES NA DIVISÃO DE EXÉRCITO, Estado Maior do Exército, Brasília, 1995.
- [8] Brasil; Exército Brasileiro; Departamento de Ciência e Tecnologia; Centro Integrado de Telemática do Exército; <http://www.citex.eb.mil.br/>, acessada em: dezembro de 2018.
- [9] Brasil; Estado Maior do Exército; C-11.1 - Manual de Campanha – EMPREGO DAS COMUNICAÇÕES, Estado Maior do Exército, Brasília, 1997.
- [10] Brasil; Estado Maior do Exército; PORTARIA Nº 032 - EME, DE 19 DE MAIO DE 2003; Estado Maior do Exército, Brasília, 2003.
- [11] Michael, R. J.; Frater, M. R.; *Tactical communications for the digitized battlefield*. Artech House, Norwood, 2002.
- [12] Wongthavarawat, K.; Ganz, A.; IEEE 802.16 based last mile broadband wireless military networks with quality of service support; *Military Communications Conference 2003*, 2, 779.
- [13] Juarez, J. C. ; Dwivedi, A.; Hammons, A. R.; Jones, S. D.; Weerackody, V.; Nichols, R. A.; Free-space optical communications for

- next-generation military networks; IEEE Communications Magazine 2006, 44, 11.
- [14] Tanenbaum, A. S.; Redes de Computadores; Editora Campus, São Paulo, 2003.
- [15] Molisch, A. F; Wireless Communications; John Wiley & Sons, Nova Jersey, 2012.
- [16] Mitola, J.; Cognitive Radio: Model-Based Competence for Software Radios; Tese de Doutorado, Instituto Real de Tecnologia, Estocolmo, 1999.
- [17] Doyle, L.; Essentials of cognitive radio; Cambridge University Press, Cambridge, 2009.
- [18] Haykin, S.; Cognitive radio: brain-empowered wireless communications; IEEE Journal on Selected Areas in Communications 2005, 23, 201.
- [19] Hu J.; Wellman, M.; Multiagent reinforcement learning: Theoretical framework and an algorithm; 15th International Conference on Machine Learning 1998, 15, 242.
- [20] Linden, R; Algoritmos Genéticos; Brasport, Rio de Janeiro, 2006.
- [21] Reed, J.; Bostian, W.; Understanding the issues in software defined cognitive radio; IEEE International Symposium on Dynamic Spectrum Access Networks 2005, 5, 611.
- [22] Akyildiz, I. F.; Lee, W.; Vuran, M. C.; Mohanty, S.; Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey; Computer Networks Journal 2006, 50, 2127.
- [23] Sabharwal, A.; Schniter, P.; Guo, D.; Bliss, D. W.; Rangarajan, S.; Wichman, R.; In-band full-duplex wireless: Challenges and opportunities. IEEE Journal on Selected Areas in Communications 2014, 32, 1637.
- [24] Riihonen, T.; Korpi, D.; Rantula, O.; Valkama, M.; On the prospects of full-duplex military radios; Military Communications and Information Systems International Conference 2017, 16, 1.
- [25] Wang, B.; Wu, Y.; Liu, K. R.; Clancy, T. C.; An anti-jamming stochastic game for cognitive radio networks; IEEE Journal on Selected Areas in Communications 2011, 29, 877.
- [26] Camilo, M.; Moura, D.; Galdino, J.; Salles, R. M.; Anti-jamming defense mechanism in cognitive radios networks; Military Communications Conference 2012, 11, 1.

Dimensionamento do reforço ao cisalhamento com compósitos de resina e fibras em vigas-parede de concreto

Fábio G. Ávila*, Ana M. A. J. Abreu, Luiz A. V. Carneiro

Instituto Militar de Engenharia (IME)

Praça General Tibúrcio, 80, 22290-270, Praia Vermelha, Rio de Janeiro, RJ, Brasil.

*fabiogavila@gmail.com

RESUMO: Este trabalho apresenta estudos encontrados na literatura sobre o ensaio de vigas-parede de concreto armado reforçadas com compósitos de resina e fibras (FRP) colados. A partir dos resultados de ensaios destes estudos, propõe-se uma metodologia para o dimensionamento do reforço ao cisalhamento de vigas-parede com FRP. Constatou-se que o tipo, a configuração e taxa de armadura de reforço, a resistência do concreto à compressão, a relação entre o vão de cisalhamento e a altura útil e a taxa de armadura interna das vigas são os parâmetros mais influentes na resistência das vigas-parede de concreto armado reforçadas com FRP colados.

PALAVRAS-CHAVE: Reforço. Vigas-parede. Concreto armado. Fibras de carbono.

ABSTRACT: This paper presents some studies about tests of reinforced concrete deep beams strengthened with fiber reinforced plastics (FRP) bonded encountered in literature. From test results of these studies, design methodology is proposed for shear strengthening of deep beams with bonded FRP. It was found that the type, configuration and reinforcement ratio of FRP, the compressive strength of concrete, the shear span ratio and the internal reinforcement ratio of beams are the most influential parameters in the strength of strengthened reinforced concrete deep beams with bonded FRP.

KEYWORDS: Strengthened. Deep beams. Reinforced concrete. Carbonfibers

1 . Introdução

Vigas-parede são vigas com sua altura grande o suficiente de modo que não são consideradas elementos lineares. A hipótese de Bernoulli não se aplica no caso de vigas-parede, pois as seções não permanecem planas sob a ação das cargas, mesmo para um material perfeitamente elástico [1]. Vigas-parede são as vigas altas em que a relação entre o seu vão e a sua altura (L/h) é inferior a 2, em vigas biapoiadas, e inferior a 3, em vigas contínuas [2].

Entretanto, não há consenso entre as principais normas internacionais sobre os limites da razão L/h que diferenciam as vigas e as vigas-parede

[3]. Na Inglaterra, vigas-parede são definidas como as vigas com esta razão menor que 2 para vigas biapoiadas e menor que 2,5 para vigas contínuas [4]; no Canadá, o limite é 1,25, para vigas biapoiadas, e 2,5, para vigas contínuas [5]; e nos Estados Unidos da América, todas as vigas com a razão menor que 4 são vigas-parede [6].

As vigas-parede são frequentemente utilizadas em diversas aplicações estruturais, tais como paredes diafragmas, pontes, reservatórios de água, peças pré-moldadas de concreto protendido, fundações, silos, abrigos subterrâneos, estruturas offshore e edifícios altos [7].

O reforço de elementos estruturais, como vigas e colunas, tornou-se vital hoje em dia,

devido a muitas circunstâncias inevitáveis, como revisão das condições de carregamento, mudança nas condições de ocupação e deterioração da estrutura existente devido a efeitos ambientais [8]. Compósitos de fibras de alta resistência não metálicas, feitos de carbono, vidro ou aramida, encapsuladas em uma matriz polimérica sob a forma de arames, barras, fios, telas armadas em cruz, lâminas, tiras ou mantas têm mostrado grandes potenciais como reforço de concreto, particularmente onde a durabilidade é a preocupação principal [9].

Estes materiais são conhecidos como compósitos de resina e fibras (FRP), podendo ser de três tipos: compósitos de resina e fibras de carbono (CFRP), compósitos de resina e fibras de vidro (GFRP) e compósitos de resina e fibras de aramida (AFRP). Dentro os três tipos de FRP, o CFRP é o que apresenta melhor eficácia no aumento da capacidade resistente ao cisalhamento de vigas [10].

O CFRP propicia adicional reforço na flexão ou cisalhamento e a confiabilidade na aplicação deste material depende de como ele está ligado ao concreto e quanto de tensão o concreto pode transferir para o compósito de CFRP [11]. O CFRP oferece ao engenheiro projetista excelentes propriedades que não estão disponíveis em materiais tradicionais. Este material de reforço é resistente à corrosão e possui maior resistência e rigidez em relação ao aço [12].

As aplicações de FRP em engenharia civil podem ser divididas em três grandes grupos: novas construções, reparação e reconstrução [13]. O primeiro reparo no local por FRP colado externamente foi em 1991 e, em 1995, após o terremoto HyogokenNanbu, no Japão, houve um aumento repentino na utilização de materiais FRP [13]. Desde então, o reforço de estruturas com o

uso de materiais FRP colados externamente tem sido estudado em todo o mundo.

2. Estudos de reforço em vigas-parede

Em 2002, o reforço de oito vigas-parede subdimensionadas ao cisalhamento, com tiras de CFRP em diferentes inclinações (0° , 45° e 90°) foram estudadas com dois tipos de carregamento (uma carga central ou duas nos terços do vão), conforme esquematizado na **figura 1** [14]. Concluiu-se que o reforço a 45° foi o mais eficaz e que a 0° , o menos de acordo com o apresentado na **tabela 1**.

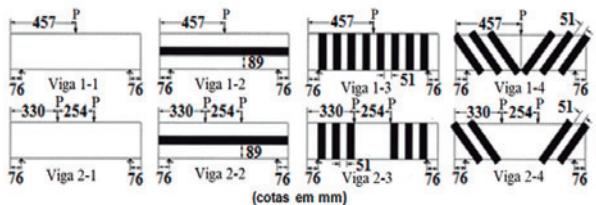


Fig. 1 – Representação das vigas-parede estudadas em 2002. Fonte: [14].

Em 2005, avaliou-se o comportamento do reforço de seis vigas-parede idênticas com o uso de mantas (2 camadas), tiras (50 mm x 1,2 mm) e telas armadas em cruz (áreas da seção transversal de $6,6\text{ mm}^2$, $17,5\text{ mm}^2$ e $39,2\text{ mm}^2$) de CFRP em diferentes configurações, como mostra a **figura 2** [9]. Concluiu-se que o uso de sistemas de FRP aumentou a capacidade resistente da viga estudada, como ilustra a **tabela 2**, e que o reforço com telas armadas em cruz coladas com orientação normal demonstrou ser o sistema mais eficaz com relação à quantidade de material utilizado.

Em 2014, apresentou-se um modelo analítico para vigas-parede de concreto armado reforçadas

ao cisalhamento externamente com tiras coladas de CFRP [15]. Um programa baseado no método dos elementos finitos foi utilizado para modelar as vigas-parede estudadas em 2002. Concluiu-se que a modelagem apresentou uma excelente correlação com os resultados experimentais.

Ainda em 2014, estudou-se o comportamento de doze vigas-parede de mesmas características (**figura 3**) submetidas a duas cargas simétricas em seis diferentes posições [16]. Metade das vigas foi reforçada com folhas de CFRP colada. Concluiu-se que, para a mesma configuração de carregamento, o reforço conduziu à diminuição da deformação transversal da viga-parede e ao aumento da sua capacidade resistente (**figura 4**).

Neste mesmo estudo, constatou-se ainda que, quanto maior a relação entre o vão de cisalhamento e a altura útil da viga (a/d), maior sua deformação transversal e menor sua carga de ruptura. A partir deste estudo, foi estabelecida uma relação empírica para ajustar o valor da deformação principal de tração em vigas-parede reforçadas com CFRP, calculado por meio do método das bielas e tirantes [17].

Em 2015, estudou-se o comportamento de vigas-parede de concreto auto-adensável reforçadas ao cisalhamento com folhas de CFRP [12]. Para isso, foram feitas onze vigas-parede com concreto auto-adensável, projetadas para romper ao cisalhamento, sendo dez delas reforçadas com folhas de CFRP coladas em diferentes configurações, cujos resultados são apresentados na **tabela 3**.

Tab. 1 – Resultados obtidos no estudo de 2002.

Viga	Carga última (kN)	Flecha para carga última (mm)
1-1	94,3	2,308
1-2	97,9	2,074
1-3	168,1	3,292

Viga	Carga última (kN)	Flecha para carga última (mm)
1-4	193,5	5,224
2-1	144,6	1,943
2-2	204,2	2,745
2-3	207,7	3,104
2-4	240,7	2,784

Fonte: [14].

Tab. 2 – Resultados encontrados no estudo de 2005.

Viga	Carga de primeira fissura		Carga última (kN)
	Flexão (kN)	Diagonal (kN)	
B0	130	280	1.040
B1	160	-	1.402
B2	150	370	1.400
B3	195	490	1.410
B4	130	350	1.292
B5	190	330	1.482

Fonte: [9].

Neste estudo, concluiu-se que, em todas as vigas-parede reforçadas, o reforço com o uso de CFRP levou ao aumento, em média, de 33% em relação à carga máxima da viga sem reforço e à redução da flecha da viga para o mesmo nível de carga. O uso de folhas de CFRP inclinadas de 45° em relação ao eixo longitudinal da viga propiciou um maior acréscimo na carga última e um maior decréscimo na flecha e na abertura das fissuras em relação ao uso de folhas verticais (90°).

3. Proposta de dimensionamento do reforço ao cisalhamento

Baseando-se nos resultados experimentais dos artigos apresentados, este trabalho propõe a metodologia descrita a seguir para o dimensionamento do reforço ao cisalhamento de vigas-parede sem abertura com a utilização de compósitos de resina e fibras colados externamente.

3.1 Condições gerais

Esta metodologia considera que, no vão de cisalhamento (a), definido pela região compreendida entre o apoio e o ponto de aplicação do carregamento, a força cortante (V) pode ser decomposta em uma força de compressão, que corresponde ao caminho da carga até o apoio, e uma força de tração, ortogonal a essa força de compressão, conforme apresentado na **figura 5**.

O ângulo de inclinação das diagonais de compressão em relação ao eixo longitudinal da viga (θ) é definido pelo arco cuja tangente é a razão entre a sua altura útil (d) e o seu vão de cisalhamento (a), não sendo menor que 30° . A determinação da força cortante resistente de cálculo relativa à ruína das diagonais comprimidas de concreto (V_{Rd2}) é feita segundo o Modelo de Cálculo II (v. **equações 1 e 2**), onde α é a inclinação da armadura transversal em relação ao eixo longitudinal [2].

$$V_{Rd2} = 0,6\alpha_{v2}f_{cd}b_wz\sin^2\theta(\cot\alpha + \cot\theta) \quad (1)$$

$$\alpha_{v2} = (1 - f_{ck}/250), \text{ com } f_{ck} \text{ em MPa} \quad (2)$$

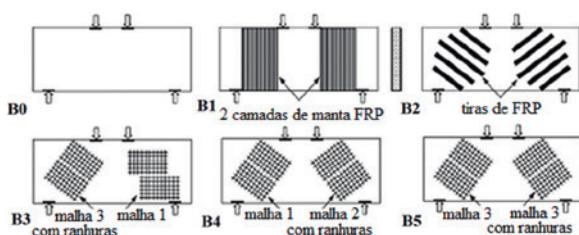


Fig. 2 – Vigas-parede estudadas em 2005. Fonte: [9].

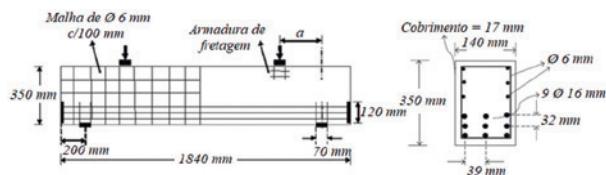


Fig. 3 – Vigas estudadas em 2014. Fonte: [16].

O braço de alavanca z deve ser calculado conforme as **equações 3, 4 e 5** [18]. No caso de vigas-parede com $L/h > 2$, estas são calculadas como se vigas esbeltas fossem [2] (v. **equações 5 a 10**), onde M_{sd} é o momento fletor solicitante de cálculo, α_c é o parâmetro de redução da resistência do concreto na compressão (v. **equações 7 e 8**) e λ é a relação entre a profundidade y do diagrama retangular de compressão equivalente e a profundidade efetiva x da linha neutra na ruptura (v. **equações 9 e 10**).

$$z = 0,6L, \text{ para } L/h \leq 1 \quad (3)$$

$$z = 0,15h[3 + (L/h)], \text{ para } 1 < L/h \leq 2 \quad (4)$$

$$z = d - \lambda x/2, \text{ para } L/h > 2 \quad (5)$$

$$x = \frac{\alpha_c f_{cd} \lambda d \pm \sqrt{(\alpha_c f_{cd} \lambda d)^2 - 2\alpha_c f_{cd} \lambda^2 M_{sd}}}{\alpha_c f_{cd} \lambda^2} \quad (6)$$

$$\alpha_c = 0,85, \text{ para } f_{ck} \leq 50 \text{ MPa} \quad (7)$$

$$\alpha_c = 0,85[1,0 - (f_{ck} - 50)/200], \text{ para } f_{ck} > 50 \text{ MPa} \quad (8)$$

$$\lambda = 0,80, \text{ para } f_{ck} \leq 50 \text{ MPa} \quad (9)$$

$$\lambda = 0,80 - (f_{ck} - 50)/400, \text{ para } f_{ck} > 50 \text{ MPa} \quad (10)$$

A força cortante resistente de cálculo relativa à ruína por tração diagonal (V_{Rd3}) é calculada pela **equação 11**, onde V_c é a parcela de força cortante resistida por mecanismos complementares ao modelo em treliça, V_{sw} é a parcela de força cortante resistida pela armadura de cisalhamento na direção ortogonal à diagonal comprimida e V_{fw} é a parcela

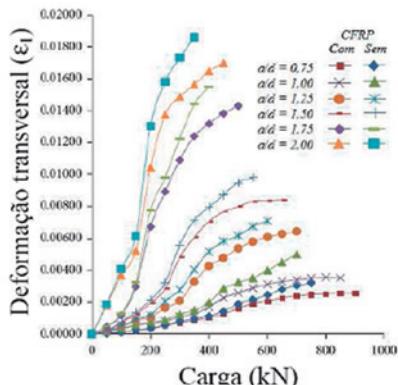


Fig. 4 – Resultados obtidos no estudo de 2014. Fonte:[16].

Tab. 3 – Resultados encontrados no estudo de 2015.

Viga	Tipo de reforço	Carga de fissuração (kN)	Carga última (kN)
SCCUS2-1	Sem reforço (viga de referência)	175	400
SCCUS2-2	Folhas a 90° em 2 lados, com altura de 300 mm, espaçadas a 50 mm	190	490
SCCUS2-3	Folhas a 45° em 2 lados, com altura de 300 mm, espaçadas a 50 mm	211	496
SCCUS2-4	Folhas em 3 lados (em forma de U), com altura de 300 mm, espaçadas a 50 mm	215	500
SCCUS2-5	Envolvimento total com folhas espaçadas a 50 mm	225	520
SCCUS2-6	Folhas a 90° em 2 lados, com altura de 300 mm, espaçadas a 20 mm	200	504
SCCUS2-7	Folhas a 45° em 2 lados, com altura de 300 mm, espaçadas a 20 mm	217	507
SCCUS2-8	Folhas em 3 lados (em forma de U), com altura de 300 mm, espaçadas a 20 mm	227	512
SCCUS2-9	Envolvimento total com folhas espaçadas a 20 mm	241	532
SC-CUS2-10	Folhas a 90° em 2 lados, com largura de 80 mm e altura de 300 mm, espaçadas a 50 mm	221	509
SC-CUS2-11	Folhas verticais e horizontais em 2 lados, com altura de 300 mm, espaçadas a 50 mm	200	498

Fonte: [12].

de força cortante resistida pelo reforço ao cisalhamento na direção ortogonal à diagonal comprimida.

$$V_{Rd3} = V_c + (V'_{sw} + V'_{fw}) / \cos \theta \quad (11)$$

A parcela V_c é obtida interpolando-se linearmente quando $V_{c0} < V_{sd} < V_{Rd2}$ e adotando-se $V_c = V_{c0}$, quando $V_{sd} \leq V_{c0}$ e $V_c = 0$, quando $V_{sd} = V_{Rd2}$, onde V_{c0} é calculado através da **equação 12** [2]. O valor de f_{ctd} , em MPa, pode ser determinado pelas **equações 13 a 15**, onde $f_{ct,sp}$ é a resistência à tração por compressão diametral do concreto e $f_{ct,f}$ é a resistência à tração na flexão do concreto.

$$V_{c0} = 0,6 f_{ctd} b_w d \quad (12)$$

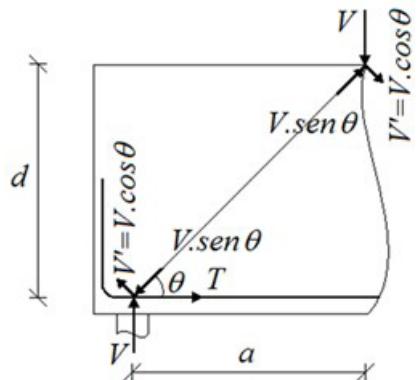


Fig. 5 – Representação esquemática da decomposição da carga em vigas-parede.

$$f_{ctd} = 0,9 f_{ct,sp} / \gamma_c \quad (13)$$

$$f_{ctd} = 0,7 f_{ct,f} / \gamma_c \quad (14)$$

$$f_{ctd} = 0,21 f_{ck}^{2/3} / \gamma_c \quad (15)$$

Nesta metodologia, a parcela V'_{sw} considera a resistência fornecida pela armadura transversal e pela armadura de pele, que é frequentemente utilizada em vigas-parede. Para o caso mais comum de armadura de cisalhamento, composta por estribos verticais e horizontais, o valor de V'_{sw} pode ser calculado por meio da **equação 16**, onde s é o espaçamento entre os estribos verticais, $A_{sw,v}$ é

a área da seção transversal do estribo vertical, $A_{sw,h}$ é a área da seção transversal do estribo horizontal, n_h é o número de estribos horizontais que compõe a armadura de pele (desconsiderando a armadura longitudinal e a armadura de montagem).

$$V'_{sw} = \frac{d \cdot \cot \theta}{s} A_{sw,v} f_{ywd} \cos \theta + n_h A_{sw,h} f_{ywd} \sin \theta \quad (16)$$

A contribuição da armadura de pele deve ser desconsiderada caso $\tan \theta > d/a$. A figura 6 apresenta esquematicamente as componentes da parcela V'_{sw} . Para o cálculo da parcela V'_{fw} , deve-se considerar o tipo do reforço utilizado para se definir sua deformação e seu módulo de elasticidade, bem como sua espessura e configuração, contendo o número de camadas, a inclinação, a largura e o espaçamento de cada camada de reforço. As equações propostas para o cálculo desta parcela são apresentadas nos itens 3.2.2 a 3.2.5.

3.2 Condições específicas

Esta seção descreve os procedimentos específicos que devem ser adotados, tendo em vista a singularidade dos principais casos de reforço ao cisalhamento em vigas-parede.

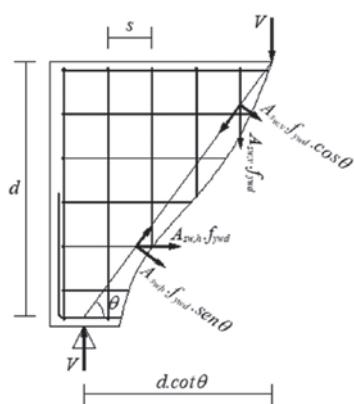


Fig. 6 – Representação esquemática das componentes da parcela V'_{sw} .

3.2.1 Vigas com $L > 2h$

Tendo em vista a divergência na definição de vigas-parede por algumas normas internacionais, existem vigas que não são consideradas vigas-parede segundo a norma brasileira ($L/h > 2$), mas são consideradas por outras. Desta forma, para vigas com altura inferior à metade do seu vão teórico, deve-se realizar os seguintes procedimentos:

- Quando a razão a/d for menor que a tangente de 30° , o valor do ângulo θ será igual a 30° e a contribuição da armadura de pele em V'_{sw} deverá ser desconsiderada.
- Quando uma carga concentrada for aplicada a uma distância $a \leq 2d$ do eixo teórico do apoio em vigas com apoio direto, a verificação da resistência relativa à ruína por tração diagonal (V_{Rd3}) deve ser feita considerando a força cortante de cálculo reduzida pelo fator $a/(2d)$, denominada nesta metodologia como V_{red} , conforme prescreve a alínea b) do item 17.4.1.2.1 da norma brasileira [2].
- A redução supracitada não modifica a força cortante de cálculo (V_{sd}) para a verificação da resistência relativa à ruína por compressão diagonal (V_{Rd2}) e para o cálculo da parcela V_c .

