MODELAGEM DE UM SISTEMA DE SEGURANÇA IMUNOLÓGICO

RESUMO

O sistema imunológico humano provê uma rica fonte de inspiração para a segurança de redes de computadores. Além de representar um modelo bastante próximo das condições em que a maioria das redes de computadores se encontra, o sistema imunológico humano possui uma série de características desejáveis a um sistema de segurança. Este trabalho apresenta uma visão geral da analogia entre segurança de computadores e imunologia, e propõe um modelo de sistema de segurança baseado no sistema imunológico humano.

ABSTRACT

The human immune system provides a rich source of inspiration for computer network security. Besides the fact that it represents a model that very closely resembles the conditions in which most computer networks exist, the human immune system has many features that are desirable for a security system. This work presents an overview of the analogy between computer security and immunology, and proposes a security system model based on the human immune system.

1 Introdução

Do ponto de vista tradicional de segurança de computadores, é possível garantir a integridade de um sistema observando alguns critérios. É necessário especificar e implantar corretamente uma política de segurança, implementar corretamente os programas e configurar adequadamente o sistema [10]. Porém, na prática, observa-se que as políticas de segurança, as implementações dos programas e as configurações dos sistemas podem conter falhas, tornando a segurança imperfeita [6, 8].

Um nível melhor de segurança pode ser atingido adotando-se recursos adicionais e melhores modelos, que representem de maneira mais próxima as condições em que a maioria das redes de computadores se encontra—um ambiente hostil e sujeito a falhas. É possível encontrar na natureza um modelo de defesa que apresenta uma série de características desejáveis a um sistema de segurança: o sistema imunológico humano.

O sistema imunológico, por ser capaz de garantir a sobrevivência de um indivíduo durante cerca de 70 anos, mesmo que ele se depare, a cada dia, com bactérias e vírus potencialmente mortais, apresenta um paralelo bastante forte com a segurança de redes de computadores.

A analogia entre problemas de segurança e processos biológicos foi inicialmente reconhecida no início de 1987, quando o termo "vírus de computador" foi introduzido por Adelman [1]. E a conexão entre imunologia e segurança de computadores teve início em 1994 com as publicações [2, 3], desencadeando uma série de outros trabalhos.

Os trabalhos iniciais concentravam-se em mecanismos isolados do sistema imunológico e como eles poderiam ser aplicados para melhorar a segurança de um sistema ou rede de computadores. Mais recentemente, as pesquisas passaram a considerar a estrutura de funcionamento do sistema imunológico como modelo de desenvolvimento de um sistema de segurança, baseando-se em uma série de princípios característicos do sistema de defesa do corpo humano.

Entretanto, a maioria dos esforços concentra-se no desenvolvimento de sistemas de detecção de intrusão, o que explora apenas uma parte do modelo oferecido pelo sistema imunológico. Utilizando esse modelo, idealiza-se um sistema de segurança que, além de detectar anomalias, seja capaz de elaborar um plano de resposta especializado e efetuar o contra-ataque. E talvez o mais importante, um sistema que possua a mesma capacidade de aprendizado e adaptação do sistema imunológico, podendo reagir a ataques desconhecidos.

Nesse sentido, este trabalho apresenta um modelo de arquitetura de segurança baseado no sistema de defesa do corpo humano e está organizado como segue. A seção 2 descreve sucintamente as estruturas e o funcionamento do sistema imunológico humano. Na seção 3 é apresentada uma visão geral da analogia entre segurança de computadores e imunologia. A seção 4 propõe uma modelagem para um sistema de segurança imunológico. E, por fim, a seção 5 tece conclusões acerca deste trabalho.

2 Sistema imunológico humano

É impossível entender o sistema imunológico como modelo de um sistema de segurança de computadores sem antes compreender seu funcionamento. Nesta seção são apresentadas as estruturas básicas

do sistema imunológico humano e as etapas da resposta imunológica.

2.1 Organização estrutural

O sistema imunológico é dividido em sistema inato e sistema adaptativo [9]. O sistema inato é caracterizado por sua natureza congênita e por sua capacidade limitada de diferenciar um agente patogênico de outro, reagindo de maneira semelhante contra a maioria dos agentes infecciosos. O sistema inato constitui a primeira linha de defesa contra a ação de micróbios, e sua resposta, por não ser específica para um determinado micróbio, é, na maioria das vezes, insuficiente. Seus principais componentes são as barreiras físicas e químicas, como a pele e os ácidos gástricos; as proteínas do sangue, incluindo as proteínas complementares e outros mediadores de inflamação; e as células conhecidas como fagócitos (macrófagos, monócitos e neutrófilos), responsáveis pela eliminação de partículas estranhas.

