Algoritmo Imunoinspirado Aplicado a Segurança Computacional

Thiago Giroto Milani, Fabricio Aparecido Breve
Universidade Estadual Paulista "Julho de Mesquita Filho"- UNESP
Rio Claro, Brasil
tmilani@rc.unesp.br, fabricio@rc.unesp.br

***Resumo*—Com o grande crescimento da área de informática e inovação tecnológica (era digital) cresce cada vez mais a necessidade de dispositivos e algoritmos capazes de aprender e reconhecer padrões. Segurança computacional se torna cada vez mais essencial com toda essa evolução, pois assim como a tecnologia cresce, os incidentes de segurança crescem duas vezes mais rápido. Com isso surge a necessidade de integrar essas duas vertentes da evolução, inteligência computacional e segurança de computadores, trazendo dispositivos e ferramentas capazes de reconhecer padrões de incidentes de segurança através da inteligência computacional.**

Área: Inteligência Computacional

I. Introdução

Recentemente vemos um grande avanço na utilização de algoritmos e modelos computacionais para resolver problemas ligado a biologia. Isso influenciou o crescimento de pesquisas na área de biologia computacional e bioinformática, sendo cada vez mais frequente a criação de grupos e eventos internacionais e nacionais ligados ao assunto. Conforme comentado por [5]

Cada vez mais os usuários de computador têm-se deparado com problemas mais sérios de segurança de computadores e redes. Com o crescimento da internet desde sua criação, e cada vez mais serviços e aplicações para ela, acabou sendo um bem essencial, e de extrema necessidade, trazendo assim mais destaque para pessoas mal-intencionadas disseminar vírus e *malware* com intenção de roubar ou atacar algum usuário ou empresa. São cada vez mais comuns notícias de invasão de sites.

Existem várias ferramentas de detecção de intrusão no mercado, porém poucas incorporam técnicas de aprendizado para o aprimoramento automático. Atualmente as ferramentas disponíveis em geral apresentam problemas de desempenho e são principalmente baseadas em comparação em assinaturas com banco de dados já existentes ou gerados pelo sistema.

Para se obter um maior desempenho é preciso uma maior base de dados de ataques, porém com um grande custo em eficiência. Outro grande problema com essas ferramentas é o grande número de falsos positivos gerados.

Além de já trabalhar com segurança computacional a algum tempo, com experimentos utilizando *honeypot* uma motivação pessoal para este artigo é a grande quantidade de ataques e mensagens de spam que são disseminados todos os dias, além dos recentes acontecimentos em escala global como o *malware wannacry*, o qual ocasionou a paralisação de diversos serviços essenciais totalmente dependentes da tecnologia e internet.

O objetivo da pesquisa é analisar algoritmos de segurança computacional imunoinspirados. Especificamente à pesquisa de algoritmos imunoinspirados aplicados a análise, classificação e reconhecimento de padrões em mensagens de *spam*.

II. Conceitos e técnicas

A.Detecção e Contenção de Redes de Computadores

Os sistemas de detecção de intrusão mais conhecidos são os *IDS* (*Intrusion Detection System*) que segundo [4] podem ser divididos da seguinte forma:

HIDS – Host-Based Intrusion Detection System: Monitora o sistema utilizando informações locais, como arquivos de *logs* ou agentes de auditoria. Ele pode ser capaz de monitorar acessos e alterações em arquivos e processos no sistema, entre outros aspectos. Exemplos de HIDS são o *Tripewire*, o *AIDE*, *OSSEC* e o *Sentrytools*.

NIDS – Network-Based Intrusion Detection System: Monitora o tráfego da rede, geralmente utilizando a interface da rede em modo promíscuo, como se fosse um *sniffer*. Exemplo de NIDS são o *Snort* e o *AAFID*.

Hybrid IDS – Hybrid Intrusion Detection System: Utiliza monitoração do sistema e de redes ao mesmo tempo para prevenir ataques. Com isso se tem o melhor dos sistemas HIDS e NIDS. Exemplo nessa categoria é o *Prelude*.