3.2.2 Reforço unidirecional com FRP

A parcela V'_{fw} do reforço em uma única direção com FRP pode ser calculada por meio da **equação 17**, onde α_f é o ângulo que as fibras do reforço fazem com o eixo longitudinal da viga, ϵ_{fud} é a deformação última de cálculo do compósito (definida no item 3.2.5), E_f é o módulo de elasticidade do reforço, t_f e b_f são a espessura total e a largura de cada tira do reforço, e n_f é a quantidade de tiras que atravessa

a diagonal comprimida, expressa pela **equação 18**, onde s_f é a distância entre o centroide de duas tiras consecutivas. A **figura 7** apresenta esquematicamente esta contribuição.

$$V'_{fw} = 2 \cdot n_f b_f t_f E_f \varepsilon_{fud} \operatorname{sen}(\alpha_f + \theta) \quad (17)$$

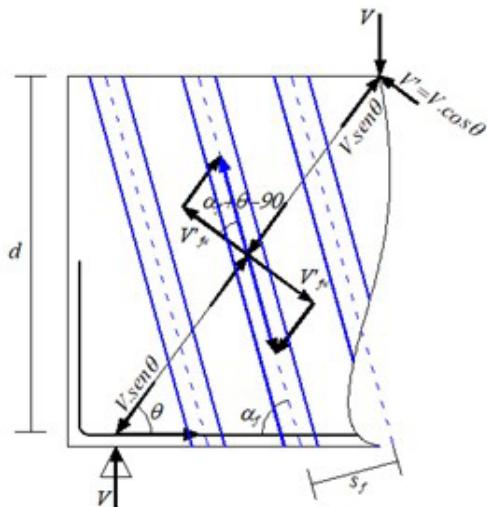


Fig. 7 – Representação esquemática da parcela V'_{fw} do reforço unidirecional.

$$n_f = \frac{d}{s_f} (\operatorname{cot} \alpha_f + \operatorname{cot} \theta) \operatorname{sen} \alpha_f \quad (18)$$

3.2.3 Reforço multidirecional com FRP

Para o cálculo da resistência ao cisalhamento do reforço com FRP em várias direções, soma-se a contribuição de cada uma das direções, obtidas de forma similar ao reforço unidirecional, conforme apresentado na **equação 19** e **figura 8**. Entretanto, apenas em uma direção, a deformação do reforço será igual à deformação última de cálculo. Na determinação da deformação do reforço em cada direção, considera-se que o deslocamento relativo à tração da diagonal comprimida seja perpendicular a mesma, sendo obtida por meio da **equação 20**.

$$V'_{fw} = \sum_i 2 \cdot n_{fi} \cdot b_{fi} \cdot t_{fi} \cdot E_{fi} \cdot \varepsilon_{fi} \cdot \operatorname{sen}(\alpha_{fi} + \theta) \quad (19)$$

$$\varepsilon_{fi} = \frac{\Delta L}{L_{fi}} \cdot \cos(\alpha_{fi} - \theta) \leq \varepsilon_{fud} \quad (20)$$

3.2.4 Reforço contínuo com FRP

Um caso particular de reforço é o com FRP colado externamente de forma contínua e abrangendo a totalidade do vão de cisalhamento. Para este tipo de reforço, a parcela n_f da **equação 17** (ou n_{fi} da **equação 19**, conforme o caso) é igual a 1. Quando este reforço contemplar a altura total da viga em suas faces laterais, pode-se considerar que toda a tensão de tração transversal que atua no concreto naquela região seja resistida pela armadura transversal e pelo reforço, caso a resistência relativa à ruína das diagonais comprimidas de concreto (V_{Rd2}) seja insuficiente.

Dessa forma, a força cortante resistente de cálculo relativa à ruína das diagonais comprimidas de concreto pode ser expressa pela **equação 21**. Embora a mudança de seu estado de tensões altere a tensão resistente máxima do concreto das diagonais comprimidas em regiões com

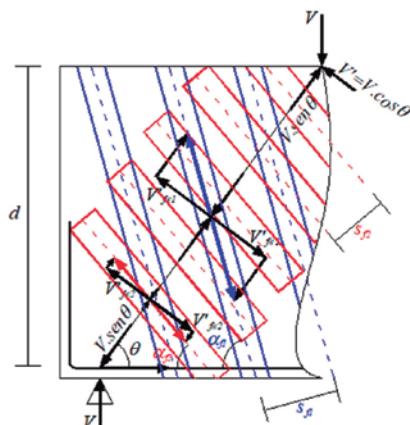


Fig. 8 – Representação esquemática da parcela V'_{fw} do reforço em duas direções.

tensões de tração transversal ($f_{cd2} = 0,60 \alpha_v f_{cd}$) pararegiões sem tensões de tração transversal ($f_{cd1} = 0,85 \alpha_v f_{cd}$), a tensão resistente máxima no concreto é limitada ao valor de sua região nodal ($f_{cd3} = 0,72 \alpha_v f_{cd}$), uma vez que esta biela possui nós onde conflui um tirante tracionado (nós CCT). Caso seja adotada esta mudança, a parcela V_c deverá ser desconsiderada.

$$V_{Rdz'} = 0,72 \alpha_v f_{cd} b_w z \operatorname{sen}^2 \theta (\cot \alpha + \cot \theta) \quad (21)$$

3.2.5 Deformação última de cálculo do compósito

Embora todos os compósitos de resina e fibras utilizados como reforço possuam deformação última informada pelo respectivo fabricante, este valor não pode ser utilizado diretamente para o cálculo da resistência do reforço, uma vez que existem outros mecanismos de falha do reforço, além da ruptura da fibra. Estes mecanismos são o descolamento e o destacamento.

Descolamento como o desligamento dos materiais concreto e de reforço, ocorrido na interface material de reforço – cola, sem haver camada de concreto aderida ao material de reforço [19]. Já destacamento é entendido como sendo a separação entre o reforço e o elemento reforçado de maneira que haja concreto aderido ao material de reforço [19].

Tendo em vista que, quando ocorrer qualquer mecanismo de falha do reforço, o colapso da viga será iminente, a deformação última de cálculo do reforço (ε_{fud}) deve ser limitado. Então, para esta metodologia, foram utilizadas as equações prescritas na norma americana[20] para o cálculo da deformação efetiva do reforço ao cisalhamento, quando não há sistema de ancoragem deste reforço, e as equações propostas

em[21], caso contrário.

O projetista deverá considerar todos os modos de falha possíveis na estrutura de concreto armado reforçado com compósitos de resina e fibras e usar uma deformação efetiva (ou deformação última de cálculo) representativa do modo da falha crítica [20]. Para vigas de concreto armado reforçadas externamente por compósito de resina e fibras com envelopamento total, a perda do engrenamento dos agregados do concreto foi observada ocorrendo com deformações menores que a deformação última da fibra [20]. Para evitar este tipo de falha, a deformação máxima de projeto deverá ser limitada pela **equação 22**, onde ε_{fu} é a deformação última do reforço.

$$\varepsilon_{fud} = 0,75 \varepsilon_{fu} \leq 0,004 \quad (22)$$

Para sistemas de reforço com compósitos de resina e fibras que não envelopam toda a seção (colados em duas ou três faces), foi observado que o destacamento do concreto ocorre antes da perda do engrenamento dos agregados da seção. Por esta razão, tensões de aderência foram analisadas para determinar a utilidade destes sistemas e o nível de deformação efetiva que pode ser alcançado. Dessa forma, a referida norma propõe que a deformação efetiva seja calculada usando um coeficiente de redução do efeito de ligação K_v aplicável ao cisalhamento, conforme **equação 23**.

$$\varepsilon_{fud} = K_v \varepsilon_{fu} \leq 0,004 \quad (23)$$

O coeficiente de redução do efeito de ligação é uma função da resistência do concreto, da configuração do reforço utilizado e a rigidez do compósito. Este coeficiente pode ser calculado pelas **equações 24 a 28**. O comprimento efetivo da

ligação, L_v , é o comprimento em que a maior parte da tensão de aderência é mantida e é calculado pela **equação 25**, onde n_f é o número de camadas do reforço, t_f é a sua espessura de cada camada em mm e E_f é o seu módulo de elasticidade em MPa.

$$K_v = \frac{k_1 k_2 L_e}{11900 \varepsilon_{fu}} \leq 0,75 \quad (24)$$

$$L_e = \frac{23300}{(n_f t_f E_f)^{0,58}} \quad (25)$$

$$k_1 = \left(\frac{f_{cm}}{27} \right)^{2/3} \quad (26)$$

$$k_2 = \frac{d_{fv} - L_e}{d_{fv}}, \text{ para reforço colado em "U"} \quad (27)$$

$$k_2 = \frac{d_{fv} - 2L_e}{d_{fv}}, \text{ para reforço colado em duas faces} \quad (28)$$

O coeficiente de redução do efeito de ligação também depende de dois fatores de modificação, k_1 e k_2 , que consideram a resistência do concreto e a configuração do reforço utilizado, respectivamente, conforme as **equações 26, 27 e 28**, onde f_{cm} é a resistência à compressão média do concreto em MPa e d_{fv} é a altura efetiva do reforço ao cisalhamento em mm (**figura 9**).

Esta norma ressalta que as equações para a determinação do K_v foram validadas para elementos em regiões com alto cortante e baixo momento fletor, entretanto ainda não foram confirmadas para reforço ao cisalhamento em áreas sujeitas à combinação de altas tensões de flexão e cisalhamento ou em regiões com momento negativo. Contudo, este fator de redução é suficientemente conservativo para estes casos.

Ancoragens mecânicas usadas nas extremidades do reforço propiciam ao mesmo

maior eficácia, possibilitando o aumento de sua deformação efetiva, calculada pela **equação 23** [20]. Dessa forma, sugere-se a utilização das **equações 29, 30 e 31** para determinar a deformação última de cálculo de reforço com fibras de carbono, quando o mesmo for ancorado na extremidade por uma tira horizontal de mesma largura para se evitar o destacamento do reforço (**equação 29**)

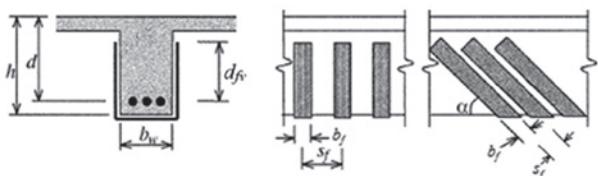


Fig. 9 – Representação esquemática das principais variáveis dimensionais do reforço ao cisalhamento.

e para se evitar o descolamento do reforço (**equações 30 e 31**) [21].

$$\varepsilon_{fud} = \frac{\tau_{lim}}{E_f \left(\frac{h_f + 4b_{ext}}{2n_f t_f} \right)} \quad (29)$$

$$\varepsilon_{fud} = 5\%, \text{ para reforço pré-fabricado} \quad (30)$$

$$\varepsilon_{fud} = 6\%, \text{ para reforço } in situ \quad (31)$$

Na **equação 29**, b_{ext} é o comprimento mostrado na **figura 10**, considerado igual ao menor valor entre a largura da tira de ancoragem e a metade da distância horizontal sem reforço entre duas tiras sucessivas, h_f é a altura do reforço, n_f é o número de camadas do reforço, t_f é a espessura de uma camada do reforço, E_f é o módulo de elasticidade do reforço e τ_{lim} é a tensão cisalhante limite a ser resistida pelo concreto, expressa pela **equação 32**, onde f_{ctd} é a resistência de cálculo à tração do concreto (**equações 13 a 15**).

$$\tau_{\lim} = 0,50 f_{ctd} \quad (32)$$

3.3 Validação da metodologia

Visando a validação da metodologia proposta de dimensionamento do reforço ao cisalhamento com FRP colados externamente em vigas-parede, aplicaram-se as equações deste trabalho, desconsiderando todos os fatores de segurança, nas vigas B0, B1 e B2 de 2005, [9], nas 12 vigas-parede estudadas em 2014, [16], e nas 336 vigas-parede da base de dados utilizadas em um estudo de vigas-parede sem reforço que possuíam armadura de cisalhamento [22].

As **figuras 11 e 12** apresentam a comparação dos resultados experimentais com os valores calculados de acordo com esta metodologia para as 15 vigas provenientes dos estudos com reforço e para as 336 vigas dos estudos sem reforço, onde percebe-se uma boa aproximação entre a resistência real e a resistência esperada. Os valores da média e desvio padrão da razão entre a resistência real ($V_{u \text{ exp}}$) e a resistência esperada ($V_{u \text{ calc}}$) para as vigas dos estudos com e sem reforço foram iguais a 1,05 e 0,07 e 1,48 e 0,86.

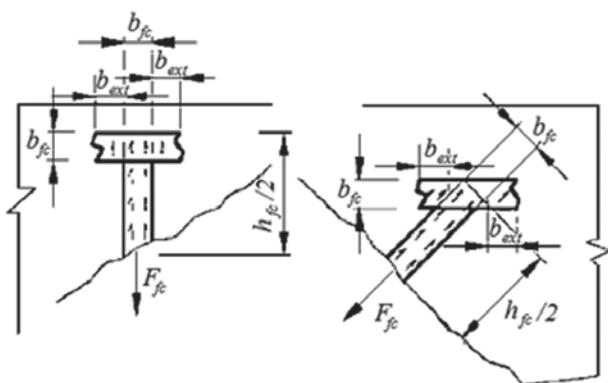


Fig. 10 – Representação esquemática da Área de Contato do Compósito com o Concreto.

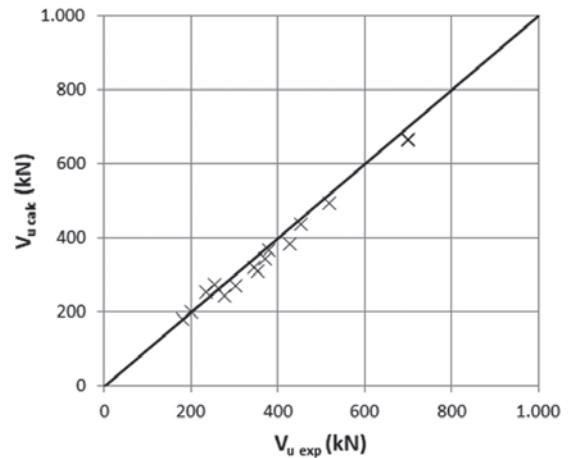


Fig. 11 – Comparação dos resultados experimentais com os valores calculados para as vigas dos estudos com reforço.

4. Conclusões

Constatou-se que o tipo, a configuração e taxa de armadura de reforço, a resistência do concreto à compressão, a relação entre o vão de cisalhamento e a altura útil e a taxa de armadura interna das vigas são os parâmetros mais influentes no comportamento das vigas-parede de concreto armado reforçadas ao cisalhamento com FRP colados.

Tendo em vista que não há um método normatizado ou proposto em outros estudos de dimensionamento de reforço ao cisalhamento de vigas-parede com materiais compósitos de resina e fibras colados externamente, este trabalho propôs uma metodologia preliminar para este dimensionamento.

Ao se comparar as cargas de ruptura determinadas com as obtidas, usando-se as equações propostas neste artigo sem os fatores de segurança previstos, encontrou-se uma boa aproximação entre elas, com a média e o desvio padrão da razão entre elas para as vigas dos

estudos com e sem reforço foram iguais a 1,05 e 0,07 e 1,48 e 0,86, respectivamente.

Entretanto, esta metodologia é válida apenas para vigas-parede sem aberturas biapoiadas sujeitas a cargas concentradas e reforçadas por folhas ou lâminas de

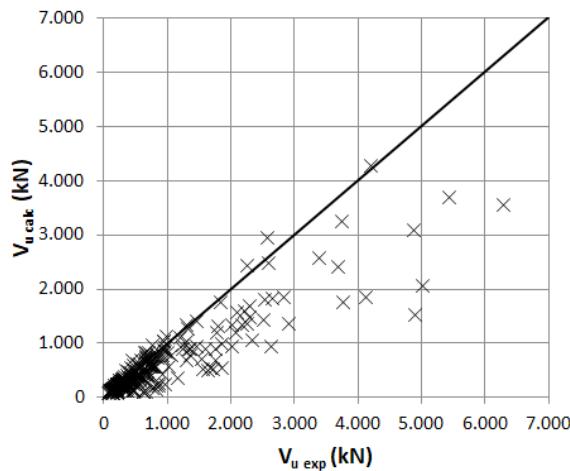


Fig. 12 – Comparação dos resultados experimentais com os valores calculados para as vigas dos estudos sem reforço.

materiais compósitos de resina e fibras, sem

ancoragem [20] ou ancoradas por tiras horizontais em suas extremidades [21].

Embora esta proposição tenha se baseado em resultados experimentais com compósitos de resina e fibras de carbono, ela pode ser usada para o dimensionamento de reforço com outros tipos de compósitos de resina e fibras, desde que colados externamente, uma vez que as equações para a determinação da deformação última de cálculo não se limita apenas ao reforço com compósitos de resina e fibras de carbono [20]. Contudo, necessita-se de estudos complementares para a determinação da deformação última de cálculo de reforço com outro tipo de ancoragem que não a apresentada neste trabalho [21].

O modelo proposto não dimensiona o reforço de vigas-parede sem armadura transversal, uma vez que seu mecanismo de ruptura sem reforço é definido pela resistência do arco ou do engaste dos dentes definidos entre fissuras de flexão consecutivas e, ao se reforçá-las, o compósito passa a atuar como estribo, modificando este mecanismo de ruptura.

Referências bibliográficas

- [1] F. Leonhardt, E. Mönnig, Construções de concreto: Princípios básicos do dimensionamento de estruturas de concreto armado. Vol.1. Interciência, Rio de Janeiro, 1977.
- [2] Associação Brasileira de Normas Técnicas, ABNT. ABNT NBR 6118 – Projeto de estruturas de concreto - Procedimento. Rio de Janeiro, 2014.
- [3] S. D. Kore, S. S. Patil, Analysis and design of R. C. deep beams using code provisions of different countries and their comparison. International Journal of Engineering and Advanced Technology, Vol. 2, p. 166-169, 2013
- [4] Construction Industry Research and Information Association. CIRIA Guide 2 – The design of deep beams in reinforced concrete. London, 1984.
- [5] Canadian Standard Association. CAN-A23.3-M84 – Design of concrete structures for buildings. Ontario, 1984.
- [6] American Concrete Institute. ACI 318M – Building Code Requirements for Structural Concrete. Farmington Hills, 2011.
- [7] K. M. Heiza, N. N. Meleka, N. Y. Elwkad, Shear strengthening of self-compacting reinforced concrete deep beams with external bonded layers. Concrete Research Letters, Vol. 3, No.4, p. 507-527, 2012.

- [8] M. R. Patel, T. Tank, Finite element modeling of RC deep beams strengthened in shear with CFRP strips. *The Indian Concrete Journal*, Vol. 88, p. 69-76, 2014.
- [9] M. R. Islam, M. A. Mansur, M. Maalej, Shear strengthening of RC deep beams using externally bonded FRP systems. *Cement & Concrete Composites*, Vol. 27, p. 413-420, 2005.
- [10] M. R. Patel, R. K. Gajjar, Shear strengthening of different beams using FRP. *International Journal of Advanced Engineering Research and Studies*, Vol. 1, p. 290-294, 2012.
- [11] M. Ekenel, V. Stephen, J. J. Myers, R. Zoughi, Microwave NDE of RC beams strengthened with CFRP laminates containing surface defects and tested under cyclic loading. *Electrical and Computer Engineering*, University of Missouri-Rolla, Rolla, p. 1-8, 2004.
- [12] J. A. Khudair, R. S. Atea, Shear behavior of self compacting concrete deep beams strengthened with carbon fiber reinforced polymer sheets. *International Journal of Engineering Research & Technology*, Vol. 4, p. 187-191, 2015.
- [13] M. B. S. Alferjani, A. A. A. Samad, B. S. Elrawaff, N. Mohamad, M. Hilton, A. A. S. Saiah, Use of carbon fiber reinforced polymer laminate for strengthening reinforced concrete beams in shear: A review. *International Refereed Journal of Engineering and Science*, Vol. 2, pp. 45-53, 2013.
- [14] J. E. Moren, Shear behavior of reinforced concrete deep beams strengthened with CFRP laminates. *Dissertação de Mestrado*. New Jersey Institute of Technology, New Jersey, 2002.
- [15] M. R. Patel, T. Tank, Finite element modeling of RC deep beams strengthened in shear with CFRP strips. *The Indian Concrete Journal*, Vol. 88, p. 69-76, 2014.
- [16] M. Panjehpour, H. K. Chai, Y. L. Voo, Strut deformation in CFRP-strengthened reinforced concrete deep beams. *The Scientific World Journal*, Vol. 2014, p. 1-9, 2014.
- [17] M. Panjehpour, A. A. A. Ali, Y. L. Voo, F. N. Aznieta, Modification of strut effectiveness factor for reinforced concrete deep beams strengthened with CFRP laminates. *Materiales de Construcción*, Vol. 64, p. 1-8, 2014.
- [18] E. C. S. Thomaz, L. A. V. Carneiro, Considerações sobre o dimensionamento de vigas paredes de concreto com o uso do método das bielas e tirantes. *52º Congresso Brasileiro do Concreto*. Fortaleza, IBRACON, 2010.
- [19] L. A. V. Carneiro, Reforço de vigas e pilares de concreto com materiais compósitos de resina e fibras. *Tese de Doutorado*. Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2004.
- [20] American Concrete Institute. *ACI 440.2R – Guide for the design and construction of externally bonded FRP systems for strengthening concrete structures*. Farmington Hills, 2008.
- [21] E. C. Cerqueira, Reforço ao cisalhamento de vigas de concreto armado com fibras de carbono. *Dissertação de Mestrado*. Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2000.
- [22] J. Liu, B. I. Mihaylov, A comparative study of models for shear strength of reinforced concrete deep beams. *Engineering Structures*, Vol. 112, p. 81-89, 2016.

Avaliação da corrosão de ligas de magnésio por ensaios de imersão e evolução de hidrogênio

Ana Caroline Crema de Almeida Fontes, Carlos Nelson Elias

Instituto Militar de Engenharia (IME)

Praça General Tibúrcio, 80, 22290-270, Praia Vermelha, Rio de Janeiro, RJ, Brasil.

elias@ime.eb.br

RESUMO: O presente estudo teve como objetivo avaliar as propriedades de corrosão por imersão e a evolução de hidrogênio de duas ligas não comerciais de magnésio ($Mg-39Ca$ e $Mg-3Gd-5Ca$) desenvolvidas para utilização como implantes bioabsorvíveis. Foram realizados ensaios de corrosão por imersão e captação de hidrogênio. As análises foram realizadas nas amostras antes e após o tratamento térmico de homogeneização a 400 °C durante 10 horas. Os resultados mostraram que o tratamento térmico de homogeneização aumenta a resistência à corrosão das ligas. A liga $Mg-39Ca$ submetida ao tratamento térmico apresentou as menores taxas de corrosão por imersão e a menor taxa de evolução de hidrogênio.

PALAVRAS-CHAVE: Liga Mg-Ca. Liga Mg-Ca-Gd. Corrosão por imersão. Evolução de hidrogênio.

1. Introdução

O Magnésio (Mg), por apresentar baixa resistência à corrosão, normalmente é usado na forma de liga para melhorar esta propriedade [1]. Entre todos os metais aplicados na engenharia, o Mg tem potencial de corrosão (E_{corr}) de aproximadamente -1,7 V em soluções contendo íons cloreto e é o metal que possui o menor potencial padrão ($E_0 = 2,37$ V a 25 °C) [2].

Apesar da baixa resistência à corrosão para aplicação na engenharia, na área dos biomateriais a elevada taxa de degradação característica do Mg e suas ligas possuem propriedades adequadas para emprego como implantes bioabsorvíveis.

ABSTRACT: The present study aimed to evaluate the immersion corrosion properties and hydrogen evolution of two non-commercial magnesium alloys ($Mg-39Ca$ and $Mg-3Gd-5Ca$) developed for use with bioabsorbable implants. Corrosion tests by immersion and hydrogen uptake were performed. Analyzers were performed on samples before and after the homogenization heat treatment at 400 °C for 10 hours. The results showed that the homogenization heat treatment increases the corrosion resistance of the alloys. The heat-treated $Mg-39Ca$ alloy presented the lowest rates of immersion corrosion and the lowest evolution rate of hydrogen.

KEYWORDS: Mg-Ca alloy. Mg-Ca-Gd alloy. Corrosion by immersion. Hydrogen evolution.

A principal vantagem destes implantes é o fato de ser desnecessária uma segunda cirurgia para sua remoção, reduzindo ou excluindo os inconvenientes a longo prazo apresentados por implantes permanentes [2,3]. No entanto, as aplicações das ligas à base de Mg ainda são modestas devido às altas taxas de degradação e consequente perda de propriedades mecânicas em ambientes com níveis de pH entre 7,4 e 7,6, que condiz com o pH de sistemas fisiológicos [4]. Outra característica das ligas de Mg é que, ao reagir com soluções aquosas, é produzido gás hidrogênio podendo resultar na formação de bolsas de gás na área implantada [5].

Portanto, apesar da corrosão do Mg ser

interessante do ponto de vista de fabricação de implantes absorvíveis, a corrosão pode ser muito rápida e não fornecer tempo hábil para o organismo consolidar a fratura e ocorrer a regeneração óssea. Esta a razão da importância dos elementos de liga que serão adicionados ao Mg para modificar sua resistência à corrosão e controlar as reações envolvidas no processo durante a degradação.