Em contraste com o sistema inato, o sistema adaptativo é capaz de identificar especializadamente um determinado agente patogênico, permitindo uma resposta mais eficiente. Além disso, ele é capaz de "memorizar" um agente infeccioso e responder mais vigorosamente a novas exposições a esse micróbio. Os componentes do sistema adaptativo são os linfócitos (linfócitos T e linfócitos B) e seus produtos, como os anticorpos.

Os mecanismos de ambos os sistemas, inato e adaptativo, constituem um sistema integrado de defesa em que um grande número de células e moléculas agem cooperativamente [9]. O sistema inato não só provê a primeira linha de defesa, mas também atua em diversas etapas da resposta do sistema adaptativo.

2.2 Resposta imunológica

Os imunologistas tradicionalmente descrevem o problema resolvido pelo sistema imunológico como o problema de distinguir o normal (ou *self*) do estranho (ou *nonself*) e eliminar o que for estranho. O *self* é tido como as células e moléculas do corpo, e o *nonself* é qualquer material estranho, particularmente bactérias, parasitas e vírus.

A distinção entre *self* e *nonself* é uma tarefa bastante difícil por algumas razões [9]. Primeiro, os componentes do corpo humano são contruídos a partir da mesma matéria prima dos *nonself*, basicamente proteínas. O sistema imunológico detecta a presença de um agente patogênico através da percepção das proteínas desse invasor. Além disso, a quantidade de padrões diferentes de proteínas que o sistema imunológico deve reconhecer é muito maior que a capacidade do corpo humano de gerar esses padrões.

Uma vez que o agente patogênico transpôs as barreiras iniciais do sistema inato (a pele, por exemplo), inicia-se a resposta imunológica. O processo todo é baseado no reconhecimento de proteínas estranhas, encontradas na superfície dos agentes infecciosos, chamadas antígenos. Esse reconhecimento se dá pela reação das proteínas da superfície das células do sistema imunológico, chamadas receptores, com os antígenos dos invasores.

A reação entre os receptores e os antígenos é determinada por suas propriedades físicas e químicas, de modo que essa reação é altamente específica e cada receptor reconhece um conjunto limitado de antígenos estruturalmente relacionados [9]. A figura 1 ilustra essa reação.

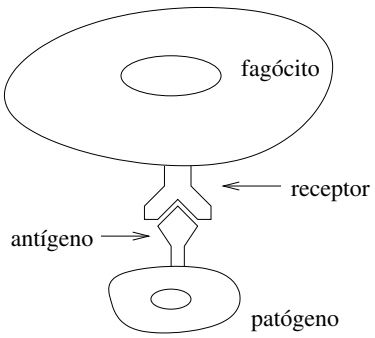

Figura 1: Reação entre receptor e antígeno.

A habilidade de reconhecer a maioria dos antígenos requer uma grande diversidade de receptores. Essa diversidade é parcialmente atingida pela geração de receptores através de um processo genético que introduz um fator aleatório [9]. Porém, essa geração aleatória pode resultar em receptores que reagem com proteínas *self*, causando problemas de autoimunidade.

O sistema imunológico resolve o problema de autoimunidade através de um processo chamado seleção negativa. Durante esse processo, as células do sistema imunológico, recém criadas, passam por um estágio de maturação em um órgão chamado timo, onde a maioria das proteínas do corpo circulam. As células que reagem com alguma dessas proteínas são eliminadas.

Uma vez na circulação, se o receptor de uma célula do sistema imunológico reage com antígenos, em uma concentração suficiente, um reconhecimento ocorre, disparando um conjunto complexo de eventos que leva à eliminação dos micróbios. Esse conjunto de eventos é denominado resposta imunológica, e pode ser dividido em três fases [9], como segue.

Fase de detecção

Esta fase é iniciada com a detecção de microrganismos estranhos, por parte dos fagócitos presentes na corrente sanguínea e tecidos. Os fagócitos possuem uma capacidade intrínseca de reconhecer microrganismos estranhos [9], uma vez que são uma espécie de coletores de lixo responsáveis pela eliminação de partículas indesejáveis.

Ocorre ainda um processo chamado opsonização, onde uma série de proteínas complementares revestem a superfície dos micróbios, aprimorando a capacidade de detecção dos fagócitos. Esse aprimoramento se dá pela reação dos receptores dos fagócitos com essas proteínas. As proteínas complementares participam também no processo de inflamação, causando o aumento do suprimento de sangue e migração de outras células de defesa para o local da infecção.

O fagócito que detectar algum agente patogênico, reage ao invasor através do processo de fagocitose[1]. Com a dissolução do micróbio no interior do fagócito, este passa a produzir moléculas de proteína, chamadas de moléculas MHC (*Major Histocompatibility Complex*), contendo fragmentos dos antígenos do micróbio. Tais moléculas MHC são expelidas pela membrana plasmática dos fagócitos, ficando presas na superfície.