B.Mensagens de Spam

Uma grande maioria dos ataques computacionais, roubos de informação, e invasões de sistemas e redes se originam de mensagens de *spam* não detectadas ou pouco conhecimento por parte dos usuários para detecção manual desse tipo de mensagem.

Uma mensagens de *spam* pode ser de vários tipos como um *e-mail* falso, com informações genéricas com a intenção de enganar o leitor a ponto de extrair informações sigilosas, envio de propagandas não autorizadas, ou até mesmo

VII Workshop do Programa de Pós-Graduação em Ciência da Computação: "Interação entre Academia e Empresa", Unesp, Rio Claro, 14 e 15 de setembro de 2017.

ocasionar a instalação de aplicativos e ferramentas que irão proporcionar o roubo de informação ou acesso indevido a maquina e arquivos sigilosos (Figura 1). Uma definição mais formal diz que *spam* é o termo usado para se referir aos *e-mails* não solicitados, que geralmente são enviados para um grande número de pessoas. Quando este tipo de mensagem possui conteúdo exclusivamente comercial também é referenciado como UCE (*Unsolicited Commercial E-mail*). [2].

Fig. 1. Pseudo Código de SIAs Inspirado na Rede Idiotípica

Nesse trabalho os autores desenvolvem um classificador a partir de um sistema dinâmico não-linear baseado em uma rede de anticorpos modelado por equações diferenciais.

III. METODOLOGIA DE DESENVOLVIMENTO

Segundo [3], a ideia de utilizar SIAs em segurança de computadores surgiu de forma relativamente precoce, durante a evolução dos estudos em SIAs. Inicialmente as preocupações eram diferentes e os estudos focados em outras áreas.

Uma grande visão de SIAs para detecção de intrusão pode ser encontrada em [15]. Dois fatos observados por esses autores na época da publicação foram: i) todos os trabalhos avaliados por eles eram baseados em modelos com distinção *self* X *nonself*; ii) os trabalhos em sua grande maioria eram em seleção negativa.

Obviamente, novas abordagens surgiram após as publicações, porém "IDS Imunoinspirados" ainda têm muito espaço para crescer e muitas áreas a serem exploradas.

A detecção de SPAM utilizando SIAs ainda é pouco explorada. O primeiro modelo conhecido para classificação de mensagens de e-mail utilizando SIA foi proposto em [16]. O algoritmo atribui um peso para cada detector (linfócito), que é incrementado quando ocorre o reconhecimento do *spam* e decrementa em caso de mensagens legítimas. Conforme [3] aplicado a um conjunto de 1200 mensagens, conseguiu identificar 90% de *spam* e 99% das mensagens legítimas, após um treinamento de 1600 mensagens de *spam* e 1200 mensagem legítimas.

Outros trabalhos mais recentes, conseguiram índices maiores que 99% conforme mostrado em [3].

IV. CONSIDERAÇÕES FINAIS

Por fim o trabalho visa o desenvolvimento de um algoritmo baseando em sistemas imunológicos artificiais inspirado na rede idiotípica para a classificação e reconhecimento de padrões em mensagens de *spam*, tendo como principal comparativo para os experimentos o algoritmo e testes mencionado no artigo, "Algoritmos Imunoinspirados Aplicados em Segurança Computacional: Utilização de Algoritmos Inspirados no Sistema Imune para Detecção de Intrusos em Redes de Computadores" de J. Q. Uchôa (2009).

REFERÊNCIAS

1. S. Forest, S. A. Holfmeyr, A. Somayaji, "Computer Imunology" Comunications of ACM, Vol. 40, ed. 10, pag. 88 – 96, 1997;
2. Cert.br. Centro de estudos, resposta e tratamento de incidentes de segurança no brasil, cartilha de segurança para internet. Cartilha de Segurança para a Internet, 2016, 2016. URL https://cartilha.cert.br
3.
4. T. Oda, T. White, "Increasing the Accuracy on a Spam-Detecting Artificial Immune System". In: Proceedings of the Congress on Evolutionary Conputation (CEC 2003), Canberra, Autralia, Decenber, 2003, Australia: IEEE, 2003, v.1, p. 390 – 396;