2. Metodologia

No presente trabalho foram estudadas duas novas composições de ligas à base de Mg: magnésio-cálcio (Mg-Ca) e magnésio-cálcio-gadolínio (Mg-Ca-Gd). As ligas foram fabricadas por fundição pelo grupo do Professor Dexue Liu (Lanzhou University of Technology, Lanzhou - China) e fornecidas na forma de barras. As barras das ligas no estado como fundidas foram cortadas nas dimensões de 1,0 x 1,0 x 0,1 cm. Para os cortes foi utilizada a cortadeira modelo Brilliant 210 e o disco de Coríndon com 0,6 mm de espessura (ATM - GmbH).

Antes dos ensaios de corrosão, as amostras foram submetidas a polimento da superfície semelhante à preparação metalográfica. Inicialmente, foram usadas lixas de carboneto de silício (SiC) de granulometria 320 (Arotec), seguido pelo polimento com pasta de diamante (Struers) de 9 µm, 6 µm e 3µm. O polimento foi finalizado com sílica coloidal (Struers).

Para o ensaio de imersão, foram utilizadas 3 amostras de cada liga. As amostras tiveram suas dimensões medidas com um paquímetro (para cálculo da área de superfície) e suas massas quantificadas em uma balança analítica do modelo Ohaus Adventurer da Toledo, com precisão de 3 casas decimais. Em seguida,

cada amostra foi inserida em um tubo que foi preenchido com 50 mL de soro fisiológico e os tubos foram armazenados em estufa a 36 °C.

Após 24 horas de imersão, as amostras foram retiradas dos tubos, lavadas, fotografadas, limpas em solução ácida à base de H_2SO_4 e pesadas novamente. Os tubos tiveram suas soluções trocadas por novas e as amostras foram imersas mais uma vez. Esse procedimento foi realizado 3 vezes (24, 48 e 72 horas). Este ensaio foi uma adaptação da Norma Técnica NACE/ASTM TM0169 G0031 12A Standard Guide for Laboratory Immersion Corrosion Testing of Metals uma vez que ainda não existem Normas específicas para as aplicações em questão.

Para a captação de hidrogênio (H_2) a amostra foi imersa em NaCl (9 mg/mL) e coberta com um funil cuja saída dava para o interior de um tubo graduado. Esse tubo também estava cheio de soro fisiológico e conforme o hidrogênio ia sendo liberado ele ocupava o espaço do soro no interior do tubo fazendo o nível de solução baixar. Após 17 horas de imersão o ensaio foi finalizado. O volume de líquido deslocado foi considerado como sendo o volume de hidrogênio produzido. Este ensaio foi baseado na metodologia proposta por SONG [6].

3. Resultados e discussão

Nas **figuras 1-A, B e C** são mostradas as amostras de Mg-39Ca não homogeneizada após o ensaio de degradação durante 24, 48 e 72 horas de imersão em solução salina. Pode-se observar que a corrosão ocorreu em toda a superfície das amostras, porém com profundidades de penetração heterogênea. Após as 72 horas de ensaio foi difícil de remover a amostra da imersão sem ocorrer a fratura. Como pode-se

observar na **figura 1-C**, a liga ficou reduzida a uma pequena porção de material sem qualquer semelhança com a forma cilíndrica inicial. Nas **figuras 1-D, E e F** são mostradas as imagens de uma amostra da mesma liga após o tratamento de homogeneização com imersão durante 24, 48 e 72 horas, respectivamente. A corrosão também ocorreu por toda a superfície. Em uma análise visual verificou-se que a degradação após o tratamento térmico de homogeneização foi mais homogênea e com menor perda de massa em relação a antes da homogeneização (**figuras 1-A, B e C**). A corrosão foi homogênea em toda a superfície da amostra sem torná-la frágil permitindo o manuseio após o ensaio.

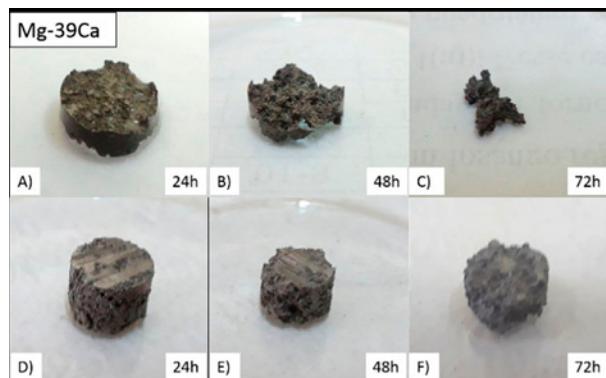


Fig. 1 - Amostras da liga Mg-39Ca após 24, 48 e 72 horas de imersão em soro fisiológico a 36 °C. Imagens A, B e C antes do tratamento térmico de homogeneização e D, E e F após a homogeneização.

A corrosão da liga Mg-5Gd-3Ca sem tratamento térmico foi homogênea (**figuras 2-A, B e C**). Na amostra Mg-5Gd-3Ca ocorreu a deposição de óxido branco na superfície, o qual não foi identificado na amostra Mg-39Ca.

Nas **figuras 2-D, E e F** são mostradas as imagens das amostras homogeneizada após 24, 48 e 72 horas de imersão. Visualmente, os grupos de amostras com e sem tratamento

térmico sofreram corrosões com comportamento similares sem grandes diferenças na quantidade de material degradado.

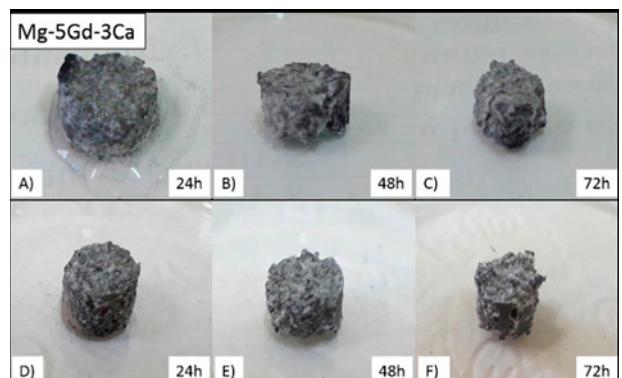


Fig. 2 - Amostras da liga Mg-5Gd-3Ca após 24, 48 e 72 horas de imersão em soro fisiológico a 36 °C. Imagens A, B e C referentes antes do tratamento térmico de homogeneização e D, E e F após a homogeneização.

No gráfico da **figura 3** são apresentados os dados da evolução da perda de massa em função do tempo de imersão das amostras. Os valores são referentes à massa de material no período do ensaio em relação à massa inicial, por isso todas as massas iniciais das amostras iniciam em 100%. A perda de massa apresentou um comportamento linear. Com a aplicação de regressões foi possível obter as equações das retas e fazer extrapolação dos resultados. Para a amostra Mg-39Ca no estado como não homogeneizada, a relação entre a perda de massa e o tempo de imersão foi $M = -1,27 \times h + 101,84$, onde M é a massa (em mg) e h o tempo de imersão em horas. Para a amostra após o tratamento de homogeneização, a equação foi $M = -0,66 \times h + 106,02$. No caso da amostra Mg-5Gd-3Ca foram encontradas as equações $M = -0,81 \times h + 99,17$ (não homogeneizada) e $M = -0,67 \times h + 99,30$ (homogeneizada).

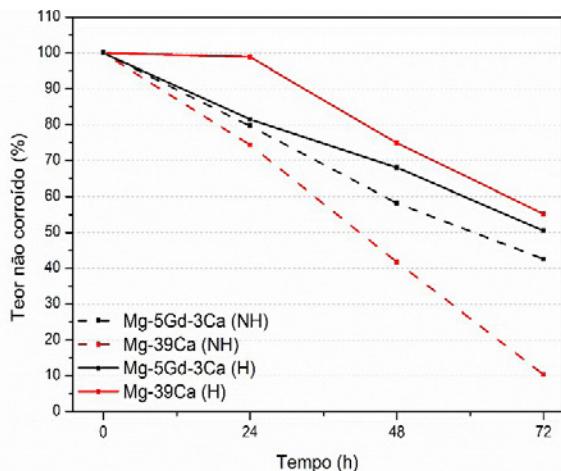


Fig. 3 – Perda da massa da amostra em função do tempo de imersão em soro fisiológico. Amostras não homogeneizadas (NH) e homogeneizadas (H).

Na **tabela 1** são apresentados os valores dos parâmetros das equações da reta calculados para cada uma das 4 condições de ensaio. A quarta coluna é uma estimativa da extração para um tempo de 500 horas de ensaio. Todos os valores são negativos, indicando a perda de massa da amostra. A ordem decrescente de perda de massa foi a Mg-39Ca (NH), seguida por Mg-5Gd-3Ca (NH), Mg-5Gd-3Ca (H) e Mg-39Ca (H).

Tab. 1 – Equações das retas obtidas por regressão linear da variação da perda de massa em função do tempo de imersão em soro fisiológico das ligas não homogeneizada (NH) e homogeneizada (H).

Mg-39Ca		
	Equação da reta	$h = 500$ h
NH	$M = 1,27 \times h + 101,84$	-533,16
H	$M = 0,66 \times h + 106,02$	-223,98
Mg-5Gd-3Ca		
	Equação da reta	$h = 500$ h
NH	$M = -0,81 \times h + 99,17$	-305,83
H	$M = -0,67 \times h + 93,30$	-235,70

Para estimar as taxas de corrosão foi utilizada a **equação 1** proposta pela norma NACE TM0169/G31-12a:

$$\text{Taxa de corrosão} = \frac{k - w}{\text{Área} \cdot T - D} \quad (1)$$

Onde,

K: constante para obter a taxa em mm/ano

W: perda de massa, em gramas

Área: área, em cm^2

T: tempo de imersão, em horas

D: densidade, em g/cm^3

A taxa de corrosão da liga Mg-39Ca sem tratamento térmico foi de aproximadamente 284,49 mm/ano. A liga Mg-5Gd-3Ca apresentou taxa de corrosão de 153,11 mm/ano no estado antes da homogeneização. Após o tratamento térmico de homogeneização a liga Mg-39Ca apresentou taxa de corrosão de 133,35 mm/ano e a Mg-5Gd-3Ca 119,97 mm/ano. Embora a menor perda de massa em função do tempo seja a da liga Mg-39Ca (H) (**tabela 1** e **figura 3**), a menor estimativa de taxa de corrosão calculada pela **equação 1** foi a da liga Mg-5Gd-3Ca (H). A causa provável desse resultado é a diferença de densidade das ligas. A liga Mg-5Gd-3Ca possui densidade levemente mais elevada, de $1,79\text{g}/\text{cm}^3$ e a liga Mg-39Ca, $1,65\text{g}/\text{cm}^3$.

Os resultados mostraram que a liga sem adição de gadolínio apresentou degradação superior (cerca de 1,8 vezes) no estado não homogeneizado. As amostras homogeneizadas da liga Mg-39Ca apresentaram taxa de corrosão semelhante a liga com Gd. Analisando as ligas individualmente, observou-se que com

o tratamento térmico ocorreu queda de 53% na taxa de corrosão da amostra Mg-39Ca. Este resultado mostra que o tratamento térmico aumenta a resistência à corrosão desta liga. No gráfico da **figura 4**, estão indicadas as taxas de corrosão com as barras de desvio padrão para cada material em cada condição de tratamento.

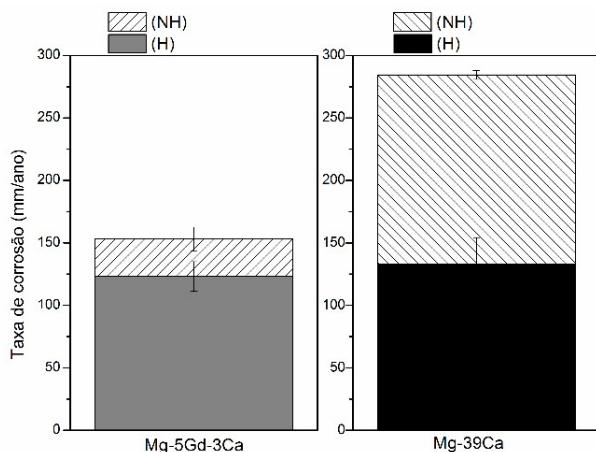


Fig. 4 - Taxas de corrosão médias (mm/ano) das ligas, não homogeneizada (NH) e homogeneizada (H).

Na **figura 5** são apresentadas as imagens dos ensaios de captação de H_2 para as amostras da liga Mg-39Ca, antes e após o tratamento térmico de homogeneização, onde as setas vermelhas indicam as posições das amostras. A **figura 5-A** é referente ao momento inicial da imersão da amostra, sem tratamento térmico. Desde o momento em que a amostra foi inserida no recipiente contendo soro fisiológico, houve a evolução de bolhas com fluxo visível a olho nu. Observou-se que após 17 horas de imersão, as amostras da liga Mg-39Ca (NH) estavam completamente degradadas e, por este motivo, o tempo de ensaio foi interrompido e padronizado para este tempo. O volume

de H_2 captado, foi aproximadamente 38 mL, apresentando a maior produção de hidrogênio entre todas as amostras.

Mostra-se na **figura 5-C** a amostra Mg-39Ca (H) após a homogeneização. Pode-se observar que o comportamento foi totalmente diferente em relação às amostras MG-39Ca (NH). Desde a imersão da amostra houve pouca liberação de H_2 e após 17 horas a amostra ainda apresentava volume sólido considerável. No detalhe do canto inferior direito, pode-se ver a amostra seca e fora da solução. Foram detectados poucos pontos de corrosão na superfície, mas sem afetar a integridade física da amostra. A captação de H_2 neste caso foi de 17 mL.

Ainda na **figura 5**, mostra-se as imagens do ensaio para a amostra Mg-5Gd-3Ca antes e após a homogeneização. Na **figura 5-E**, o momento inicial da imersão da liga sem tratamento térmico e na **figura 5-F** a mesma amostra após 17 horas. Quando a amostra foi imersa, houve a formação de um fluxo de bolhas, porém em menor quantidade que a amostra Mg-39Ca (NH). O eletrólito ficou turvo com um tom esbranquiçado e a amostra ao ser retirada e seca, estava com pontos de corrosão e identificada a deposição de um óxido branco na superfície. A captação de H_2 foi de 31mL. Pode-se observar na **figura 5-G** a amostra da liga Mg-5Gd-3Ca após a homogeneização e no momento inicial da imersão. O comportamento foi semelhante ao anterior. Após 17 horas (**figura 5-H**), o eletrólito apresentava alteração na cor e a amostra ao ser retirada e seca, também apresentava deposição de um óxido branco. A captação de H_2 neste caso foi de 27,5 mL apresentando a terceira maior produção de hidrogênio.

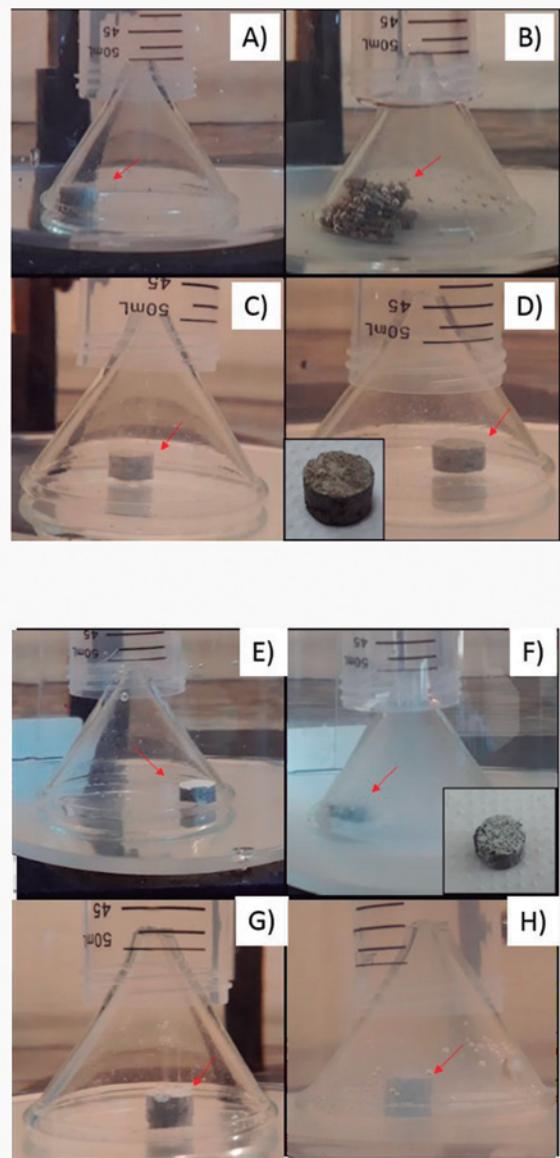


Fig. 5 – Ensaios de captação de hidrogênio. (A) início do ensaio da amostra Mg-39Ca sem tratamento térmico; (B) após 17 horas de imersão, mostrando a amostra totalmente desfeita; (C) amostra com tratamento térmico no início do ensaio; (D) amostra após 17 horas de imersão com poucos danos superficiais; (E) início do ensaio da amostra Mg-5Gd-3Ca sem tratamento térmico; (F) amostra Mg-5Gd-3Ca após 17 horas de imersão; (G) amostra com tratamento térmico

no início do ensaio e (H) após 17 horas de imersão.

Os dados da variação da liberação de hidrogênio com o tempo de ensaios são mostrados na **figura 6**. As retas indicam a variação do volume de H_2 (mL) em função do tempo (horas) de imersão em soro fisiológico, com as respectivas equações das retas. A amostra que apresentou a maior evolução de hidrogênio foi a Mg-39Ca (NH) e a equação da reta foi $V_{H_2} = 2,21 \times \text{tempo}(h)$. A situação que apresentou a segunda maior produção de hidrogênio foi para a amostra Mg-5Gd-3Ca (NH), com a equação da reta $V_{H_2} = 1,82 \times \text{tempo}(h)$. No caso da Mg-5Gd-3Ca (H), a produção de hidrogênio ficou em terceiro lugar com equação da reta $V_{H_2} = 1,62 \times \text{tempo}(h)$. Por fim, a amostra Mg-39Ca (H) apresentou o melhor resultado, com menor produção de hidrogênio e equação da reta $V_{H_2} = 1,03 \times \text{tempo}(h)$. A produção de H_2 pela amostra Mg-39Ca (H) foi inferior a 50% da amostra Mg-5Gd-3Ca (H) que apresentou a menor resistência à corrosão.

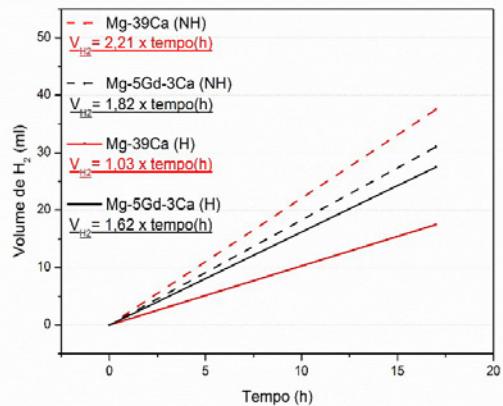


Fig. 6 – Captação de hidrogênio (mL) em função do tempo (horas) e as equações que regem as retas.

5. Conclusões

- O tratamento térmico de homogeneização das ligas Mg-5Gd-3Ca e Mg-39Ca aumenta a

resistência à degradação em solução salina e em soro fisiológico.

- As ligas sem adição de Gd apresentaram melhor resistência à corrosão com o tratamento térmico com queda de 53% na taxa de corrosão em relação à sem tratamento térmico.

- A liga Mg-5Gd-3Ca apresentou redução de 22% na taxa de corrosão.

- A amostra Mg-39Ca homogeneizada foi a

que produziu menor evolução de gás hidrogênio entre as quatro condições. O tratamento térmico reduziu a produção de H₂ em 50% para esta amostra.

Agradecimentos

Os autores agradecem ao PPGCM/IME e CNPq (141884/2017-7).

Referências bibliográficas

- [1] SONG G L, ATRENS A. Corrosion Mechanisms of Magnesium Alloys. *Advanced Engineering Materials*. 1999; 1:11-33.
- [2] STAIGER M P, PIETAK A M, HUADMAI J, DIAS G. Magnesium and its alloys as orthopedic biomaterials: A review. *Biomaterials*. 2006; 27(9):1728-1734.
- [3] CHEN Y, XU Z, SMITH C, SANKAR J. Recent advances on the development of magnesium alloys for biodegradable implants. *Acta Biomaterialia*. 2014; 10(11):4561-4573.
- [4] KANNAN M B, RAMAN R K S. In vitro degradation and mechanical integrity of calcium-containing magnesium alloys in modified-simulated body fluid. *Biomaterials*. 2008; 29(15):2306-2314.
- [5] PERSAUD-SHARMA D, MCGORON A. Biodegradable Magnesium Alloys: A Review of Material Development and Applications. *Journal of biomimetics, biomaterials, and tissue engineering*. 2012; 12:25-39.
- [6] SONG G, ATRENS A, STJOHN D. An hydrogen evolution method for the estimation of the corrosion rate of magnesium alloys. *Magnesium Technology*. 2001;255-262.

Análise de comportamento de malware utilizando redes neurais recorrentes - uma abordagem por intermédio da previsão de opcodes

Davi Gomes de Albuquerque, Lincoln de Queiroz Vieira, Ricardo Sant'Ana* e Julio Cesar Duarte

Instituto Militar de Engenharia (IME)

Praça General Tibúrcio, 80, 22290-270, Praia Vermelha, Rio de Janeiro, RJ, Brasil.

*ricardo.santana@ime.eb.br

RESUMO: A internet é palco de milhares de ataques cibernéticos todos os dias. Uma das maiores ameaças que afetam empresas e até usuários comuns são malware. Companhias de antivírus tentam aumentar a eficácia dos métodos de detecção de vírus, mas o grande aumento do número de variações de malware com o uso de técnicas como obfuscation aumenta a cada dia, tornando seu trabalho cada vez mais complexo. A utilização de redes neurais tem se mostrado cada vez mais presente na construção de algoritmos decisórios, inclusive na definição de classificadores de malware. Esse trabalho tem o objetivo de aplicar redes neurais recorrentes para prever os opcodes de um malware e, em seguida, classificá-los. Essa abordagem é inovadora pois, nos trabalhos analisados, não encontramos uma solução que utilize a previsão de opcodes como entrada para um classificador de família de malware. O classificador de famílias obteve uma acurácia média de 92%.

PALAVRAS-CHAVE: Classificação. Malware. Previsão. Opcodes. Redes Neurais Recorrentes. Long-Short Term Memory

1. Introdução

Apesar do desenvolvimento significativo de segurança da informação, o número de programas maliciosos, chamados de *malware*, vêm crescendo de forma espantosa a cada ano. Segundo o relatório da empresa *McAfee*, quase 150 milhões de novos *malware* foram criados durante os 8 primeiros meses de 2018 [1].

Malware são software projetados com intenções maliciosas que podem ser usados para espionagem, extorsão, sabotagem e para

ABSTRACT: The internet is the scene of thousands of cyber attacks every day. One of the biggest threats affecting businesses and even ordinary users is malware. Antivirus companies are trying to increase the effectiveness of virus detection methods, but the huge increase in the number of malware variations using techniques such as obfuscation is increasing every day, making their work much more complex. The use of neural networks has been shown to be increasingly present in the development of decision algorithms, including in the definition of malware classifiers. This work aims to apply recurrent neural networks to predict the opcodes of a malware and then classify them. This approach is innovative because, in the studies analyzed, we did not find a solution that uses the prediction of opcodes as input to a classifier of the malware family. The family classifier obtained an average accuracy of 92%

KEYWORDS: Classificação. Malware. Previsão. Opcodes. Redes Neurais Recorrentes. Long-Short Term Memory

executar tarefas não autorizadas [2]. *Malware* podem, ainda, ser dos mais variados tipos como *worms*, *trojans*, *rootkits*, *spyware*, *adware*[3] e pertencerem a diversas famílias.

O método clássico para detecção de vírus utiliza um banco de dados de assinaturas digitais de *malware* que realiza a busca por padrões de código malicioso em arquivos. Porém, técnicas que alteram levemente o código malicioso burlam as assinaturas digitais dos antivírus com relativa facilidade. A criação de métodos para detectar

malware através de seu comportamento, por outro lado, é bastante onerosa [4].

O processamento de grande quantidade de informação relacionada a *malware* exige um esforço enorme das companhias de antivírus. Aliado a isso, existe o aumento do número de famílias de *malware* e de novas variantes dentro da família [5].

Assim, o trabalho de um analista, devido a essa evolução de códigos maliciosos, tem ficado cada vez mais complexo, cansativo e custoso. É fundamental, portanto, a automatização de tarefas dentro da análise, classificação e detecção do *malware*.

O presente trabalho teve a finalidade de elaborar, implementar e testar um algoritmo para detecção de famílias de *malware* utilizando redes neurais recorrentes. Mais especificamente, foram utilizadas dois tipos de redes neurais: uma *Long Short Term Memory* (LSTM) para realizar a predição de *opcodes*, onde se utilizou um modelo para cada família e, em seguida, uma rede neural artificial (RNA) para classificar as amostras em famílias. A partir dos modelos LSTM e RNA treinados, os *opcodes* extraídos de uma amostra de *malware* são utilizados como entrada para cada um dos modelos de LSTM. As acurárias de predição de *opcodes* de cada modelo LSTM serve de entrada para a RNA que produzirá uma predição de família da amostra.

Este artigo foi estruturado da seguinte forma: a Seção 2 descreve os principais básicos necessários ao entendimento do artigo. A Seção 3 apresenta a implementação desenvolvida e os principais resultados obtidos. A seção 4, apresenta uma breve discussão e conclusão do trabalho, além de sugestões para trabalhos futuros.

2. Referencial teórico

Nessa seção, serão apresentados os conceitos relacionados a *malware* e análise de *malware*, uma introdução sobre redes neurais recorrentes e LSTM, além das definições dos principais hiperparâmetros do modelo LSTM.

2.1 Malware e análise de malware

Malware é um *software* criado para ter efeitos indesejados ou danosos a um computador e oferecer grande ameaça para a segurança de sistemas computacionais [6]. Esses códigos maliciosos, muitas vezes, exploram vulnerabilidades presentes em programas, como problemas de programação na lógica de um código ou má administração do uso de memória [7].

Os vários tipos de estruturas e comportamentos de códigos maliciosos permitem agrupá-los em várias categorias. De acordo com [8], os *malware* podem ser classificados em três gerações baseando-se em três características do código malicioso: *payload*, método de propagação e utilização das vulnerabilidades. Outra classificação, baseada no comportamento do código malicioso, é proposta por [9], que divide os *malware* em *worm*, *backdoors*, *botnet*, *downloader*, *information-stealing malware*, *launcher*, *rootkit*, *scareware* e *spam-sending malware*.

Além disso, devido a evolução das técnicas de detecção de *malware* com antivírus, os autores de *malware* tem criado códigos cada vez mais complexos. A técnica mais utilizada hoje por antivírus é a de assinatura digital, em que um *malware* é identificado quando apresenta, em seu código, trechos considerados maliciosos ou quando o resultado de uma função de *hash* do

executável encontra-se numa tabela pré-definida. Essa evolução dos *malware* permite uma nova classificação: *Malware* polimórficos e *malware* metamórficos.

Malware polimórficos [10] e *malware* metamórficos [11] são capazes de alterar seus códigos através da troca da ordem de execução de instruções, introdução de instruções desnecessárias, transposição de código, substituição de instruções equivalentes entre outras técnicas. Duas são as consequências do uso dessas técnicas: a detecção desses *malware* pelos antivírus torna-se mais difícil e há um incremento no crescimento do número de variações de *malware*.

A análise de *malware* é a tarefa de examinar um *malware* de forma a entender como eles funcionam, como identificá-los e como eliminá-los. Seu propósito é, geralmente, prover informações necessárias para responder a uma intrusão em uma rede, empresa ou computador específico: descobrir o que ocorreu, assegurar-se de localizar todas as máquinas e arquivos infectados. Nessa tarefa, é importante determinar exatamente o que um artefato pode fazer, como detectá-lo e como medir e conter os danos causados [9]. Além da análise, a detecção e a classificação de *malware* são duas tarefas executadas por companhias de antivírus.

Durante a análise do *malware*, os executáveis são analisados para verificar se exibem comportamento malicioso. Com as informações coletadas sobre o *malware*, é possível armazenar seus dados em banco de dados para futura referência. A análise de *malware* feita por um analista pode ser dividida em duas fases: a estática e a dinâmica. De acordo com [9], pode-se subdividir as análises estática e dinâmica em

básica e avançada.

As quatro fases, como mostra a **figura 1**, são realizadas em ciclo para que seja possível realimentar o sistema e recomeçar em qualquer fase da análise, utilizando as novas informações. Cada uma das fases será descrita a seguir.

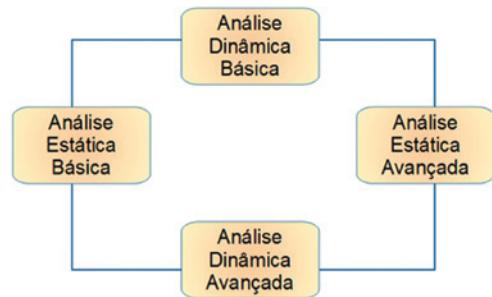


Fig. 1 – Fase da Análise de Malware de acordo com [9]: Análise estática básica, Análise dinâmica básica, Análise estática avançada e Análise dinâmica avançada.

Na análise estática básica, o artefato é analisado sem ser executado, procurando por padrões maliciosos em sua estrutura. Por meio desta análise, pode-se confirmar se um código é malicioso, adquirir informações básicas de suas funcionalidades e prover informações sobre possíveis conexões em redes. Por exemplo, por meio do aplicativo *Dependency Walker*, o analista é capaz de obter uma lista das funções chamadas dinamicamente pelo executável e, dessa forma, é possível prever certos comportamentos do artefato em análise. É uma fase rápida e que utiliza técnicas simples para levantar o máximo de informações do artefato mas é ineficaz contra *malware* mais sofisticados que utilizam técnicas de obfuscção, por exemplo. Os resultados obtidos dessa análise são adicionados a um relatório de análise.

A análise dinâmica básica consiste em

executar o *malware* em um ambiente virtual controlado como *sandbox* ou máquina virtual e observar seu comportamento. Com ferramentas previamente instaladas, é possível verificar quais novos processos são gerados pelo artefato, quais conexões de rede ele tenta criar, quais DLLs são chamadas e quais chaves de registro são modificadas. A medida que o analista descobre novas informações sobre o *malware*, ele atualiza o relatório com as novas informações encontradas.

A análise estática avançada é feita realizando a engenharia reversa do código do *malware* com o auxílio de um *disassembler*. Por meio da análise das instruções *assembly* é possível descobrir o que o programa faz (sem executá-lo). Dessa forma, a análise das instruções é interessante pois pode revelar comportamentos do artefato que não foram, por exemplo, identificados durante a fase de análise dinâmica. No entanto, essa técnica é mais complexa que as demais pois exige do analista um conhecimento profundo em programação *assembly* e conceitos específicos sobre sistemas operacionais nos quais o *malware* se encontra.

A análise dinâmica avançada é realizada por meio de um *debugger* que é utilizado para examinar o estado interno de um executável (pilha, registradores) e para controlar o fluxo de chamada de cada instrução do código. Esse tipo de análise é feita quando se deseja obter informações mais precisas sobre o comportamento do artefato. Além disso, essa fase pode ser considerada complexa, uma vez que o analista deve dominar tanto o *assembly* como as principais estruturas (pilha, registradores).

Com a grande variedade e o grande crescimento do número de *malware*, as técnicas supracitadas para análise e detecção de programas

maliciosos tem se tornado cada vez mais ineficazes – o tempo necessário para realizar uma análise manual é incompatível com a demanda. Já a detecção automatizada por meio de assinatura digital exige bancos de dados atualizados cada vez maiores e a criação de métodos para detectar *malware* através da semântica, por outro lado, é bastante complicada [4]. O método de análise proposto nesse trabalho se baseia na análise automatizada estática que, tem potencial para processar grande quantidade de amostras de *malware* é bem mais simples, quando comparada a análise automatizada dinâmica.

2.2 Redes neurais recorrentes

Redes neurais recorrentes são um tipo específico de redes neurais, onde as unidades intermediárias alimentam não apenas as camadas seguintes, mas a mesma camada ou anteriores. Ao alimentar uma célula de uma rede neural recorrente simples, é aplicada uma função de ativação, de forma que, quanto mais a informação flui pela camada intermediária, maior a quantidade de composições de funções de ativação. Ao calcular o gradiente descendente, e devido a regra da cadeia, quanto maior a sequência de composição de funções de ativação menor é o gradiente. Isso resulta em valores cada vez menores de gradiente a cada composição e, portanto, pouca ou nenhuma atualização dos pesos da rede. Esse fenômeno é chamado de *vanishing gradient*. Para solucionar esse problema foi desenvolvida a arquitetura *long short term memory* (LSTM).

2.3 Long Short Term Memory

Na arquitetura de uma LSTM existem 3 portas: porta de entrada, porta de esquecimento e porta

de saída. A porta de entrada é responsável por quanto o estado anterior influencia no estado atual. A porta de esquecimento controla quando parte da informação deve ser atualizada ou esquecida. A porta de saída controla quais partes da informação devem ser utilizadas na saída da célula. Cada porta é associada a pesos que regulam a atuação das mesmas [12]. Pode-se ver na **figura 2**, uma representação da LSTM. A chave para LSTMs é o estado da célula (*cell state*): a linha horizontal que percorre a parte superior do diagrama. A informação que flui pelo estado da célula sofre poucas modificações e, portanto, não sofre do efeito de *vanishing gradient*, ou seja, o gradiente não é atenuado com um número grande de estados [13].

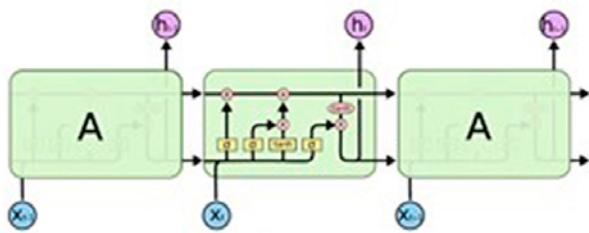


Fig. 2 – Camada LSTM expandida. Fonte: [14]

A lógica dentro da porta do esquecimento pode ser vista na **equação (1)**, a lógica da porta de entrada é dada pelas **equações (2) e (3)**. A composição da saída dessas duas portas no estado da célula (cell state) pode ser vista na **equação (4)**. A lógica do portão do esquecimento está apresentada em (5) e (6):

$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma (W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$C_t = \tanh (W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

$$C_t = f_t * C_{t-1} + i_t * C_t \quad (4)$$

$$o_t = \sigma (W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t * \tanh (C_t) \quad (6)$$

onde $*$ é o produto elemento por elemento e a função σ é a função sigmoide. A **figura 3** apresenta a identificação de cada um dos portões de acordo com a nomenclatura utilizada nas equações anteriores.

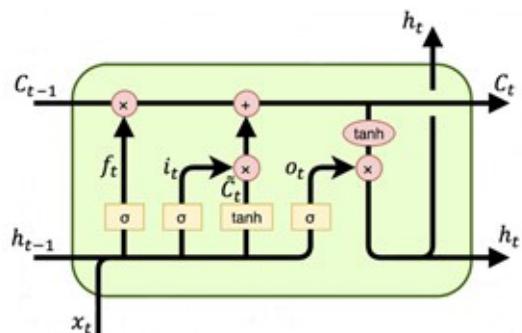


Fig. 3 – Célula LSTM detalhada. Baseado em [14].

Podemos dizer que f_t indica o quanto da informação anterior deve ser preservado, baseado na entrada x_t do estado atual e na saída h_{t-1} do estado anterior. Esse termo representa a camada da porta de esquecimento. Já i_t indica o que deve ser atualizado, agindo como a camada da porta de entrada, e C_t elenca novas informações candidatas a serem adicionadas. Combinando i_t e C_t temos a atualização do estado C_{t-1} para C_t . O termo o_t define a saída que influenciará os próximos estados, caracterizando a porta de saída. Os termos b_f , b_i e b_o presentes nas equações correspondem ao *bias*, critérios de preferência de uma hipótese de sobre outra, e os pesos W_f , W_i , W_C e W_o presentes nas camadas da LSTM estão profundamente relacionados ao processo de aprendizado, sendo esses parâmetros adaptados com o intuito de diminuir o erro e obter um modelo eficiente [14].

2.4 Hiperparâmetros do modelo LSTM

Para o presente trabalho foram realizadas

escolhas de hiperparâmetros relacionadas à implementação da LSTM na biblioteca Keras. Abaixo temos uma relação dos hiperparâmetros relevantes e sua definição.

- *lookback*: indica quantas etapas anteriores devem ser importantes na predição da próxima etapa [15].
- *batch size*: indica o número de exemplos de treinos usados em uma iteração do treino. A rede neural recorrente não pode tratar todos os exemplos de treino por limitação de memória.
- *epoch*: representa uma varredura pelo conjunto inteiro do conjunto de dados de treinamento durante o treinamento do modelo. Faz-se necessário usar mais de um *epoch*, para melhor adaptar o modelo a todo conjunto de dados [16].
- *time step*: reflete o ponto de observação da amostra, ou seja, qual o tamanho temporal da sequência de amostras que vai alimentar o modelo. Mesmo valor de *lookback*.
- otimizadores: tem função de acelerar o processo de treino produzindo resultados compatíveis. O otimizador escolhido foi o *adam* por ser o mais utilizado. Existem outros otimizadores para casos mais específicos [17].

3. Implementação e resultados

Nesta seção serão apresentados as implementações e experimentos realizados além dos resultados. Será descrito como foi realizada a extração e codificação dos *opcodes*, além do treinamento do modelo de predição de *opcodes* e do modelo de classificação em famílias com seus respectivos resultados.

3.1 Extração de Opcodes

A base de dados utilizada foi obtida do Desafio de Classificação de Malware da Microsoft de 2015 [26] e contém 10.868 amostras de *malwares* classificados em 9 famílias. Cada amostra é composta por um arquivo do tipo *asm*, contendo o código em *assembly* do *malware*, e um arquivo do tipo *byte*, com um extrato do executável em binário. Os *opcodes* foram extraídos dos arquivos *asm* por meio de rotinas de processamento de texto. A **tabela 1** mostra a quantidade original de amostras por classe.

Tab. 1 – Distribuição das famílias da malware na base de dados da Microsoft [26].

Ordem	Nome da Família	Quantidade de Amostras
1	Ramnit	1541
2	Lollipop	2478
3	Kelihos_ver3	2942
4	Vundo	475
5	Simda	42
6	Tracur	751
7	Kelihos_ver1	398
8	Obfuscator.ACY	1228
9	Gatak	1013

3.2 Codificação dos Opcodes

Para codificar os *opcodes* foi utilizada a técnica *word2vec*, que é um método para criar *embeddings* entre palavras ao transformá-las em vetores. O *word2vec*, como o nome sugere, tenta descrever palavras como forma de vetores, agrupando palavras com semânticas similares de forma mais próxima, onde cada dimensão do vetor é entendida como uma característica da palavra. Geralmente, palavras da língua inglesa, codificadas com *word2vec*, possuem dimensão 300. Ainda assim, é possível visualizar os vetores em um

gráfico duas dimensões, utilizando técnicas de redução de dimensionalidade, como por exemplo a análise de componentes principais (*Principal Component Analysis - PCA*). Na **figura 4** temos nome de países e suas respectivas capitais projetadas em 2 dimensões:

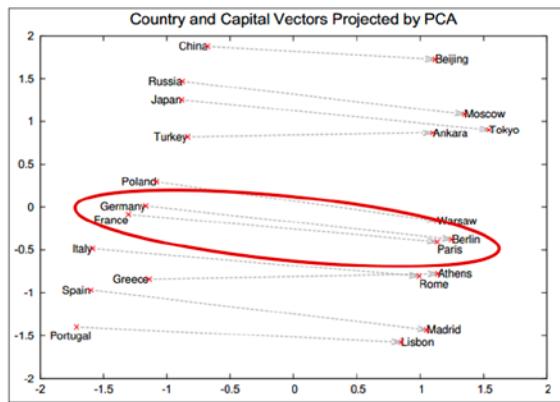


Fig. 4 – Vetores de exemplos com países e capitais. Fonte: [18]

Podemos observar que os vetores de diferença entre capital e país é quase constante, o que mostra que a técnica de *word2vec* foi capaz de compreender a relação entre esses dois tipos de palavras (nome do país e capital).

Neste trabalho foi utilizado uma codificação fruto da aplicação do método *word2vec* às instruções *assembly*. Esses vetores foram gerados por [19], criando uma base de dados de instruções *assembly* na forma de vetores de dimensão 100.

A partir dos vetores criados por [19], que consideram o *opcode* e os operandos para formar o vetor, foram criados novos vetores para representar apenas os *opcodes*. Isso foi realizado calculando a média entre todos os vetores existentes que continham o *opcode* desejado, independente do operando. Por exemplo, para calcular o vetor correspondente

ao *opcode mov* foi calculada a média entre os 107.339 vetores presentes na base de dados de [19]. O resultado deste procedimento aplicado a cada um dos *opcodes* extraídos neste trabalho foi denominado de *opcode2vec*.

Para observarmos como essa codificação representa a semântica do *opcode* apresentamos na **figura 5** quatro instruções projetadas em duas dimensões: *add* (ponto vermelho), *sub* (ponto verde), *pop* (ponto amarelo) e *push* (ponto azul). É importante ressaltar que os eixos X e Y da **figura 5** não tem significado pois são eixos de projeção do vetor de dimensão 100.

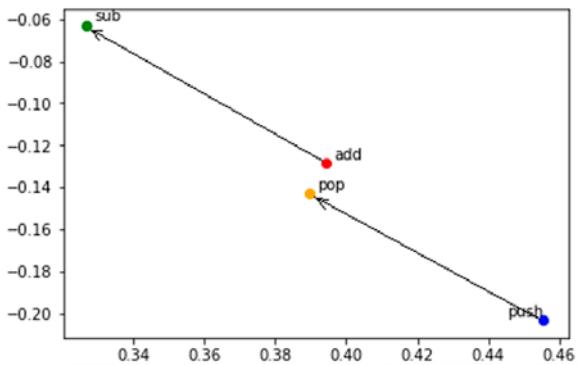


Fig. 5 – Exemplo de vetores add, sub, pop e push projetados no plano e utilizados neste trabalho.

Da **figura 5**, podemos observar que a diferença *sub - add* é aproximadamente igual à diferença *pop - add*. Desta forma, podemos concluir que o vetor diferença representa função oposta. Essa codificação de *opcodes* foi utilizada no conjunto de dados de treinamento e testes. A **figura 6** apresenta a sequência *push, push, call, jnz, mov, ..., push, int* e *xor* sendo codificada para vetores. Podemos observar que *opcodes* iguais são codificados para vetores iguais.

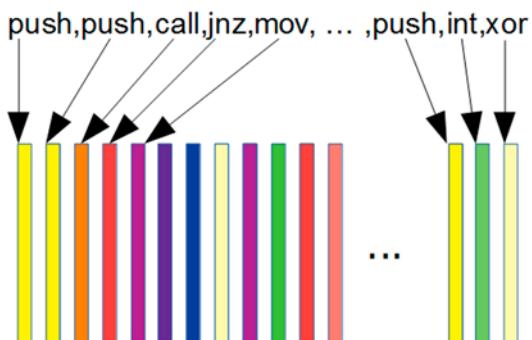


Fig. 6 – Processo de codificação de uma sequência de opcodes em vetores utilizando opcodes2vec.

Para alimentar o modelo LSTM, o tamanho da sequência temporal (*lookback*) foi definida como 5, ou seja, sequências de entrada de 5 opcodes consecutivos (entrada) para predizer o sexto elemento, a saída desejada.

Exemplificando, um *malware* que contém a sequência de opcodes: op1 (tempo 1), op2 (tempo 2), op3 (tempo 3), op4 (tempo 4), op5 (tempo 5), op6 (tempo 6) tem como entrada (op1, op2, op3, op4, op5) e como rótulo de saída op6.

A figura 7 apresenta um exemplo onde, a partir de uma sequência de opcodes codificados foram geradas 3 sequências contendo a entrada de tamanho 5 e a saída desejada (sexto elemento da sequência).

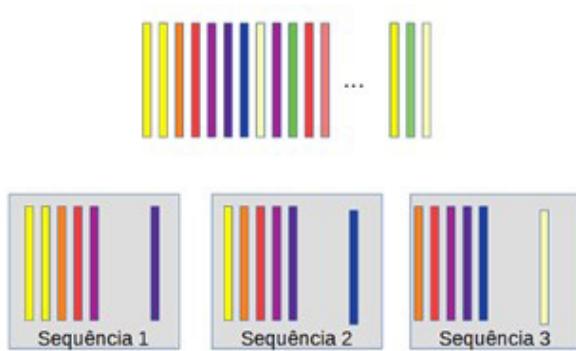


Fig. 7 – Sequência de vetores codificados gerando as primeiras sequências com loopback igual a 5.

3.3 Treinamento dos modelos de predição de *Opcodes*

As entradas das sequências de *opcodes* codificadas de cada classes e a saída desejada servem para treinar a LSTM. A figura 8 apresenta uma sequência de vetores codificados (entrada) e a saída desejada (rótulo) utilizados para treinar a rede LSTM para predição *opcodes*. Para gerar um modelo LSTM para cada classe, esse processo é repetido para cada sequência de entrada e saída desejada das amostras daquela classe. Ao final, teremos 9 modelos preditores de *opcodes*: um para cada família de *malware*.

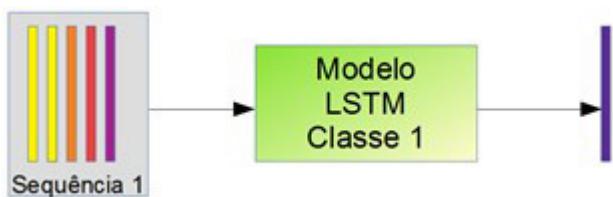


Fig. 8 – Cada sequência de vetores codificados de tamanho 5 o rótulo de saída são utilizados para treinar a rede LSTM.

A dimensão da LSTM é a dimensão do vetor de saída e o *time step* do modelo é igual ao valor do *lookback*, ou seja, 5. O treinamento foi feito minimizando o erro quadrático médio, até a convergência do erro, com *batch size* = 100 por 100 épocas.

A taxa de acertos da predição de *opcodes* para o conjunto de testes, utilizando elementos da própria família, é apresentado na **tabela 2** dentro das respectivas classes.

Tab. 2 – Avaliação do treino dos modelos de predição de opcodes usando opcode2vec

Nome da Família	Acurácia de Predição de opcodes
Família 1	40,6%
Família 2	37,9%
Família 3	73,4%
Família 4	53,6%
Família 5	94,0%
Família 6	91,3%
Família 7	69,1%
Família 8	80,3%
Família 9	90,1%

Da tabela, podemos concluir que os modelos LSTM foram capazes de modelar as amostras de treinamento de cada uma das classes. As classes 1, 2 e 4 ficaram com as menores acuráncias na predição de *opcodes*; provavelmente o modelo LSTM utilizado não foi capaz de modelar a diversas sequências encontradas nesta classe. Ainda assim, isso não é um fator limitante pois o objetivo principal é que o modelo da classe deva ser capaz de fazer predições da própria classe melhor que os outros modelos.

A saída dos modelo LSTM – que é a acurácia de predição de *opcodes*, foi utilizada para treinar uma RNA.

3.4 Treinamento da RNA para classificação

Após o treino dos modelos utilizando LSTM, cada um dos nove modelos realizou predições sobre cada uma das amostras das nove famílias e as acuráncias foram obtidas. A **figura 9** apresenta as sequências de uma determinada amostra alimentando os 9 preditores de *opcodes* e produzindo um vetor de acuráncias de dimensão 9.

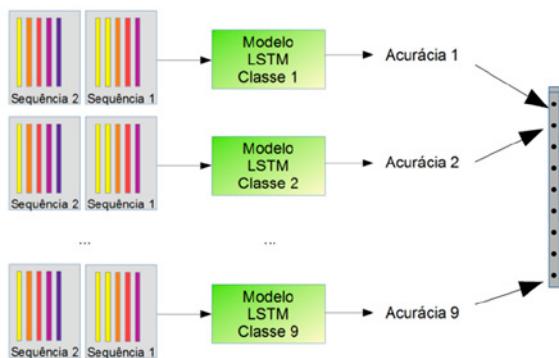


Fig. 9 – As sequências de opcodes de uma determinada amostra alimentam os 9 modelos preditores de opcodes. Um vetor de acuráncias de dimensão 9 é produzido. Na figura apresentamos apenas as duas primeiras sequências de uma determinada amostra.

Com os resultados das predições (vetor de acuráncias) de cada amostra de *malware*, foi possível treinar uma rede neural artificial para classificar cada amostra em famílias. A rede neural utilizada possui duas camadas sendo que a primeira camada com 9 neurônios, duas camadas escondidas com 12 e 10 neurônios respectivamente e a camada de saída com 9 neurônios. A **figura 10** apresenta um vetor de acuráncias (para as diversas amostras de *malware*) utilizados para treinamento de rede neural artificial.

A última camada foi utilizada uma função de ativação chamada de *softmax*, que força a soma de todas as saídas ser igual a 1, permitindo uma interpretação probabilística de cada uma das 9 saídas da rede neural.

Para o treinamento da rede neural artificial, foram utilizadas 80% do conjunto total de amostras de treinamento, totalizando 8604 amostras. A entrada da rede neural artificial é um vetor de 9 posições: a acurácia de predição de *opcodes* para cada uma das 9 famílias.

Dessa forma, a rede neural artificial aprendeu, portanto, as relações entre as acuráncias de forma a classificar corretamente os *malware* em famílias. Para o teste da rede neural artificial, todo o conjunto de teste foi utilizado.

vetor de acuráncias

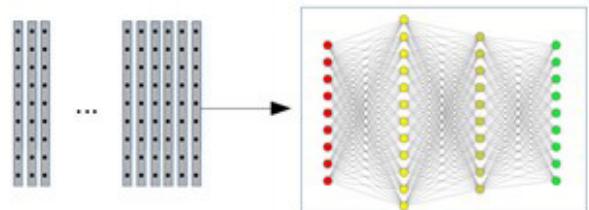


Fig. 10 – As sequências de opcodes de uma determinada amostra alimentam os 9 modelos preditores de opcodes. Um vetor de acuráncias de dimensão 9 é produzido. Na figura apresentamos apenas as duas primeiras sequências de uma determinada amostra.

Para definir os melhores hiperparâmetros da rede neural artificial, foram realizados experimentos com diversos valores de hiperparâmetros:

- batch size: 2,5, 10, 20;
- epochs: 500, 1000, 1500;

Função de perda: foram testadas as funções binary cross entropy e mean squared error.

A rede neural artificial obteve melhores resultados com os seguintes hiperparâmetros:

- batch size: 20;
- epochs: 1000;
- Função de perda: *mean squared error*.

A acurácia do modelo para o conjunto de dados de teste foi de 92%. A **tabela 3** mostra os resultados da rede neural artificial utilizando outras métricas para cada uma das famílias.

Considerando que a distribuição de amostras entre as famílias não é balanceada, consideramos utilizar o *f1-score*. O *f1-score* é uma medida de desempenho comumente utilizada para classificações com múltiplos rótulos [21]. É calculada com a média harmônica dos valores abrangência e precisão, ou seja, ela busca um equilíbrio entre esses valores [22], penalizando-o pelo menor dos valores. De acordo com [22], valores altos de precisão e abrangência significam que a classe está bem definida no modelo.

Dessa forma, as famílias mais bem definidas no modelo são as famílias 3,4 e 8, pois seus valores de *f1-score* são os mais altos. Com exceção da família 5, todas as famílias produziram um *f1-score* acima de 0.80.

Tab. 3 – Resultados do classificador

Nome da Família	Precisão	Abrangência	f1-score
Família 1	0,88	0,92	0,90
Família 2	0,84	0,97	0,90
Família 3	1,00	1,00	1,00
Família 4	0,97	0,96	0,97
Família 5	1,00	0,21	0,35
Família 6	0,94	0,87	0,91
Família 7	1,00	0,83	0,91
Família 8	0,97	0,86	0,92
Família 9	0,91	0,74	0,81

Possivelmente, o valor baixo de *f1-score* para a família 5 se justifique pela pequena quantidade de amostras disponíveis dessa classe. O modelo LSTM preditor de *pcodes* da classe 5 obteve uma acurácia de 94% - no entanto, esse modelo não deve ser capaz distinguir amostras da classe 5 de outras, produzindo um resultado ruim para o *f1-score* de classificação da rede neural artificial.

Para comparação de resultados, Rocca [23] é referência principal. Esse autor realiza a extração de *pcodes* da mesma base de dados da *Microsoft* utilizada neste trabalho. Em seguida, implementa uma codificação de bigramas de *pcodes* utilizando o *Term Frequency - Inverse Document Frequency* - TF-IDF [24]. O TF-IDF tenta indicar a importância de uma palavra (*opcode*) de um binário em relação a uma coleção de binários.

A classificação em famílias é feita por meio de uma rede auto-codificadora pré-treinada para reconstruir as entradas TF-IDF. A saída da metade codificadora dessa rede é usada como entrada para uma rede neural artificial para classificação em famílias. Desta forma, pode-se observar que essa solução possui as mesmas limitações de classificação de um artefato em uma família deste trabalho, ao utilizar como informação para classificação apenas os *pcodes* das amostras. Desta forma, os resultados passam a ser comparáveis com o que foi desenvolvido neste trabalho.

A **tabela 4** apresenta os resultados obtidos pelo classificador [23]. Podemos observar que os valores de *f1-score* foram, no geral, superiores aos alcançados neste trabalho. Pode-se observar, no entanto, algumas semelhanças: a família 5 obteve, em ambos os casos, um valor de *f1-score* mais baixo, possivelmente por causa da pouca quantidade de amostras da família 5. Além disso, o maior *f1-score* foi obtido em ambos os casos na família 3, pois essa

possui a maior quantidade de amostras.

Tab. 4 – Métricas por classe para o classificador final – Rede 4, 4 camadas ocultas, conjunto de dados de bigramas. Fonte: [23]

Nome da Família	Precisão	Abrangência	f1-score
Família 1	0,97	0,94	0,96
Família 2	0,97	0,98	0,97
Família 3	0,99	0,99	0,99
Família 4	0,96	0,89	0,93
Família 5	0,72	0,72	0,72
Família 6	0,89	0,96	0,92
Família 7	0,97	0,95	0,96
Família 8	0,92	0,95	0,93
Família 9	0,97	0,94	0,96

Podemos observar que, em sua grande maioria, o modelo proposto obteve melhores resultados de *f1-score* que os apresentados neste trabalho [23]. De qualquer forma, devido a limitação de tempo, não tivemos oportunidade de avaliar o impacto do ajuste de diversos hiperparâmetros no *f1-score* da classificação de cada família. Essas sugestões de melhoria serão apresentadas na seção seguinte na forma de trabalhos futuros.

4. Discussão e conclusão

A criação de modelos preditivos de *opcodes* para cada uma das 9 famílias de *malware* permite a implementação de um classificador em famílias baseado nas acurárias dos modelos preditores de forma automatizada, motivados pela dificuldade em classificar um grande volume de artefatos.

Uma dificuldade inesperada foi encontrada na extração dos *opcodes*, pois foi necessário o desenvolvimento de um *script* especializado em extrair *opcodes* dos textos exportados pelo IDA – seria mais interessante utilizar bibliotecas especializadas para essa extração, mas para isso seria necessário possuir o binário do artefato malicioso. Tão importante quanto a correta extração de opcodes foi a correta codificação dos opcodes. Neste trabalho,

os resultados mais interessantes foram obtidos utilizando uma codificação de opcodes baseadas no word2vec utilizando uma base de codificação desenvolvida [19].

Foram montados 9 modelos preditores de opcodes usando rede neural recorrente baseada em LSTM: consideramos que os resultados obtidos na predição de opcodes pelos 9 modelos foi adequada para desenvolver um classificador de família baseado na saída dos preditores de opcodes.

Os resultados apresentados pelo classificador foram satisfatórios, apesar do resultado ruim para a família 5. Além disso, os resultados obtidos em [23], que também utiliza a informação de extração de opcodes, foram superiores aos apresentados nestes trabalhos, motivando a sugestão de trabalhos futuros que possam melhorar os resultados aqui apresentados.

Assim, como trabalhos futuros, pode-se sugerir:

- Extrair opcodes e operandos, pois dessa forma são obtidas mais informações para o modelo de predição. Para isso, seria necessário gerar uma base mais completa de codificação baseada no word2vec que codifique opcodes e operandos juntos.
- Aplicar as técnicas aqui apresentadas em outra base de dados de malware, que não estivesse limitada a 10.000 amostras.
- Utilizar outras codificações de opcodes/operandos como fastText [20] ou BERT [21].
- Ajustar alguns hiperparâmetros como: o time step para predição, o número de épocas de forma a garantir a convergência do erro (por exemplo, por meio do parâmetro early stop).
- Utilizar uma estrutura LSTM mais complexa, como LSTM empilhadas, sempre levando em conta a possibilidade de overfitting e maior duração dos treinamentos.

Referências bibliográficas

- [1] [1] RAJ SAMANI, C. B. McAfee Labs Threads Report. Disponível em: <<https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-dec-2018.pdf>>. Acesso em: 31 dec. de 2018.
- [2] [2] RAYMOND J. CANZANESE, J. Detection and Classification of Malicious Processes Using System Call Analysis. 2015. 1 f. Tese (Doctor in Philosophy) – Drexel University, Filadélfia, Estados Unidos, 2015.
- [3] [3] GAVRILUT, D.; CIMPOESU, M.; ANTON, D. ; CIORTUZ, L. Malware detection using machine learning. Proceedings of the International Multiconference on Computer Science and Information Technology, v. 4, p. 735–741, 2009.
- [4] [4] UCCI, D.; ANIELLO, L. ; BALDONI, R. Survey on the usage of machine learning techniques for malware analysis. Computers & Security, v. 81, p. 123–147, 2017.
- [5] [5] AHMADI, M.; ULYANOV, D.; SEMENOV, S.; TROFIMOV, M. ; GIACINTO, G. Novel feature extraction, selection and fusion for effective malware family classification. In: CODASPY, 16., 2016. Anais... [S.I.: s.n.], 2016, p. 183–194.
- [6] [6] YUXIN, D.; SIYI, Z. Malware detection based on deep learning algorithm. Neural Computing and Applications, v. 31, n. 2, p. 461–472, 2019.
- [7] [7] KUMAR, P. D.; NEMA, A. ; KUMAR, R. Hybrid analysis of executables to detect security vulnerabilities: security vulnerabilities. In: PROCEEDINGS OF THE 2ND INDIA SOFTWARE ENGINEERING CONFERENCE, 2., 2009. Anais... [S.I.: s.n.], 2009, p. 141–142.
- [8] [8] SUNG, A. H.; MUKKAMALA, S. ; XU, J. Static analyzer of vicious executables. Proceedings - Annual Computer Security Applications Conference, ACSAC, v. 0, p. 326–334, 2005.
- [9] [9] SIKORSKI, M.; HOGIN, A. Pratical Malware Analysis. 5. ed. [S.I.]: No Starch Press, 2012. 2-2 p.
- [10] [10] MATHUR, K.; HIRANWAL, S. A survey on techniques in detection and analyzing malware executable. International Journal of Advanced Research in Computer Science and Software Engineering, v. 3, p. 422–428, 2013.
- [11] [11] CHRISTODORSCU, M.; JHA, S. Static analysis of executables to detect malicious patterns. In: USENIX SECURITY SYMPOSIUM, 12., 2004. Anais... [S.I.: s.n.], 2004, p. 169–186.
- [12] [12] JONES, M. T. Arquiteturas de aprendizado profundo. Disponível em: <<https://www.ibm.com/developerworks/br/library/cc-machine-learning-deeplearning-architectures/cc-machine-learning-deep-learning-architectures-pdf.pdf>>. Acesso em: 10 mai. de 2019.
- [13] [13] BUDUMA, N.; LOCASCIO, N. Fundamentals of Deep Learning: Designing Next-Generation Machine Intelligence Algorithms. 1st. ed. [S.I.]: O'Reilly Media, Inc., 2017. ISBN 1491925612, 9781491925614.
- [14] [14] OLAH, C. Understanding LSTM Networks. Disponível em: <<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>>. Acesso em: 10 mai. de 2019.
- [15] [15] KOMPELLA, R. Using LSTMs to forecast time-series. Disponível em: <<https://towardsdatascience.com/using-lstms-to-forecast-time-series-4ab688386b1f>>. Acesso em: 10 mai. de 2019.
- [16] [16] SHARMA, S. Epoch vs Batch Size vs Iterations. Disponível em: <<https://towardsdatascience.com/epoch-vs-iterations-vs-batch-size-4dfb9c7ce9c9>>. Acesso em: 10 mai. de 2019.
- [17] [17] BROWNLEE, J. Gentle Introduction to the Adam Optimization Algorithm for Deep Learning. Disponível em: <https://machinelearningmastery.com/adam-optimizationalgorithm-for-deep-learning/>. Acesso em: 10 mai. de 2019.
- [18] [18] GOM, J. Mining English and Korean text with Python. Disponível em: <<https://www.stechstar.com/user/zbxel/AlgorithmPython/52575>>. Acesso em: 01 outubro de 2019.
- [19] [19] MASSARELLI, L.; DI LUNA, G. A.; PETRONI, F.; BALDONI, R. ; QUERZONI, L. Safe: Self-attentive function embeddings for binary similarity. In: INTERNATIONAL CONFERENCE ON DETECTION OF INTRUSIONS AND MALWARE, AND VULNERABILITY ASSESSMENT, 16., 2019. Anais... [S.I.: s.n.], 2019, p. 309–329.
- [20] [20] DI, W.; BHARDWAJ, A. ; WEI, J. Deep Learning Essentials: Your Hands-on Guide to the Fundamentals of Deep Learning and Neural Network Modeling. [S.I.]: Packt Publishing, 2018. ISBN 1785880365, 9781785880360.
- [21] [21] DEVLIN, J.; CHANG, M.; LEE, K. ; TOUTANOVA, K. BERT: pre-training of deep bidirectional transformers for language understanding. CoRR, v. abs/1810.04805, 2018. Disponível em: <<http://arxiv.org/abs/1810.04805>>. Acesso em: 3 out. de 2019.
- [22] [21] FUJINO, A.; ISOZAKI, H. ; SUZUKI, J. Multi-label text categorization with model combination based on f1-score maximization. In: PROCEEDINGS OF THE THIRD INTERNATIONAL JOINT CONFERENCE ON NATURAL LANGUAGE PROCESSING: VOLUME-II, 3., 2008. Acesso em: 3 out. de 2019
- [23] [22] BAPTISTE ROCCA. Handling imbalanced datasets in machine learning. Disponível em: <<https://towardsdatascience.com/handling-imbalanced -datasets-in-machine-learning-7a0e84220f28>>. Acesso em: 27 de janeiro de 2019.
- [24] [23] PINTO, D. R. Aprendizado Profundo Aplicado à Análise Estática de Malwares. Disponível em: <<http://www.comp.ime.eb.br/pos/arquivos/publicacoes/dissertacoes/2018/2018-Dhiego.pdf>>. Acesso em: 107 fev. de 2020.
- [25] [24] MANNING, C. D.; SCHÜTZE, H. ; RAGHAVAN, P. Introduction to information retrieval. New York, NY, USA: Cambridge University Press, 2008.
- [26] [25] KAGGLE, T. Microsoft Malware Winners' Interview: 1st place, "NO to overfitting!". Disponível em: <<http://blog.kaggle.com/2015/05/26/microsoft-malware-winners-interview-1st-place-no-to-overfitting/>>. Acesso em: 29 setembro de 2019.

Otimização da trajetória de robôs móveis diferenciais de tração usando algoritmo de otimização por enxame de partículas

Gustavo S Rodrigues^{*a}, Carlos L. M. de Souza Junior^b, Elias D. R. Lopes^b, Marcelo A. Acuña^b

^aDepartamento de Engenharia Mecânica, Universidade Estácio de Sá, Rio de Janeiro, Brasil

^bInstituto Militar de Engenharia (IME)

Praça General Tibúrcio, 80, 22290-270, Praia Vermelha, Rio de Janeiro, RJ, Brasil.

[*gustavosimao@uol.com.br](mailto:gustavosimao@uol.com.br)

RESUMO: Este artigo apresenta uma técnica de otimização não linear, a fim de otimizar a trajetória de um robô móvel diferencial não-holônomico. Os sistemas robóticos usados para realizar esse experimento consistem em um veículo com tração diferencial de duas rodas, acionado por dois motores de corrente contínua (CC) montados em um eixo comum e cada roda que pode ser conduzida independentemente para frente ou para trás, dado o movimento do robô. Este artigo também traz o modelo dinâmico do robô e o modelo do motor usado no robô. As equações cinemáticas do robô também são mostradas. O algoritmo de otimização por enxame de partículas, do inglês Particle Swarm Optimization (PSO), é aplicado para encontrar a melhor trajetória, usando o caminho mais curto e seguro entre dois pontos, evitando obstáculos fixos. O presente algoritmo é um método evolucionário que utiliza conceitos de otimização por enxames e a ideia original deste método veio da observação do comportamento social de populações de pássaros buscando um local ideal para se aninhar. O método visa equilibrar a individualidade e a sociabilidade da população para encontrar o local ótimo.

PALAVRAS-CHAVE: Cinemática. Robô de Acionamento Diferencial. Robô Não-holônomico. Otimização. PSO.

ABSTRACT: This paper presents a nonlinear optimization technique in order to optimize a trajectory of a nonholonomic differential mobile robot. The robotic systems used to perform this experiment consist in a two-wheeler differential drive vehicle, actuated by two Direct Current (DC) motors mounted on a common axis and each wheel can independently be driven forward or backward, given the robot motion. This paper also brings the dynamic model of the robot and the motor used in the robot. The kinematics equations of the robot are also shown. Particle Swarm Optimization (PSO) algorithm is applied to find the best trajectory, using the shortest and safest path between two points, avoiding fixed obstacles. The present algorithm is an evolutionary method using concepts of swarm optimization and the original idea of this method came from the observation of the social behavior of bird populations looking for an ideal nesting place. The method aims to balance the individuality and sociability of the population to find the optimal place.

KEYWORDS: Kinematics. Differential Drive Robot. Nonholonomic robot. Optimization. PSO

1. Introdução

Os robôs de acionamento diferencial são muito populares e versáteis. Eles são amplamente utilizados para fins industriais e pessoais. Com o rápido crescimento da demanda e a chegada da quarta geração de automação da indústria, o estudo de seu movimento tornou-se cada vez mais necessário. A proposta neste trabalho

consiste na tarefa simples de deslocamento ao objetivo, que exige que o robô se mova de um ponto inicial, com uma determinada pose (posição e orientação do robô) inicial, e vá para o ponto desejado com a pose desejada, evitando possíveis obstáculos anteriormente inseridos no sistema. Para tornar isso possível, um algoritmo para otimizar o caminho entre esses dois pontos é usado, um sistema de visão aérea também é

implementado para capturar a pose instantânea do robô, alimentando o algoritmo do caminho de otimização e obtendo resultados mais precisos.

2. Modelagem

Os robôs móveis com rodas constituem uma classe de sistemas mecânicos caracterizados por restrições cinemáticas que não são integráveis e, portanto, não podem ser excluídas das equações do modelo [1], o que significa que o uso da abordagem padrão para manipuladores robóticos sem restrições não é mais útil para o sistema usado neste trabalho. Como as modelagens cinemática e a dinâmica dos robôs de acionamento diferencial, que são muito específicas, não estão bem documentadas em livros, é necessário o estudo do movimento desses robôs. Os robôs de acionamento diferencial são conhecidos como um sistema sub-atuado e também possuem características não-holonômicas. Entre muitas alternativas diferentes para modelar a dinâmica de um sistema, a abordagem da dinâmica de Lagrange foi usada para modelar a dinâmica do robô neste trabalho.

2.1 Restrição Não-Holonômica

Um sistema não-holonômico está sujeito a pelo menos uma restrição não-integrável (isto é, não-holonômica) [2]. Nesse caso, o comportamento de rolagem das rodas fornece ao robô uma limitação de movimento chamada restrição não-holonômica. Essas restrições tornam o robô capaz apenas de avançar e recuar, nunca se movendo para os lados. Existem duas restrições não-holonômicas no sistema e podem ser expressas como:

- Nenhum movimento de escorregamento lateral, o que significa que o robô não pode se mover de lado;

$$\dot{x} \sin \theta - \dot{y} \cos \theta = 0 \quad (1)$$

- Restrição pura ao rolamento, o que significa que cada roda pode ter apenas um ponto de contato com o piso, como mostra a **figura 1**.

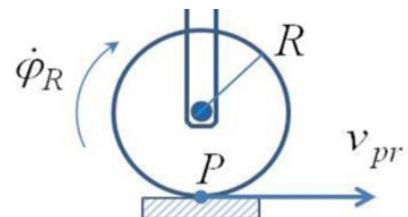


Fig. 1 – Parâmetros da roda. [3]

$$\dot{x}_R \cos \theta + \dot{y}_R \sin \theta \cos \theta = R \dot{\phi}_R \quad (2)$$

$$\dot{x}_L \cos \theta + \dot{y}_L \sin \theta \cos \theta = R \dot{\phi}_L \quad (3)$$

2.2 Modelo Cinemático

São necessários dois sistemas de coordenadas diferentes para referenciar um modelo de maneira confiável [3]. O primeiro sistema define o robô no referencial global e é chamado de sistema de coordenadas inercial e o segundo sistema define o robô em seu próprio referencial e é chamado de sistema de coordenadas do robô. Eles serão indicados como $(X_i; Y_i)$ e $(X_r; Y_r)$, respectivamente.

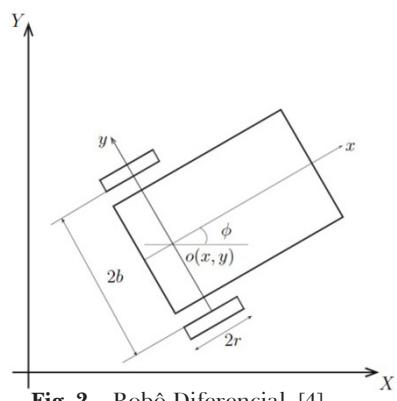


Fig. 2 – Robô Diferencial. [4]

A posição e a orientação do robô no sistema inercial estão indicadas na **equação (4)**

$$q^I = \begin{bmatrix} x_a \\ y_a \\ \theta_a \end{bmatrix} \quad (4)$$

onde (x_a, y_a) é a posição do robô (ponto 0), que é o centro das rodas motrizes na **figura 2** sobre o sistema de coordenadas global e θ é a orientação do robô no sistema de coordenadas global. R e L são a velocidade angular das rodas direita e esquerda, respectivamente.

A velocidade linear do robô e a velocidade angular são indicadas nas **equações (5) e (6)**, respectivamente.

$$\dot{x}_a^r = R \frac{(\dot{\varphi}_R + \dot{\varphi}_L)}{2} \quad (5)$$

$$\dot{\theta} = \omega = R \frac{(\dot{\varphi}_R + \dot{\varphi}_L)}{2L} \quad (6)$$

2.3 Modelo Dinâmico

A obtenção do modelo dinâmico do sistema é necessária para a resposta do sistema mais precisa. Um método muito poderoso e tradicional é usado para descrever o movimento do robô. Lagrange introduziu esse método em 1788 e, desde então, é amplamente utilizado para descrever a equação de movimento de sistemas mecânicos. O método de Lagrange usa as energias cinemáticas e potenciais do sistema para formular as equações de movimento e pode ser escrito como indicado na **equação (7)**.

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) + \frac{\partial L}{\partial q_i} = F - \Lambda^T q \lambda \quad (7)$$

Onde $L = T - V$ é a função Lagrangiana, T é a energia cinética, V é a energia potencial q_i , são as coordenadas generalizadas, F é o vetor de

força generalizada, Λ é a matriz de restrições e λ é o vetor dos multiplicadores de Lagrange.

A energia cinemática do sistema é dada pela **equação (8)**.

$$T = \frac{1}{2} m(\dot{x}_a^2 + \dot{y}_a^2) - m_c d \dot{\theta} (\dot{y}_a \cos \theta - \dot{x}_a \sin \theta) + \frac{1}{2} I_w (\dot{\varphi}_R^2 + \dot{\varphi}_L^2) + \frac{1}{2} I \dot{\theta}^2 \quad (8)$$

Como o sistema tem uma dinâmica restrita horizontal, pode-se fazer $L = T$, então a equação de movimento do sistema é dada pelo **sistema de equações (9)**.

$$\begin{aligned} m \ddot{x}_a - m d \ddot{\theta} \sin \theta - m d \dot{\theta}^2 \cos \theta &= C_1 \\ m \ddot{y}_a - m d \ddot{\theta} \cos \theta - m d \dot{\theta}^2 \sin \theta &= C_2 \\ - I \ddot{\theta} - m d \ddot{x}_a \sin \theta + m d \ddot{y}_a \cos \theta &= C_3 \\ I_w \ddot{\varphi}_R &= \tau_R + C_4 \\ I_w \ddot{\varphi}_L &= \tau_L + C_5 \end{aligned} \quad (9)$$

Onde $C_1 \dots C_5$ são os coeficientes relacionados às restrições cinemáticas. Usando o modelo cinemático, o modelo dinâmico reduzido pode ser escrito em termos das velocidades linear e angular do robô (v e ω), conforme descrito no **sistema de equação (10)**.

$$\begin{cases} \left(m + \frac{2I_w}{R^2} \right) \dot{v} - m_c d \omega^2 = \frac{1}{R} (\varphi_R + \varphi_L) \\ \left(I + \frac{2L^2}{R^2} I_w \right) \dot{\omega} - m_c d \omega v = \frac{L}{R} (\varphi_R + \varphi_L) \end{cases} \quad (10)$$

2.4 Modelo do motor de corrente contínua (CC)

A dinâmica do motor de corrente contínua (CC), amplamente difundida em muitos livros diferentes, é conhecida como:

$$\begin{cases} L \frac{di}{dt} + Ri + K_w \dot{\phi} = E \\ (I_s \ddot{\phi} - K_t i + \nu \dot{\phi} - \tau = 0) \end{cases} \quad (11)$$

Onde E é a tensão da armadura, i é a corrente da armadura, R é a resistência da armadura e L é a indutância da armadura. A inércia do eixo do motor é dada por I_s , K_t é a constante de torque do motor, K_w é a constante de tensão, ν é o coeficiente de atrito viscoso e τ é a carga dinâmica aplicada ao motor. Fazendo L igual a 0, é possível encontrar o modelo de ordem reduzida, que é descrito como:

$$I_s \ddot{\phi} + \left(\frac{K_w K_t}{R} + \nu \right) \dot{\phi} - \frac{K_t}{R} E = \tau \quad (12)$$

3. Otimização

O primeiro problema de planejamento de trajetória foi abordado por [5], onde o objetivo era encontrar o menor caminho entre um ponto inicial A até um ponto final B não levando em consideração a presença de obstáculos. Posteriormente, Murray [6] e Santilli [7] usaram a mesma abordagem de Dubins [5] para o problema de planejamento de trajetória, mas incluíram obstáculos no ambiente.

No presente trabalho, é utilizada a técnica denominada Optimização por Enxame de Partículas, do inglês *Particle Swarm Optimization* (PSO), que explora todas as soluções possíveis para o problema da obtenção de trajetória e encontra a trajetória que leva o robô de um ponto A a um ponto B através do menor caminho possível.

O algoritmo PSO é um método evolucionário proposto por Eberhart [8] que utiliza conceitos de otimização por enxames. Segundo Dulikravich [9], a ideia original deste método veio da observação do comportamento social de populações de pássaros buscando um local ideal para se aninhar.

Conforme descrito Herskovits [10], o método visa equilibrar a individualidade e a sociabilidade da população para encontrar o local ótimo. Quando a individualidade aumenta, consequentemente a procura por locais alternativos para o aninhamento cresce. Entretanto, quando a individualidade da população fica muito alta, os indivíduos podem nunca encontrar o melhor local. Por outro lado, quando a sociabilidade da população é aumentada, o indivíduo aprende mais com a experiência do vizinho e tende a não buscar novos locais para aninhar-se. Porém, se a sociabilidade fica muito alta, todos os indivíduos podem convergir para o primeiro local de mínimo encontrado, que possivelmente seja um mínimo local.

Assim, no PSO, as partículas ou indivíduos de uma população, são constantemente melhoradas graças a experiência de partículas vizinhas, mas também permitindo que cada partícula busque novas soluções que podem ser melhores que a solução das outras partículas. Neste método as partículas responsáveis pelo melhor valor da função objetivo daquela posição ($pbest$) e da história ($gbest$) são armazenadas e a cada iteração varia-se a velocidade das partículas para tender a $pbest$ e $gbest$, de forma que:

$$x_i^{k+1} = x_i^k + v_i^{k+1} \quad (13)$$

$$v_i^{k+1} = \alpha v_i^k + \beta_1 r_{1,i} (pbest - x_i^k) v_i^{k+1} + \beta_2 r_{2,i} (gbest - x_i^k) \quad (14)$$

onde x_i^k e v_i^k são, respectivamente, a posição e a velocidade da partícula na iteração .

A velocidade inicial v_i^0 é igual a zero quando k é zero, r_1 e r_2 são números aleatórios com

distribuição entre 0 e 1 e também são definidos os parâmetros α entre 0 e 1 e β entre 1 e 2.

4. Odometria

A medição da pose do robô é uma maneira muito confiável de verificar se o robô está seguindo o caminho obtido no processo de otimização. Existem muitas maneiras diferentes de obter a pose do robô. Neste artigo, uma estrutura de visão computacional de plataforma aberta, muito rápida e robusta, é usada para rastrear o robô enquanto ele está executando o caminho. A pose do robô pode ser obtida usando uma etiqueta chamada marcador fiducial, mostrada na **figura 3**. A etiqueta é fixada no robô e, com a ajuda de uma câmera com vista superior, o sistema pode obter velocidade, posição e orientação do robô.



Fig. 3 – Marcador Fiducial

O sistema chamado de *ReactIVision* consiste em um algoritmo *opensource* que fica constantemente verificando etiquetas coladas no robô. Cada etiqueta possui seu próprio número e, utilizando o ROS (*Robot Operating System*) através do protocolo TUO – protocolo de comunicação desenvolvido com a finalidade de atender os requisitos de comunicação entre interfaces tangíveis, pode-se usar o software MATLAB

para obter a pose e a velocidade do robô através da identificação da imagem, como mostrado na **figura 4**.

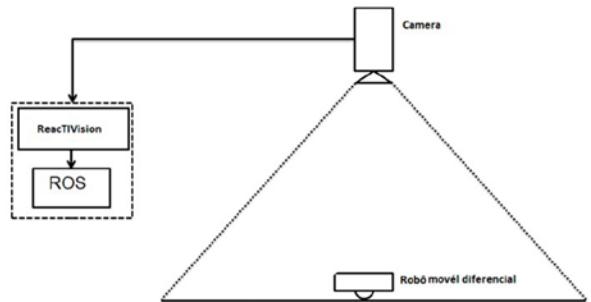


Fig. 4 – Sistema de odometria com auxílio de câmera.

5. Resultados

Como resultado parcial, a simulação alcançou o melhor caminho para o robô. A posição do objetivo foi alcançada evitando os obstáculos anteriormente inseridos na simulação, como mostra a **figura 5**.

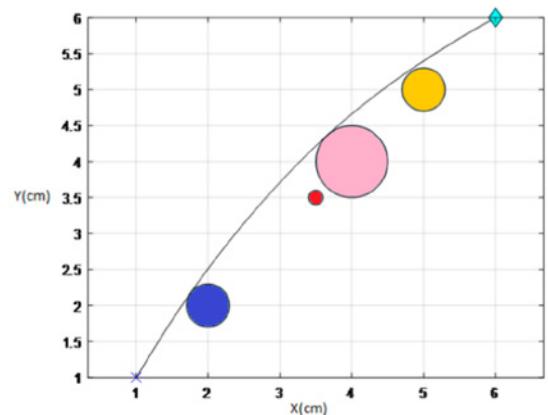


Fig. 5 – Trajetória otimizada

O X azul e o diamante são a posição inicial e a posição final, respectivamente, e os círculos azuis são os obstáculos e a linha preta é o caminho otimizado.

6. Conclusão

Pode-se concluir que o objetivo do presente trabalho foi alcançado, obtendo-se a trajetória otimizada entre dois pontos sem que houvesse colisão com os obstáculos previamente inseridos no plano.

O algoritmo PSO se mostrou bastante eficiente para o problema proposto, convergindo rapidamente para a solução, demonstrando a robustez desejada para problemas como

determinação de trajetórias ótimas de robôs de acionamento diferencial com restrições não-holonômicas.

Agradecimentos

Os autores agradecem ao Programa Pesquisa Produtividade da UNESA, ao LabRob (PUC-Rio), à Seção de Engenharia Mecânica do IME e à CAPES.

Referências bibliográficas

- [1] G. Campion, W. Chung. Wheeled robots. Springer handbook of robotics, 2008, p. 391-410.
- [2] I. Anvari. Non-holonomic Differential Drive Mobile Robot Control & Design: Critical Dynamics and Coupling Constraints. 2013. Tese de Doutorado. Arizona State University.
- [3] R. Dhaouadi, A. A. Hatab. Dynamic Modelling of Differential-Drive Mobile Robots Using Lagrange and Newton-Euler Methodologies: A Unified framework. *Advances in Robotics & Automation*, v. 2, n. 2, 2013, p. 1-7.
- [4] E. O. C. Torres. Traction Modeling and Control of a Differential Drive Mobile Robot to Avoid Wheel Slip. 2013. Tese de Doutorado. Oklahoma State University.
- [5] L. E. Dubins. On Curves of Minimal Length with a Constraint on Average Curvature, and with Prescribed Initial and Terminal Positions and Tangents. *American Journal of Mathematics*, [S.I.], v.79, n.3, 1957, p. 497–516.
- [6] J.P. Laumond, P. Jacobs, M. Taix, R. Murray. A motion planner for nonholonomic mobile robots. *Robotics and Automation, IEEE Transactions on*, [S.I.], v.10, n.5, Oct 1994, p.577–593.
- [7] A. Bicchi, G. Casalino, C. Santilli. Planning Shortest Bounded-Curvature Paths for a Class of Nonholonomic Vehicles Among Obstacles. *Journal of Intelligent and Robotic Systems*, [S.I.], v.16, n.4, 1996, p.387–405.
- [8] J. Kennedy, R. C. Eberhart. Particle swarm optimization. In *Proceedings of the 1995 IEEE International Conference on Neural Networks*, volume 4, 1995, pp. 1942–1948, Perth, Australia, IEEE Service Center, Piscataway, NJ.
- [9] M. J. Colaço, H. R. B. Orlande, G. S. Dulikravich. Inverse and Optimization Problems in Heat Transfer. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, v. 28, n. 1, 2006, p. 1-24.
- [10] M. Alba, A. G. Soares, L. R. J. Herskovits. Resolução da Cinemática Inversa de Robôs Industriais por Algoritmos de Otimização, volume 14, nº 159, XL International Sodebras Congress, Vitória, ES, março, 2019.

Rede IoT assistida por sistema de aeronaves remotamente pilotadas para apoio em operações de recuperação de desastres

Erick M. Moreira^a, Fábio Luiz Jr, Leandro Moreira^a, Jauvane C.de Oliveira^b e Paulo F. F. Rosa^a

^aInstituto Militar de Engenharia (IME)

Praça General Tibúrcio, 80, 22290-270, Praia Vermelha, Rio de Janeiro, RJ, Brasil.

^b Laboratório Nacional de Computação Científica (LNCC)

*emenezes@ime.eb.br

RESUMO: Plataformas IoT (Internet of Things) são consideradas o futuro das redes de sensores. Sua natureza embarcada facilita a adição de mobilidade e amplia drasticamente sua possibilidade de uso. Especialmente, quando usadas com veículos aéreos, essas redes podem avaliar crises rapidamente. Durante as operações de recuperação de desastres, comunicação em tempo real é essencial para coordenar as atividades. No entanto, esses cenários geralmente são privados de boa infraestrutura de comunicação. Este trabalho propõe um sistema capaz de estabelecer uma rede usando veículos aéreos, que conecta todas as partes interessadas através de uma rede de endereços IPv6. Os experimentos mostraram que um pequeno conjunto de 50 aeronaves equipadas com um dispositivo de comunicação com alcance de 50 metros é capaz de conectar até 24 elementos, espalhados por uma área de 13.000 metros quadrados. Para um conjunto de 150 aeronaves, foi possível atender a uma média de 68 agentes, em uma área de 25.600 metros quadrados.

PALAVRAS-CHAVE: Aeronave remotamente pilotada. ARP. Rede mesh.

ABSTRACT: IoT (Internet of Things) platforms are considered the future of sensor networks. Its embedded nature facilitates the addition of mobility and dramatically broadens the possibility of use for such systems. Especially when shipped on aerial vehicles, those networks can be used to assess crises quickly. During disaster recovery operations, instant communication is of the essence to coordinate activities. However, those scenarios are commonly deprived of good communication infrastructure. This work proposes a system capable of deploying an instant network mounted on UAVs, that connects all interested parts through an IPv6 mesh network. The experiments have shown that a small set of 50 UAVs equipped with a 50-meter range communication device is capable of connecting up to 24 elements spread over an area of 13,000 square meters. For a set of 150 UAVs, it was possible to serve an average of 68 agents in an area of 25,600 square meters.

KEYWORDS: Unmanned aerial vehicle. UAV. Mesh network.

1. Introdução

As aeronaves remotamente pilotadas (ARP) são uma tecnologia inestimável para muitas aplicações, como o gerenciamento de desastres [1-5], principalmente devido à sua flexibilidade e alta empregabilidade. Outras tarefas que podem ser realizadas por um conjunto de múltiplas ARPs são: inspeção [6-9], vigilância [1-2] e resposta a emergências [10-11].

Nesses contextos, a diversidade e complexidade

das demandas impossibilita o projeto de máquinas que exerçam todos esses papéis. A solução tem sido criar robôs mais simples, heterogêneos, limitados, que cooperem uns com os outros em um sistema de aeronaves remotamente pilotadas (SARP) para exceder essas limitações, demonstrando unicidade de vontade (objetivo comum) e praticando o coletivismo, a partir do autoconhecimento e da comunicação com seus vizinhos.

Normalmente, operações como essas têm um

centro de comando e controle (centro C2) e agentes de campo espalhados pela área afetada, conforme ilustrado na **figura 1**. Os agentes são responsáveis por detectar danos e reportar ao centro C2, que decide quais ações eles devem executar para mitigar o problema. A rede de comunicação é um elemento crítico desse cenário, pois é através dela que os eventos são relatados e prontamente respondidos; esse canal também possibilita o acompanhamento operacional e logístico das atividades desenvolvidas pelas equipes.



Fig. 1 – O conceito de comando e controle estabelece que o centro C2 deve coordenar as ações de todos os agentes de campo envolvidos [12]. Portanto, ele deve ter consciência situacional durante o maior tempo possível. Isto pode ser alcançado com o recebimento de relatórios dos agentes em campo e através do monitoramento do local da operação.

Um problema fundamental em áreas afetadas por desastres é a ausência ou precariedade de infraestrutura; recursos comuns como o LTE (*Long Term Evolution*, 4G) podem não estar disponíveis. Outras soluções, como caminhões de antenas móveis, postos de comando móveis e torres de celular móveis também podem ser impraticáveis devido às condições das ruas e estradas. Um SARP pode ajudar a melhorar os resultados no cenário descrito. Eles podem monitorar visualmente o local, detectando níveis de radiação, concentrações nocivas de fumaça, ou conectando agentes em solo ao centro C2. A **figura 2** apresenta uma região de desastre, provocado

por chuva ou terremotos e a solução vislumbrada para estabelecer comunicação usando múltiplos ARPs.

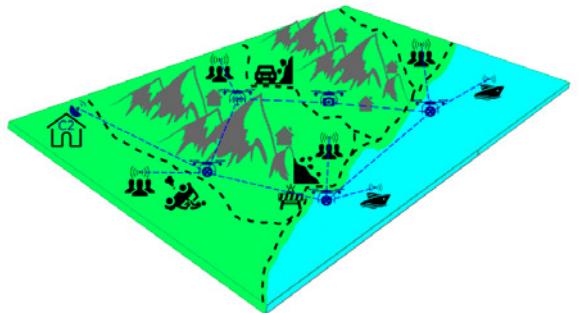


Fig. 2 – A cena mostra eventos desastrosos e equipes de resgate em campo. Devido aos bloqueios de passagem, um SARP foi empregado para estabelecer a conectividade entre os agentes e o centro C2.

Em uma solução como a descrita acima, cada ARP deve ser capaz de estabelecer uma rede de comunicação sem nenhuma infraestrutura prévia. Suas posições devem ser determinadas de acordo com algum critério, e cada veículo deve ser capaz de viajar para sua posição calculada sem colidir com os demais.

Este artigo tem como objetivo apresentar um sistema capaz de calcular essas posições, usando o menor número possível de ARPs. Em seguida, ele deve conduzir cada ARP até a sua posição, evitando colisões. E, finalmente, o sistema deve conectar todos os agentes de campo ao centro C2, através de uma rede mesh estabelecida por dispositivos IoT (Internet of Things) usando o software OpenThread [23]. As experiências mostraram que um pequeno conjunto de 50 ARPs, equipados com um dispositivo de comunicação com alcance de 50 metros, é capaz de conectar até 24 elementos, espalhados por uma área de 13.000 m². Para um conjunto de 150 ARPs com as mesmas características, foi possível

atender a uma média de 68 agentes, em uma área de 25.600 m². Já um grupo maior de 250 aeronaves, é capaz de conectar até 96 agentes em uma área de 52.900 m².

O restante do artigo está organizado da seguinte forma: a seção 2 apresenta trabalhos relacionados ao problema e a seção 3 descreve o sistema proposto; os experimentos realizados são apresentados na seção 4 e discutidos na seção 5; finalmente, a seção 6 apresenta as conclusões.

2. Trabalhos relacionados

A maioria dos trabalhos relacionados é limitada a tópicos de pesquisa, como algoritmos de roteamento [13-16] ou de posicionamento para a máxima cobertura [17-20].

É certo que uma rede assistida por SARP tem um grande potencial e ainda representa uma área muito inexplorada [21-22]. Entretanto, existem vários problemas a serem resolvidos até que uma rede desse tipo possa ser considerada estável, equilibrando alcance e largura de banda. Segundo os autores de [21-22], as redes SARP são comumente descritas como redes *ad hoc* com mobilidade, variando de acordo com o tipo de missão. Sua topologia normalmente tem um formato híbrido, mantendo uma estrutura de estrela com a estação de controle de solo (GCS, do inglês ground control station) no centro e uma estrutura de malha (*mesh*) entre as ARPs. Além disso, a topologia não tem estabilidade e pode ser encontrada em um estado estático ou em um estado altamente dinâmico, dependendo da aplicação em que é usada. Outra característica marcante dessa rede é que seus nós perdem e recuperam a conectividade continuamente, seja devido à alta mobilidade, as restrições de energia ou à presença de obstáculos entre os nós.

A abordagem de Abrajano [22] usou redes

tolerantes a atraso, com as ARPs transportando dados entre os locais de desastre e o centro C2. Embora isso seja uma abordagem interessante, essas redes não são de tempo real e seu atraso na entrega das mensagens pode interferir no processo de decisão do centro C2.

Neste trabalho buscamos uma abordagem de rede de topologia *mesh* e totalmente conectada, ou seja, nenhum nó deve se afastar além do alcance de comunicação.

3. Visão geral do sistema

O sistema proposto visa a estabelecer uma rede de comando e controle em uma área afetada por desastres para auxiliar na recuperação da região, atendendo às demandas de tráfego geradas pelas equipes de trabalho em solo. O problema a ser resolvido pode ser formalizado da seguinte forma. Dado um conjunto de demandas de tráfego D e um conjunto de ARPs disponíveis V , calcular as posições a serem ocupadas por cada veículo v , de modo que cada um esteja a uma distância máxima δ dos outros, onde a comunicação entre eles é possível. Quando todos os elementos atingirem suas respectivas posições finais, cada ARP deverá utilizar seu dispositivo IoT para criar uma rede de dados com os outros e aguardar até o final da missão. Para evitar colisões, o sistema pode ser modelado com algumas restrições, tais como planejar as trajetórias em um ambiente livre de obstáculos, e posicionar as ARPs distantes umas das outras no eixo Z (altitude).

A introdução deste sistema é dividida em três partes. A primeira seção descreve os itens essenciais da ARP e o software relacionado. Em seguida, apresentamos o algoritmo usado para calcular as posições de todas as ARPs. Por fim, mostramos o software de rede OpenThread e

como ele foi configurado.

3.1 ARP e sistema de controle

O veículo utilizado no experimento descrito neste artigo foi criado por Ricardo Maroquio Bernardo, pesquisador do Laboratório de Robótica e Inteligência Computacional do IME, é do tipo quadricóptero e foi montado com hardware comercial disponível em prateleira, como mostrado na **figura 3**. O objetivo do projeto é que a ARP seja legalmente habilitada a voar em qualquer lugar, sem restrições, de acordo com a legislação vigente; para tanto, o modelo foi desenhado para ter massa menor que 250 gramas.

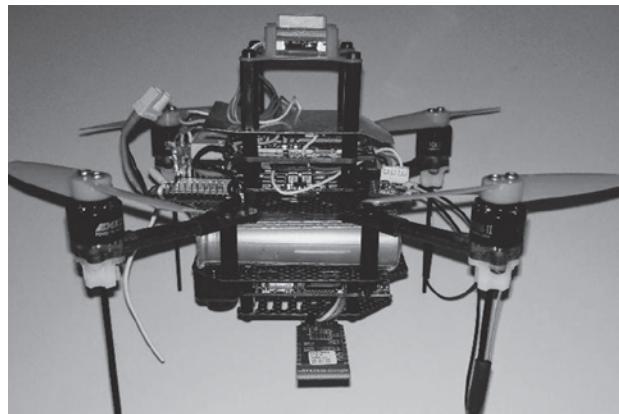


Fig. 3 – Componentes de hardware do quadricóptero: GPS Matek M8Q, controladora de voo Kakute F7 AIO, ESC Luminier 4x1, motores EMax, bateria 2S VTC6 18650 3000 mAh, computador embarcado Nano Pi Neo Air, altímetro do tipo LiDAR, sensor de fluxo ótico e dispositivo de comunicação IoT nRF52840-dongle.

O monitoramento e o controle das ARPs são feitos através de uma GCS (*Mission Planner* [24]). O *firmware* do software ArduCopter [25], na versão 3.7.0, controla a plataforma e é carregado na controladora de voo Kakute F7 AIO [26], responsável pelo sensoriamento e manobra da aeronave. Junto a essa controladora de voo, há

um Nano Pi [27], computador embarcado, que executa todos os processos não relacionados ao voo, como o software de rede e outros mecanismos de controle de supervisão distribuídos. A **figura 4** representa o funcionamento do sistema. Todos os processos Linux em execução na ARP têm acesso às variáveis de estado da aeronave, fornecidas pela controladora de voo, e podem trocar mensagens com o mundo externo através da interface de rede IPv6 provida pelo software *wpantund* [28].

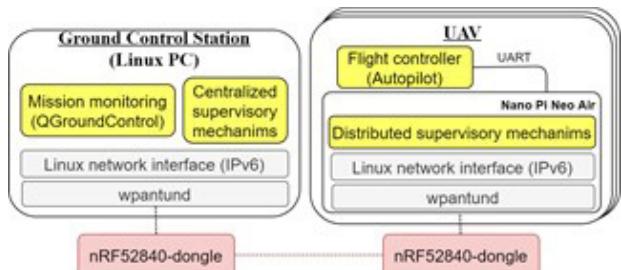


Fig. 4 – A arquitetura de software da plataforma usa o sistema operacional Ubuntu core 16.04 e a aplicação *wpantund* que une o kernel do linux ao dispositivo nRF52840-dongle. Com essa configuração, os processos Linux podem trocar mensagens IPv6/UDP com outros elementos na rede.

Além disso, esse projeto possui uma bateria 2S VTC6 18650 3000 mAh, que permite um tempo de voo de 20 a 25 minutos em boa condição climática.

3.2 Planejamento de posição

No intuito de minimizar a quantidade de veículos utilizadas, desenvolvemos um algoritmo para cálculo das posições finais das aeronaves do SARP.

O algoritmo é baseado no uso do ponto de Fermat ou ponto de Torricelli (pF). Tal ponto determina o local onde a soma das distâncias dos vértices até ele é mínima. O teorema afirma que o ponto existe no interior do triângulo, se nenhum

de seus vértices possui ângulo maior que 120° . Caso haja um vértice com ângulo igual ou maior que 120° , o pF será o próprio vértice. Portanto, teremos algumas demandas de tráfego conectadas diretamente ao centro C2 e outras conectadas por um ponto de Fermat. A primeira é chamada de subárvore direta e a segunda de subárvore de ponto de Fermat (pF). Além disso, cada subárvore tem um custo. O custo de subárvore direta é a distância entre o centro C2 e a demanda. O custo de uma subárvore de ponto Fermat é a soma de todas as distâncias que conectam o centro C2, ambas as demandas e o ponto Fermat. Os conceitos podem ser vistos na **figura 5**.

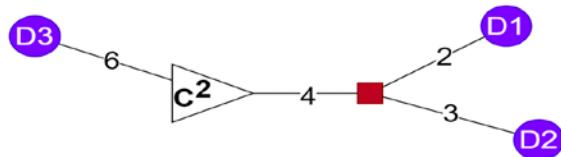


Fig. 5 – A subárvore D1-C²-D2 é do tipo ponto Fermat e têm custo 9, enquanto a subárvore C²-D3 é direta e têm custo 6.

O algoritmo 1 apresenta o planejamento de posição em pseudocódigo. Em síntese, ele consiste em utilizar o máximo de subárvore que atendam 2 demandas por meio de um pF. E, caso alguma demanda não tenha sido atendida, conectá-la diretamente ao centro C2. O resultado é uma lista de subárvore, onde cada uma atende a uma ou duas demandas. Após os cálculos, ainda é necessário percorrer todos os caminhos das subárvore selecionadas para detectar caminhos longos. Se a rota entre dois pontos for maior que a distância estipulada δ , então devemos adicionar entre eles tantas ARPs quanto forem necessários para cumprir a restrição.

Algoritmo 1: Algoritmo de cálculo posições finais

Dados: conjunto de demandas descritas por sua posição
Resultado: o conjunto de subárvore de menor custo capaz de conectar todas as demandas ao centro C2

```

1   fpTrees ← Conjunto de todas as subárvore que
2       ligam 2 demandas ao centro C2 através de um
3       ponto de Fermat, em ordem crescente de custo;
4   directTrees ← Conjunto de todas as subárvore
5       que ligam 1 demanda ao centro C2 diretamente,
6       em ordem crescente de custo;
7   selectedTrees ← [ ]; visited ← [ ]; i, j ← 1;
8   while i < fpTrees.length and j < demands.length
9       do
10      if fpTrees[i].cost ≥ demands[j].cost then
11          tree ← fpTrees[i];
12          i ← i+1;
13          demand1 ← fpTrees[i].demand1;
14          demand2 ← fpTrees[i].demand2;
15          if demand1, demand2 visited then
16              selectedTrees.add(fpTrees[i]);
17              visited.add(demand1);
18              visited.add(demand2);
19          end
20      else
21          tree ← demands[j];
22          j ← j+1;
23          demand ← directTrees[i].demand;
24          if demand visited then
25              selectedTrees.add(directTrees[i]);
26              visited.add(demand);
27          end
28      end
29  end

```

3.3. Rede IoT

Para estabelecer a rede de dados, foi usado o dispositivo nRF52840-dongle fabricado pela Nordic Semiconductor. Seu firmware implementa o software OpenThread no modo de coprocessador de rede (NCP), como mostra a **figura 6**. Como resultado, obtém-se uma interface Linux IPv6 disponível no computador complementar.

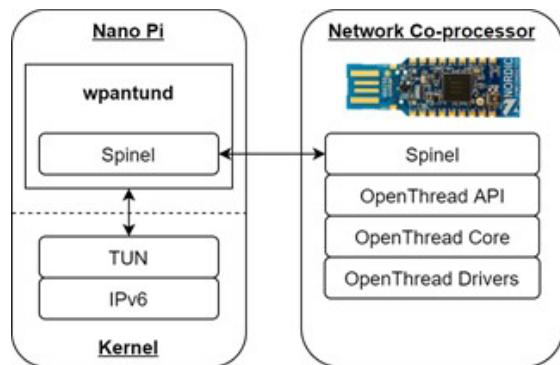


Fig. 6 – Configuração de software e hardware da rede. O computador complementar interage com outros aplicativos por meio de uma interface de rede IPv6, enquanto o wpantund controla o dispositivo NCP usando o protocolo spinel.

OpenThread (OT) é uma implementação de código-fonte aberto da especificação Thread mantida pela empresa Nest Labs. Além da rede mesh formada entre os dispositivos, outro recurso importante disponível na família de software OpenThread é o roteador de borda. Esse elemento é capaz de permitir que outras interfaces de rede entrem na rede mesh, e.g., uma interface sem fio IEEE 802.11bg. É através do roteador de borda que os agentes terrestres podem acessar o restante da rede.

4. Experimentos

As conclusões deste trabalho são derivadas da execução de três experimentos diferentes. O primeiro para determinar o alcance máximo

de comunicação entre duas ARP e, portanto, determinar o valor máximo de δ a ser usado no cálculo de posições. O segundo experimento determina a quantidade máxima de demandas que podem ser atendidas com diferentes conjuntos de ARP. E, o terceiro experimento consiste na simulação da missão completa.

No primeiro experimento, movemos progressivamente a ARP para longe da posição inicial, enquanto monitoramos sua localização e o indicador de intensidade do sinal recebido (RSSI). A **figura 7** é um registro do experimento.



Fig. 7 – O primeiro experimento analisou o alcance da comunicação com o dongle nRF52840 e a taxa de dados disponível.

O objetivo do segundo experimento é determinar quantas demandas podem ser atendidas. Um fator essencial a se considerar é o número de veículos disponíveis e o tamanho da área a ser coberta. Os parâmetros utilizados foram inspirados por [29], que projetou os três cenários descritos na **tabela 1**.

Tab. 1 – Cenários operacionais [29].

Requisito	Cenário 1	Cenário 2	Cenário 3
Área de emprego (m ²)	13.225	25.600	52.900
Número de ARP disponíveis	50	150	250

Ele propõe três cenários que consideram operações de enxame pequenas, médias e grandes, definindo para cada um a quantidade de ARP disponíveis e o tamanho das áreas de operação. Nesta fase, foi utilizado o Matlab para determinar a relação entre o número de demandas

a serem atendidas e a quantidade de aeronaves necessárias, usando o algoritmo 1.

O terceiro experimento foi projetado para ser extremamente próximo ao voo real, permitindo avaliar detalhes de um voo com vários robôs. Por esse motivo, usamos a estação de controle de solo (GCS) desenvolvida por nosso grupo de pesquisa e validada para o controle de voos reais com múltiplas aeronaves [30-31].

5. Resultados

O primeiro experimento mostrou que o dispositivo nRF52840-dongle, com antena embutida no circuito, pode atingir até 80 metros com comunicação bidirecional ativa. A **figura 8** mostra um gráfico dos registros do experimento. No eixo horizontal listamos a distância, em metros, e no eixo vertical o nível de sinal RSSI recebido, do inglês Received Signal Strength Indicator. Conforme especificação, o dispositivo possui sensibilidade de recepção máxima de -90 dBm, ou seja, níveis de sinal menores que este indicam que a mensagem não poderá ser recebida corretamente. O limiar é atingido entre 60 e 70 metros, onde há uma queda mais significativa na qualidade da comunicação. Nesse intervalo, aproximamos e afastamos o veículo algumas vezes; por isso, a **figura 8** mostra um gráfico mais caótico nesse intervalo, consequência desses movimentos.

Como toda especificação de rede, a Thread disponibiliza protocolos para resolver problemas comumente apresentados, tais como o protocolo IPv6 para resolver o endereçamento de nós; o protocolo UDP (*User Datagram Protocol*), otimizado para realizar transporte de dados fim-a-fim, aproveitando o cabeçalho de outros protocolos de camadas inferiores; e o protocolo CoAP (*Constrained Application Protocol*), para estabelecer os comandos disponíveis e estruturar

as mensagens no nível de aplicação, em redes com restrição de banda e potência.

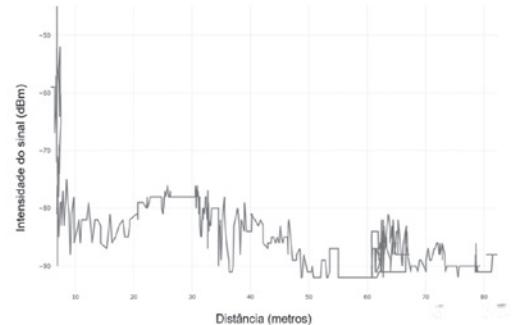


Fig. 8 – Gráfico da intensidade do sinal recebido em função da distância entre a ARP e a GCS. No eixo horizontal, temos a distância e, no eixo vertical, o nível do sinal recebido.

No entanto, todos esses recursos e seus cabeçalhos de controle possuem um custo, e normalmente esse custo interfere na taxa de transferência de dados do enlace. Para determinar a taxa máxima de transmissão utilizamos o software *iPerf3* [32], que é uma ferramenta de medição para redes. Com o auxílio da ferramenta, executamos uma análise de sessenta segundos e outra de cento e vinte segundos. A análise ocorre de forma combinada entre os dois nós: durante o período de execução, um dos nós inunda o outro com pacotes e, baseado nas perdas, o software calcula a taxa máxima de dados transmitida. Na **figura 9**, podemos ver a saída no console do comando emitido. Os parâmetros passados no comando solicitam que um dos nós inunde o outro com mensagens do tipo UDP, com uma banda máxima de 300 kbps, durante um período de 60 segundos; além disso, ele solicita que os relatórios sejam impressos a cada 6 segundos e formatados em kilobits por segundo.

Durante os testes, capturamos os pacotes com o software wireshark [33] de duas maneiras diferentes. Uma das capturas foi feita no nível da interface de software, isto é, na interface de rede do sistema operacional; a captura neste ponto nos indica o

tamanho da banda útil para o uso de aplicações.

```
menezes@gcs:~/openthread$ iperf3 --format k -6 --udp -b 300k -i 6 -t 60 --reverse --client m4
Connecting to host m4, port 5201
Reverse mode, remote host m4 is sending
[ 5] local fd11:1111:1122:0:8f5b:b4cc:e6a:6ac2 port 51170 connected to fd11:1111:1122:0:fd2d
[ ID] Interval Transfer Bitrate Jitter Lost/Total Datagrams
[ 5] 0.00-6.00 sec 74.3 KBytes 101 Kbits/sec 358103.463 ms 87/150 (58%)
[ 5] 6.00-12.00 sec 72.0 KBytes 98.2 Kbits/sec 6994.475 ms 125/186 (67%)
[ 5] 12.00-18.00 sec 74.3 KBytes 101 Kbits/sec 127.237 ms 124/187 (66%)
[ 5] 18.00-24.00 sec 72.0 KBytes 98.2 Kbits/sec 10.962 ms 124/185 (67%)
[ 5] 24.00-30.00 sec 73.1 KBytes 99.9 Kbits/sec 12.901 ms 126/188 (67%)
[ 5] 30.00-36.00 sec 72.0 KBytes 98.3 Kbits/sec 5.543 ms 124/185 (67%)
[ 5] 36.00-42.00 sec 73.1 KBytes 99.9 Kbits/sec 6.587 ms 125/187 (67%)
[ 5] 42.00-48.00 sec 72.0 KBytes 98.3 Kbits/sec 6.864 ms 123/184 (67%)
[ 5] 48.00-54.00 sec 72.0 KBytes 98.2 Kbits/sec 12.517 ms 125/186 (67%)
[ 5] 54.00-60.00 sec 73.1 KBytes 99.9 Kbits/sec 10.746 ms 127/189 (67%)
[ ID] Interval Transfer Bitrate Jitter Lost/Total Datagrams
[ 5] 0.00-60.24 sec 2.15 MBytes 300 Kbits/sec 0.000 ms 0/1870 (0%) sender
[ 5] 0.00-60.00 sec 728 KBytes 99.4 Kbits/sec 10.746 ms 1210/1827 (66%) receiver
```

Fig. 9 – Resultado da execução do software iPerf3.

A outra captura foi feita com um dispositivo nRF52840, carregado com um software de captura de pacotes 802.15.4; a captura neste ponto nos informa a capacidade máxima de transferências de dados na interface aérea. Na figura 10, apresentamos o resultado de ambas as capturas: no primeiro gráfico, a captura na interface aérea, com o valor de 195 kbps; e no segundo gráfico, a captura na interface de rede, com o valor de 100 kbps. Portanto, verificamos que o impacto da pilha de protocolos Thread na capacidade máxima de transmissão do protocolo 802.15.4 é de aproximadamente 51,28%.

A fim de estimar o valor médio de demandas que podem ser atendidas em cada cenário e o tempo médio de processamento, realizamos o segundo experimento. Nele simulamos o cálculo de posições das ARPs para estabelecer a rede de comunicação. Com base nos resultados do primeiro teste, configuramos o valor da distância máxima de enlace entre as aeronaves (δ) para 50 metros. Executamos simulações com demandas aleatoriamente posicionadas, variando a quantidade de demandas de 4 a 100, com passo incremental de 2 demandas.

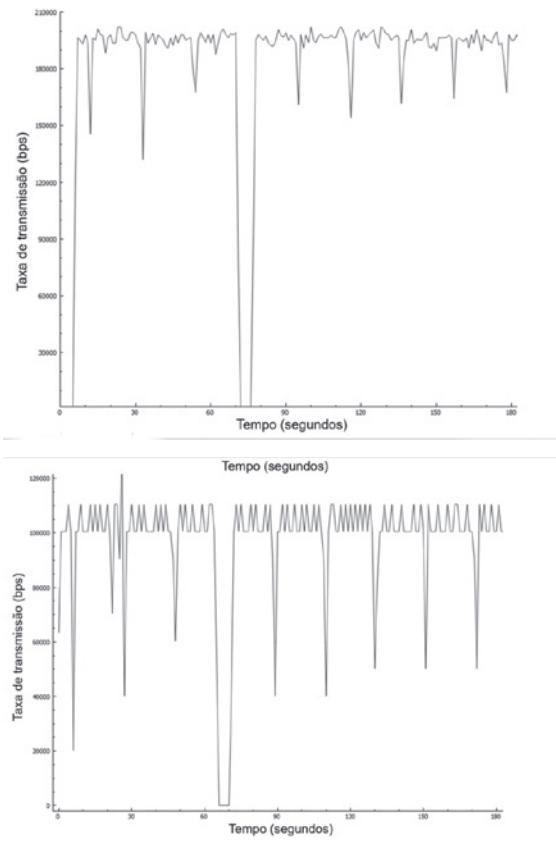


Fig. 10 – Teste usando a ferramenta iPerf3. Acima temos a taxa de dados transferida no nível de pacotes do protocolo 802.15.4 e, abaixo, a taxa de transferência na interface do sistema operacional.

Para cada quantidade de demandas, executamos 1000 simulações a fim de determinar a quantidade média de veículos necessários e o tempo de médio de execução do algoritmo. O experimento demonstrou que o cálculo das posições poderia ser realizado em tempo real e definiu o número médio de demandas, que podem ser atendidas com diferentes conjuntos de ARPs. A **figura 11** representa graficamente os resultados de quantas demandas podem ser atendidas (gráfico superior) e o tempo de execução do algoritmo (gráfico inferior).

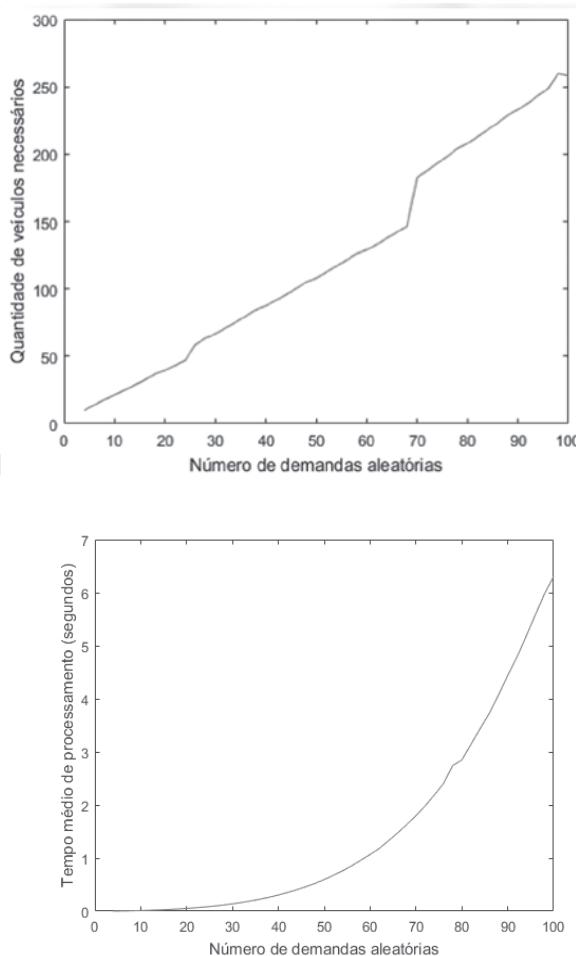


Fig. 11 – Cálculo de posição usando pontos de Fermat: O gráfico superior lista o número de veículos necessários (eixo

vertical) para atender a um determinado número de agentes terrestres (eixo horizontal). Observa-se o aumento linear da capacidade de serviço do sistema junto com o aumento do número de veículos disponíveis. O gráfico inferior mostra o tempo médio em segundos (eixo vertical) necessário para executar o cálculo da posição, considerando o número de agentes terrestres a serem atendidos (eixo horizontal).

Após a análise dos resultados do segundo experimento, extraímos os requisitos operacionais, resumidos na **tabela 2**. Conforme indicado, em média, para um conjunto de 50 ARPs, podemos atender até 24 demandas; para um conjunto de 150 ARPs, podemos atender até 68 demandas; e para um conjunto de 250 ARPs, podemos atender até 96 demandas. Apesar da curva de crescimento, vale ressaltar que o tempo máximo de processamento não ultrapassou 7 segundos. Tal resultado é consequência da simplicidade do algoritmo desenvolvido e, principalmente, da baixa quantidade de veículos disponíveis para uso nos cenários operacionais apresentados. O experimento também apurou que a economia de veículos necessários varia de 26,30% a 38,82%, em relação ao método de planejamento direto de trajetórias; onde cada demanda é atendido por uma subárvore direta de aeronaves.

Tab. 2 – Estimativas médias de planejamento de trajetória com ponto de Fermat.

Resultados	Cenário 1	Cenário 2	Cenário 3
Número médio de demandas atendidas	24	68	96
Duração média da execução do algoritmo (segundos)	0,0787	1,6326	5,5796
Economia de veículo em relação ao planejamento direto (%)	26,3014	33,7374	38,8220

No terceiro experimento, o algoritmo foi

integrado na GCS desenvolvida por Silva [34] no intuito de viabilizar o ensaio em voo do sistema completo. De posse dessa implementação, pudemos apresentar uma prova de conceito inicial e reutilizar um componente já testado em outros trabalhos [30-31, 34]. Na **figura 12**, podemos observar um centro C2, três demandas dadas como entrada para o algoritmo de cálculo de posições e onze ARPs. Note que há uma árvore de ARPs conectando todas as demandas à sua raiz, o centro C2.

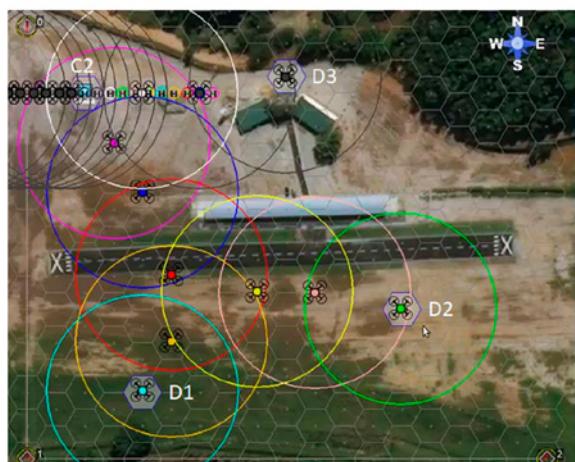


Fig. 12 - A simulação mostra três demandas conectadas ao centro C2 através de uma rede formada por 11 ARPs.

6. Conclusão

Veículos aéreos não tripulados têm um potencial considerável para alavancar operações de recuperação de desastres. Além da aplicação descrita neste artigo, as ARPs podem ajudar a monitorar as atividades das equipes em solo, avaliar as áreas afetadas e capturar dados para reconstrução 3D das áreas de risco.

Este artigo descreveu uma solução para empregar um conjunto de ARPs para formar uma rede sobre uma área afetada por desastres, conectando todos os agentes em solo ao centro de comando e controle. Os resultados dos experimentos mostram que essa rede é viável com o dispositivos do tipo Thread, apesar da limitação de alcance já esperada do dispositivo nRF52840-dongle.

Outra consideração interessante do sistema é o mecanismo de continuidade. Atualmente, usamos o sistema de substituição de ARP para tarefas contínuas [35].

Agradecimentos

Os autores agradecem a Ricardo Maroquio Bernardo por compartilhar seu projeto de quadricóptero e por ajudar na montagem das ARPs utilizadas nos experimentos.

Referências bibliográficas

- [1] Saska, M.; Chudoba, J.; Precil, L.; Thomas, J.; Loianno, G.; Tresnak, A.; Vonasek, V.; Kumar, V. Autonomous deployment of swarms of micro-aerial vehicles in cooperative surveillance. 2014 International Conference on Unmanned Aircraft Systems, ICUAS 2014 - Conference Proceedings, v. 1, p. 584–595, 2014.
- [2] Šulák, V.; Kotuliak, I.; Čižák, P. Search using a swarm of unmanned aerial vehicles. 2017 15th International Conference on Emerging eLearning Technologies and Applications (ICETA), v. 1, p. 1–6, Oct 2017.
- [3] Kato, A.; Wakabayashi, H.; Hayakawa, Y.; Bradford, M.; Watanabe, M.; Yamaguchi, Y. Tropical forest disaster monitoring with multi-scale sensors from terrestrial laser, uav, to satellite radar. In: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2017. p. 2883–2886.

- [4] Hartman, J.; Barzilov, A.; Novikov, I. Remote sensing of neutron and gamma radiation using aerial unmanned autonomous system. 2015 IEEE Nuclear Science Symposium and Medical Imaging Conference, NSS/MIC 2015, v. 1, p. 1–4, 2016.
- [5] Micconi, G.; Aleotti, J.; Caselli, S.; Benassi, G.; Zambelli, N.; Zappettini, A. Haptic guided uav for detection of radiation sources in outdoor environments. In: 2015 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED-UAS). Mexico: IEEE, 2015. p. 265–271.
- [6] Yu, H.; Yang, W.; Zhang, H.; He, W. A uav-based crack inspection system for concrete bridge monitoring. In: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2017. p. 3305–3308.
- [7] Ikeda, T.; Yasui, S.; Fujihara, M.; Ohara, K.; Ashizawa, S.; Ichikawa, A.; Okino, A.; Oomichi, T.; Fukuda, T. Wall contact by octo-rotor uav with one dof manipulator for bridge inspection. In: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2017. p. 5122–5127.
- [8] Li, Z.; Mu, S.; Li, J.; Wang, W.; Liu, Y. Transmission line intelligent inspection central control and mass data processing system and application based on uav. In: 2016 4th International Conference on Applied Robotics for the Power Industry (CARPI). 2016. p. 1–5.
- [9] Shukla, A.; Karki, H. Teleoperation of a uav by a 6-dof parallel manipulator with hybrid mapping for inspection of industrial structures. In: 2017 17th International Conference on Control, Automation and Systems (ICCAS). 2017. p. 883–888.
- [10] Kumar, S. A.; Vanualailai, J. A Lagrangian UAV swarm formation suitable for monitoring exclusive economic zone and for search and rescue. In: 2017 IEEE Conference on Control Technology and Applications (CCTA). [S.I.]: IEEE, 2017. p. 1874–1879. ISBN 978-1-5090-2182-6.
- [11] Paula, P. de S.; Castro, M. F. de; Paillard, G. A. L.; Sarmento, W. W. F. A swarm solution for a cooperative and self-organized team of uavs to search targets. 2016 8th Euro American Conference on Telematics and Information Systems (EATIS), v. 1, p. 1–8, April 2016.
- [12] Ministério da Defesa, EMCFA. Doutrina para o sistema militar de comando e controle, 3a Edição. 2015. Disponível em: <www.defesa.gov.br>. Acesso em: 26 de fevereiro de 2018.
- [13] Dong, S. Y. Optimization of olsr routing protocol in uav ad hoc network. In: 2016 13th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWA-MTIP). 2016. p. 90–94.
- [14] Hussen, H. R.; Choi, S.; Park, J.; Kim, J. Performance analysis of manet routing protocols for uav communications. In: 2018 10th International Conference on Ubiquitous and Future Networks (ICUFN). 2018. p. 70–72. ISSN 2165-8536.
- [15] Mahmud, I.; Cho, Y. Adaptive hello interval in fanet routing protocols for green uavs. IEEE Access, v. 7, p. 63004–63015, 2019. ISSN 2169-3536.
- [16] Zhang, Q.; Jiang, M.; Feng, Z.; Li, W.; Zhang, W.; Pan, M. Iot enabled uav: Network architecture and routing algorithm. IEEE Internet of Things Journal, v. 6, n. 2, p.3727–3742, April 2019. ISSN 2327-4662.
- [17] Ruan, L.; Wang, J.; Chen, J.; Xu, Y.; Yang, Y.; Jiang, H.; Zhang, Y.; Xu, Y. Energy-efficient multi-uav coverage deployment in uav networks: A game-theoretic framework. China Communications, v. 15, n. 10, p. 194–209, Oct 2018. ISSN 1673-5447.
- [18] Li, X.; Xu, X. Uav comprehensive coverage to users in urban environment. In: 2018 9th International Conference on Software Engineering and Service Science (ICSESS). 2018. p. 503–506. ISSN 2327-0594.
- [19] Urosevic, U.; Veljovic, Z.; Pejanovic-Djurisic, M. Improving cellular coverage through uavs. In: 2017 20th International Symposium on Wireless Personal Multimedia Communications (WPMC). 2017. p. 68–73. ISSN 1882-5621.

- [20] Li, B.; Chen, C.; Zhang, R.; Jiang, H.; Guo, X. The energy-efficient uav-based bs coverage in air-to-ground communications. In: 2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM). 2018. p. 578–581. ISSN 2151-870X.
- [21] Gupta, L.; Jain, R.; Vaszkun, G. Survey of important issues in uav communication networks. IEEE Communications Surveys Tutorials, v. 18, n. 2, p. 1123–1152, Second quarter 2016. ISSN 1553-877X.
- [22] Abrajano, G.; Favila, C.; Luo, C. Y.; Trono, E.; Lagazo, D.; Sevilla, B.; Honrado, J.; Solpico, D.; Yu, J.; Chua, K.; Mamaradlo, J.; Jose, C.; Yao, C. J.; dela Cruz, J.; Ancheta, E.; Domingo, A.; Ong, J.; Dautuin, J.; Yasumoto, K.; Libatique, N.; Tangonan, G. Demonstrations of post-disaster resilient communications and decision-support platform with uavs, ground teams and vehicles using delay-tolerant information networks on sub-ghz frequencies. In: 2017 IEEE Global Humanitarian Technology Conference (GHTC). 2017. p. 1–8.
- [23] Nest Labs. OpenThread. 2019. Disponível em: <<https://openthread.io/>>. Acesso em: 19 de janeiro de 2019.
- [24] ArduPilot Team. Mission planner. 2019. Disponível em: <<https://ardupilot.org/planner/>>. Acesso em: 19 de janeiro de 2019.
- [25] ArduPilot Team. Firmware arducopter. 2019. Disponível em: <<https://ardupilot.org/copter/>>. Acesso em: 19 de janeiro de 2019.
- [26] ArduPilot Team. Controladora de voo kakute F7 AIO. 2019. Disponível em: <<https://ardupilot.org/copter/docs/common-holybro-kakutef7aio.html>>. Acesso em: 19 de janeiro de 2019.
- [27] Friendly ARM. Computador Nano Pi Neo Air. 2019. Disponível em: <http://wiki.friendlyarm.com/wiki/index.php/NanoPi_NEO_Air>. Acesso em: 19 de janeiro de 2019.
- [28] Wpantund Team. wpantund software. 2019. Disponível em: <<https://github.com/openthread/wpan-tund>>. Acesso em: 22 de maio de 2019
- [29] DARPA. OFFensive Swarm Enabled Tactics (OFFSET). 2017. Disponível em: <<https://www.fbo.gov/utils/view?id=015bda98560d82d7568501376efaebdf>>. Acesso em: 02 de dezembro de 2017.
- [30] Silva, L. C. B. da; Bernardo, R. M.; Oliveira, H. A. de; Rosa, P. F. F. Multi-uav agent-based coordination for persistent surveillance with dynamic priorities. In: 2017 International Conference on Military Technologies (ICMT), 2017. p. 765–771.
- [31] Silva, L. C. B. da; Bernardo, R. M.; Oliveira, H. A. de; Rosa, P. F. F. Unmanned aircraft system coordination for persistent surveillance with different priorities. In: 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE), 2017. p. 1153–1158. ISSN 2163-5145
- [32] ESnet - Lawrence Berkeley National Laboratory. iperf3. 2019. Disponível em: <<https://github.com/esnet/iperf>>. Acesso em: 22 de maio de 2019.
- [33] Combs, G. Wireshark. 2019. Disponível em: <<https://www.wireshark.org/>>. Acesso em: 22 de maio de 2019.
- [34] Silva, L. C. B. Coordenação de multi-agentes robóticos para controle de VANT sem monitoramento persistente por cobertura de área com prioridades dinâmicas. Tese de doutorado, Instituto Militar de Engenharia, Rio de Janeiro, RJ, Brasil, 2018.
- [35] Oliveira, N.; Moreira, E. M.; Rosa, P. F. F. Particle swarm optimization algorithm implementation for multiple drones control in continuous task simulation. In: 16th IEEE Latin American Robotics Symposium (SBR-LARS 2019). Rio Grande, Brazil: IEEE, 2019.

A Influência do Teor de Austenita Revertida no Comportamento Mecânico do Aço Maraging 250

Daniel Fonseca da Cunha*, **Alaelson Viera Gomes**, **Achille Arantes Bassi**, **Luiz Paulo de Mendonça Brandão**
Instituto Militar de Engenharia (IME)

Praça General Tibúrcio, 80, 22290-270, Praia Vermelha, Rio de Janeiro, RJ, Brasil.

*danielfonsecacunha@gmail.com

RESUMO: *Aços maraging possuem limite de resistência e tenacidade à fratura superiores aos aços convencionais de alta resistência, sendo classificados como aços de ultra-alta-resistência. Tais propriedades advém da composição desses aços aliada ao processamento termomecânico aos quais são submetidos. Entretanto, durante o tratamento térmico de envelhecimento, a ocorrência de uma fase austenítica conhecida como austenita revertida provoca grande impacto nas propriedades mecânicas do material.*

A fim de se determinar a influência desta porção de fase austenítica no comportamento mecânico do aço maraging 250, buscou-se analisar neste trabalho o aumento desta fração em função da temperatura de envelhecimento e o decréscimo de dureza atrelada a esse aumento.

Para tanto, realizou-se uma análise por Difração de Raios X (DRX) a fim de se determinar a porcentagem das fases presentes no material e ensaios de micro dureza e tração, de acordo com as normas ASTM E384 e ASTM E8/E8M.

Por fim, conclui-se que a presença da fase martensítica, presente na liga, é maior do que a prevista para o sistema Fe-Ni, com apreciável decréscimo das propriedades analisadas.

PALAVRAS-CHAVE: *Aço. Maraging 250. Caracterização.*

ABSTRACT: *Maraging steels have higher strength and fracture toughness limits than conventional high strength steels and are classified as ultra-high strength steels. Such properties come from the composition of these steels allied to the thermomechanical processing to which they are submitted. However, during aging heat treatment, the occurrence of an austenitic phase known as reversed austenite has a major impact on the mechanical properties of the material.*

In order to determine the influence of this austenitic phase portion on the mechanical behavior of maraging 250 steel, this work aimed to analyze the increase of this fraction as a function of aging temperature and the decrease of hardness and mechanical strength associated with this increase. For this purpose, an X-Ray Diffraction analysis (DRX) was performed to determine the percentage of phases present in the material and micro hardness and tensile tests according to ASTM E384 e ASTM E8/ E8M standards.

Finally, it can be concluded that the presence of martensitic phase present in the alloy is higher than expected for Fe-Ni system, with appreciable decrease of the properties analyzed.

KEYWORDS: *Steel. Maraging 250. Characterization*

1. Introdução

A principal característica dos aços Maraging 18%Ni é apresentar boas propriedades de tração e de tenacidade para vários níveis de resistência, possuindo um vasto campo de aplicações. Essa característica única é obtida através do endurecimento por precipitação de compostos intermetálicos em uma matriz martensítica tenaz, após um simples tratamento

térmico de envelhecimento. Dentre os compostos intermetálicos formados estão: Ni_3Mo , Ni_3Ti , Fe_2Mo .

Primeiramente o aço é submetido a uma etapa de solubilização ou austenitização por volta de 820 °C com resfriamento ao ar e posteriormente a uma etapa de envelhecimento por volta de 480 °C [1]. Durante tais tratamentos ocorrem transformações de fases que são determinantes para o comportamento mecânico destes aços.

Tais transformações podem ser analisadas a partir do diagrama de fases metaestável do sistema Fe-Ni (**figura 1**), onde podem ser observadas a transformação da austenita para a martensita durante o resfriamento e a transformação reversa da martensita para a austenita no aquecimento.

A martensita formada durante o resfriamento nos aços maraging possui estrutura cúbica de corpo centrado (CCC), diferentemente da martensita formada no sistema Fe-C que possui estrutura tetragonal de corpo centrado (TCC).

Esta transformação de fase ocorre por cisalhamento da rede cristalina e não por difusão (adifusional) [2]. Como consequência disso, a martensita gerada possui densidade de discordâncias superior à austenita original, o que aumenta o número de sítios preferenciais para a nucleação de precipitados intermetálicos durante o envelhecimento [3].

Com relação à reversão da martensita em austenita, a mesma ocorre durante o aquecimento, onde a temperatura de início da formação da austenita no aquecimento (A_s) e a temperatura de fim da formação da austenita no aquecimento (A_f) são bem superiores às M_s e M_f (temperaturas de início e fim da transformação martensítica, respectivamente), caracterizando uma histerese no ciclo térmico [4].

A austenita formada nesta transformação é denominada de austenita revertida e seu aparecimento está ligado a exposição a altas temperaturas e tempos prolongados de tratamento térmico de envelhecimento, conhecido como superenvelhecimento. A taxa em que esta reação de reversão ocorre é fortemente dependente da composição do aço maraging e está intimamente ligada a diminuição das propriedades mecânicas destes aços [5].

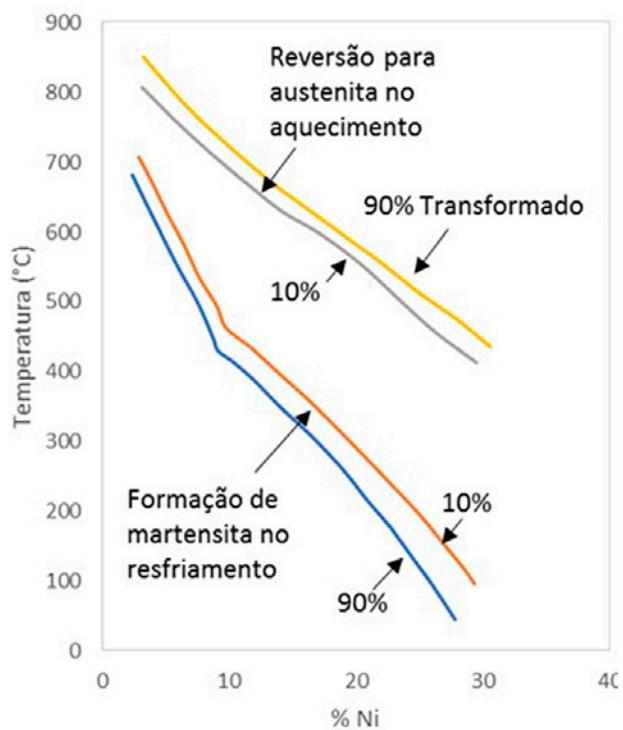


Fig. 1 – Diagrama de transformação metaestável no sistema Fe-Ni. Fonte: SILVA; MEI, 2010.

A cinética de formação da austenita revertida é influenciada pelos teores de níquel, cobalto, molibdênio e titânio. Maiores teores de níquel e molibdênio aceleram o processo de reversão, enquanto maiores teores de cobalto e de titânio retardam esta formação [6] e [7].

As propriedades mecânicas dos aços maraging são determinadas a partir dos tratamentos termomecânicos aos quais são submetidos. A **tabela 1** apresenta valores típicos de propriedades mecânicas do aço maraging 250 quando solubilizados a 820°C por 1 hora e envelhecidos a 480°C por 3 horas [8].

Tab. 1 – Propriedades mecânicas típicas do aço maraging 250 quando solubilizado e envelhecido.

Propriedades	Solubilizado	Solubilizado e Envelhecido
Limite de Escoamento (MPa)	799	1660
Limite de Resistência (MPa)	1010	1690
Deformação total (%)	19	11
Redução de Área (%)	72	50
Dureza (HRC)	29	49

Fonte: SILVA; MEI, 2010

De acordo com Schmidt e Rohrbach [2], quanto maior for o limite de resistência menor será a tenacidade a fratura do material, entretanto, estas propriedades são superiores para os aços maraging da classe 18% Ni.

Com relação à dureza, tanto os testes conduzidos na International Nickel Company [1], quanto as análises de Pardal et al [9] apontam para um expressivo ganho de dureza do material quando envelhecido em comparação com somente solubilizado.

Entretanto, o estudo de Pardal et al [9] aponta para um decréscimo de dureza de aços maraging classe 300 quando o tratamento térmico é conduzido a temperaturas mais elevadas que 480 °C e por um longo período de tempo.

2. Materiais e métodos

A fim de se determinar o aumento da fração de austenita revertida em função da temperatura de envelhecimento, realizou-se este processo a temperaturas de 480°C, 510°C, 530°C, 550°C e 570°C por um período de 3 horas com posterior resfriamento a ar.

O material analisado encontra-se de acordo com as normas AMS 6512-H [10] e MIL-S-46850-D [11] e foi fornecido pela empresa Villares Metals S.A. Foi produzido em fornos a vácuo pelos processos VIM

e VAR e com composição das amostras, segundo certificado de qualidade fornecido pela referida empresa, detalhada na **tabela 2**.

Tab. 2 – Composição química do aço maraging 250 fornecido pela Villares Metals S.A.

Elemento	Percentual em Peso (%)	Elemento	Percentual em Peso (%)
C	0,04	Ti	0,47
Si	0,05	Al	0,107
Mn	< 0,01	B	0,018
P	< 0,05	Zr	0,012
S	0,011	Mg	< 0,005
Co	8,49	Ca	< 0,0010
Cr	0,08	N	< 0,0010
Mo	5,2	O	< 0,0010
Ni	18,23	Fe	Rest
Cu	0,02		

A análise do material se deu através da técnica de Difração de Raios X (DRX), que tem por finalidade determinar as fases presentes no material de acordo com a temperatura de envelhecimento a qual foram submetidos, onde se procurou identificar o aparecimento e a evolução da austenita revertida presente nas amostras atrelado ao aumento de temperatura. Além da microdureza do material, através do ensaio de microdureza Vickers e do ensaio de tração, foi investigada a mudança de propriedades mecânicas em função do tratamento térmico submetido.

3. Resultados e Discussões

Tab. 3 – Porção das fases martensíticas e austeníticas presente nas amostras.

Temperatura de Envelhecimento (°C)	Como Recebido	480	510	530	550	570
Martensita (%)	100	96,43	91,07	83,43	76,5	61,38
Austenita Revertida (%)	0	3,57	8,93	16,57	23,5	38,62

As **figuras 2 a 7** representam o resultado da análise de DRX com os respectivos planos cristalográficos identificados.

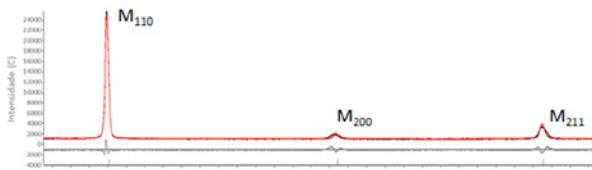


Fig. 2 – Difratograma do material como recebido.

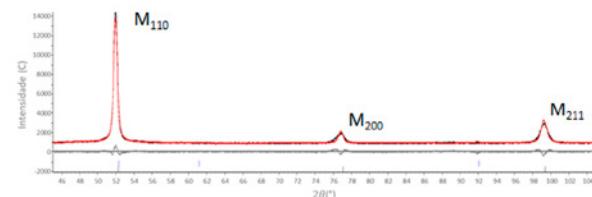


Fig. 3 – Difratograma do material envelhecido

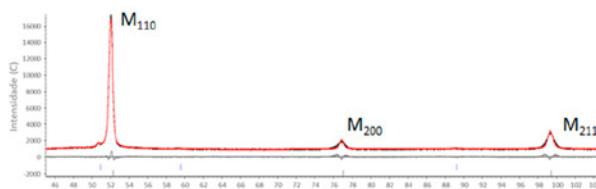


Fig. 4 – Difratograma do material envelhecido a 510°C.

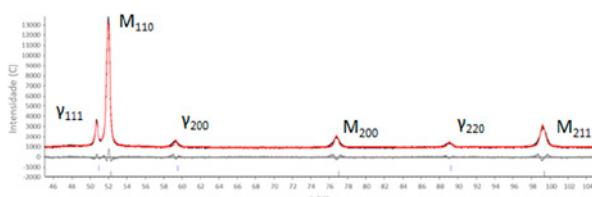


Fig. 5 – Difratograma do material envelhecido a 530°C.

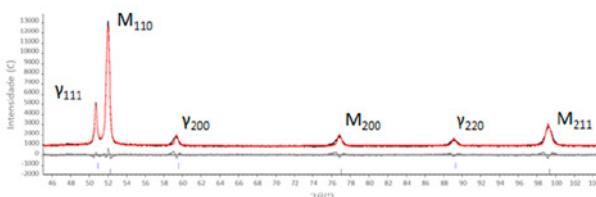


Fig. 6 – Difratograma do material envelhecido a 530°C.

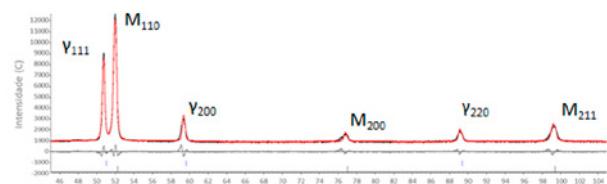


Fig. 7 – Difratograma do material envelhecido a 570°C.

Fica evidente o aumento da fase austenítica com o aumento da temperatura de tratamento, através da identificação dos planos (111), (200) e (220).

Fazendo-se uso do diagrama de transformação metaestável do sistema Fe-Ni (**figura 1**), fica evidente que os resultados obtidos se encontram fora da faixa de reversão austenítica, como observado na **figura 8**. Esse fato pode ser explicado através do efeito dos demais elementos de liga e do processo de tratamento térmico empregado, o qual não se utilizou forno com atmosfera controlada.

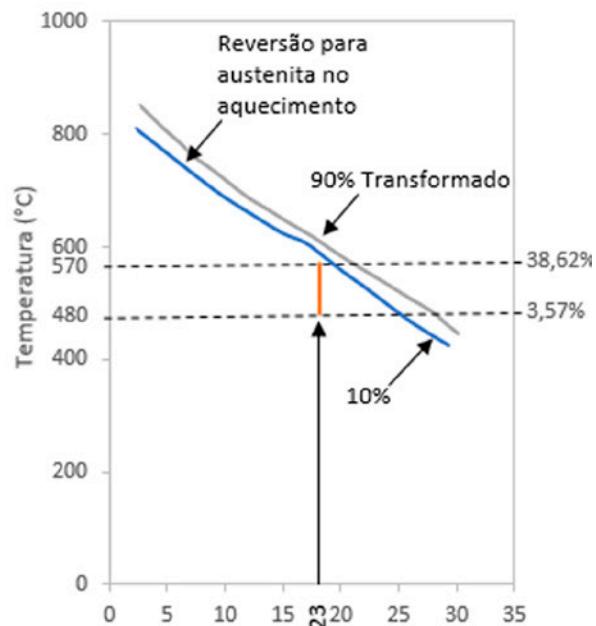


Fig. 8 – Fração de austenita revertida e sua relação com o diagrama de transformação metaestável no sistema Fe-Ni.
Fonte: Adaptado de SILVA; MEI, 2010.

Com relação ao ensaio de dureza, foi utilizada a técnica de microdureza Vickers segundo a norma ASTM E384 [12], onde foram efetuadas cinco indentações (impressões) em pontos aleatórios em cada amostra com carga aplicada de 0,2 kgf (1,961 N) e duração de 15 s. O ensaio foi realizado com o durômetro Shimadzu modelo HMV-G2 no Laboratório de Ensaios Mecânicos do Instituto Militar de Engenharia (IME).

A **tabela 4** apresenta os valores de dureza obtidos na escala HV.

Tab. 4 –Medidas de dureza.

Temperatura de envelhecimento (°C)	Dureza (HV)						Média	Desvio Padrão
	480	545	541	541	549	537		
510	525	525	529	529	525	526,6	2,19	
530	492	499	499	499	495	496,8	3,19	
550	478	485	485	485	482	483	3,08	
570	429	435	435	429	429	431,4	3,29	
Como recebido	299	297	295	300	297	297,6	1,95	

Os ensaios de tração foram realizados com a máquina universal de ensaio EMIC, modelo EMIC DL10000 com célula Trd 28, no Laboratório de Ensaios Mecânicos do Instituto Militar de Engenharia (IME).

Os corpos de prova possuem as dimensões em concordância com a norma ASTM E8/E8M [13]. Foram testados corpos de prova envelhecidos a 480 °C por 3 horas e como recebidos (somente solubilizados). Tal escolha se baseou na intenção de estimar o ganho de resistência mecânica no melhor cenário de tratamento de acordo com o resultado apresentado nos ensaios de dureza e compará-lo com o resultado do material somente solubilizado.

Foram confeccionados nove corpos de prova de tração e divididos em dois grupos. O grupo

A com 4 amostras recebeu somente o tratamento de solubilização e o grupo B com 5 amostras foi solubilizado e envelhecido a 480 °C por 3 horas.

A **tabela 5** mostra os valores encontrados para os corpos de prova do grupo A. Entretanto, um corpo de prova teve que ser descartado durante o ensaio.

Tab. 5 – Resultado do Ensaio de Tração para o Grupo A.

	Corpo de Prova 1	Corpo de Prova 2	Corpo de Prova 3	Média	Desvio Padrão
Resistência Mecânica (MPa)	945	931	977	951	24
Limite de Escoamento (MPa)	873	900	657	810	133

A **tabela 6** mostra os valores encontrados para os corpos de prova do grupo B.

Entretanto, para este grupo, um corpo de prova também teve que ser descartado durante o ensaio.

Tab. 6 – Resultado do Ensaio de Tração para o Grupo B.

	Corpo de Prova 1	Corpo de Prova 2	Corpo de Prova 3	Corpo de Prova 4	Média	Desvio Padrão
Resistência Mecânica (MPa)	1704	1609	1738	1744	1699	62
Limite de Escoamento (MPa)	511	960	1694	971	1034	489

4. Conclusão

Após a análise dos resultados, pode-se dizer que a fração de austenita revertida começa a ocorrer a temperaturas em torno de 500°C quando

submetida a 3 horas de tratamento térmico.

As análises de DRX evidenciaram um aumento progressivo da porção de austenita revertida no material associada ao aumento da temperatura de envelhecimento, ocorrendo um aumento mais acentuado acima de 550 °C.

As medidas de dureza do material evidenciaram uma diminuição da dureza das amostras em função do aumento da temperatura de envelhecimento, o que está relacionado ao aparecimento da fase austenítica.

O tratamento térmico de envelhecimento realizado a 480 °C por 3 horas se apresentou como satisfatório, potencializando as propriedades

mecânicas do material e dentre as condições testadas se mostrou a melhor opção de tratamento térmico.

Por fim, pode-se afirmar que embora o aço maraging 250 possua excelente potencial para aplicações que requerem alta performance, a escolha do tratamento térmico ao qual o material será submetido definirá suas propriedades. Como observado nos resultados obtidos, variações na ordem de 20 °C já são suficientes para que se tenha uma grande variação de propriedades mecânicas. Sendo recomendado, portanto, um excelente controle de temperatura na etapa de envelhecimento.

Referências bibliográficas

- [1] INCO – The International Nickel Company. 18 Per Cent Nickel Maraging Steels: Engineering Properties, Nickel Development Institute, 1976.
- [2] Schmidt, M.; Rohrbach, K.; Heat Treating of Maraging Steels, Heat Treating, ASM Handbook, ASM International, v. 4, p. 219-228, 1991.
- [3] Rohrbach, K.; Schmidt, M.; Maraging Steels. Properties and Selection: Irons, Steels, and High-Performance Alloys, ASM Handbook, ASM International, v.1. p. 1869-1887, 1990.
- [4] Santana, S.I.V. Microestrutura de um Aço Maraging 350 Submetido a Tratamentos Termomecânicos. Tese de Doutorado, Instituto Militar de Engenharia, Rio de Janeiro, 2017.
- [5] Rao, M. N. Progress in Understanding the Metallurgy of 18% Nickel Maraging Steels. International Journal of Materials Research, v.11, p.1594-1607, 2006.
- [6] Peters, D. T. A Study of Austenite Reversion During Aging of Maraging Steels. Transactions of ASM, v. 61, p. 62-74, 1968.
- [7] Ahmed, M.; Nasim, I.; Husain, S. W. Influence of Nickel and Molybdenum on the Phase Stability and Mechanical Properties of Maraging Steels. Journal of Materials Engineering and Performance, v. 3, p. 248-254, 1994.
- [8] Silva, A. L. V. C.; Mei P. R. Aços e Ligas Especiais. São Paulo. Blucher, 2010.
- [9] Pardal, J.M. et al. Influence of Temperature and Aging Time on Hardness and Magnetic Properties of the Maraging Steel Grade 300. Journal of Materials Science. v. 41. p. 2276-2281. 2007.
- [10] SAE – Society of Automotive Engineers. AMS6512 – H: Steel, Maraging, Bars, forgings, Tubing, and Rings 18Ni – 7.8Co – 4.9Mo – 0,10Al Consumable Electrode Vacuum Melted, Annealed. 2017.
- [11] United States Defense Standard. MIL-S-46850-D: Steel: Bar, Plate, Sheet, Strip, forgings, and Extrusions, 18 Percent Nickel Alloy, Maraging, 200 ksi, 250 ksi, 300ksi, High Quality. 1991.
- [12] ASTM – American Society for Testing Materials. E384 – 17: Standard Test Method for Microindentation Hardness of Materials. 2017.
- [13] ASTM – American Society for Testing Materials. E8/E8M – 16a: Standard Test Methods for Tension Testing of Metallic Materials. 2016.



PÓS-GRADUAÇÃO NO IME

Bolsas: CAPES, CNPq e FAPERJ

Mestrado

- Engenharia Mecânica
- Engenharia Nuclear
- Engenharia Elétrica
- Engenharia de Transportes
- Sistemas e Computação

Mestrado e Doutorado

- Química

- Ciência dos Materiais
- Engenharia de Defesa

Apoio:



Matrículas: Fluxo Contínuo

Tel.: (21) 2546-7114 – Fax: (21) 2546-7089 – www.ime.eb.br – sd1@ime.eb.br