Fase de ativação do sistema adaptativo

Esta fase inicia-se com o processo de apresentação dos antígenos estranhos. Durante esse processo, os fagócitos envoltos por moléculas MHC "apresentam" os fragmentos de antígenos aos linfócitos T, onde estes reconhecem um antígeno específico nas moléculas MHC da superfície dos fagócitos. Os linfócitos T possuem receptores em sua superfície que reagem com um tipo específico de antígeno, permitindo a identificação precisa do agente patogênico.

Os linfócitos T, ao reagirem com os antígenos, passam a multiplicar-se, gerando três tipos de clones [9]:

- *linfócitos T ajudantes*: estimulam a ação dos fagócitos e a multiplicação dos linfócitos B, através da liberação de citozinas (proteínas);
- *linfócitos T citotóxicos*: capazes de identificar as células contaminadas, através das moléculas MHC com fragmentos de antígenos que tais células produzem, e destruí-las;
- *linfócitos T de memória*: guardam informações a respeito do agente patogênico em sua carga genética, permitindo uma resposta mais aprimorada a uma nova exposição a esse micróbio.

Os linfócitos B, que possuem anticorpos específicos para o agente patogênico, também reagem com os antígenos, iniciando sua multiplicação e diferenciação em dois tipos de células:

- *células plasmáticas*: linfócitos B responsáveis pela produção de anticorpos;
- *linfócitos B de memória*: semelhantes aos linfócitos T de memória.

[1] Processo pelo qual uma célula envolve uma partícula com seu próprio corpo [9].

Dessa forma, é recrutado um exército especializado para combater o agente infeccioso.

Fase de contra-ataque

Nesta fase, as células plasmáticas passam a produzir anticorpos que reagem especificamente com os antígenos dos micróbios, impedindo que estes contaminem outras células. Além disso, os anticorpos que revestiram a superfície dos micróbios permitem que os fagócitos os identifiquem com maior precisão.

Os linfócitos T citotóxicos passam a destruir as células infectadas e, conforme a concentração de antígenos estranhos diminui, os estímulos químicos são gradativamente inibidos, até que o contra-ataque chega ao fim.

3 Imunologia computacional

Nesta seção é discutida, mais detalhadamente, a analogia entre o sistema imunológico e segurança de redes de computadores.

3.1 Paralelos entre imunologia e segurança de computadores

Alguns paralelos entre o sistema imunológico e a segurança de redes de computadores são imediatos, como mostra a tabela 1.

Sistema Imunológico	Segurança de Redes
Sistema de filtragem, composto de cílios, pele, mucosas e ácidos	*Firewall*
Amígdalas	Bode expiatório, *booby traps* armados pelo administrador de segurança
Inserção de DNA de vírus no interior das células atacadas	*Buffer overflow* e instalação de *trojan horses*
Detecção dos agentes patogênicos pelos macrófagos	Sistema de detecção de intrusão
Produção de moléculas MHC com fragmentos de antígenos	Comportamento dos processos
Ativação dos linfócitos específicos	Análise forense
Antígeno dos agentes patogênicos	Assinatura digital do ataque
Memória imunológica	Alimentação de uma base de dados
Destruição das células contaminadas e dos agentes patogênicos	Medidas de contenção, como finalização de processos e encerramento de conexões

Tabela 1: Paralelos entre o sistema imunológico e segurança de redes de computadores.

Além desses paralelos, é possível identificar algumas abstrações que conectam o sistema imunológico com a segurança de redes e sistemas. Por exemplo, é possível imaginar o corpo humano como sendo um *host* em uma rede, e os processos executando nesse *host* como as células do corpo. Ou ainda, uma rede

pode ser vista como o corpo e os *hosts* dessa rede como as células. Essas abstrações provêem possíveis arquiteturas para um sistema de segurança que explore a analogia com o sistema imunológico.

3.2 Princípios do sistema imunológico

O estudo do sistema imunológico revela um conjunto de princípios organizacionais que podem servir de base para o desenvolvimento de um sistema de segurança [6]. Esses princípios são apresentados como segue.

Descentralização e localidade

No sistema imunológico, a ação de uma célula de defesa não é iniciada por algum mecanismo centralizador, mas quando alguma condição anormal é detectada pela célula. Essa característica descentralizada indica a inexistência de um ponto central de falha no sistema, garantindo uma maior robustez. Além de descentralizado, o sistema imunológico é capaz de agir em diferentes localidades paralelamente, não se fazendo necessária a mobilização de todo o "exército" para um determinado ponto de infecção.

Multi-camada

Nenhum mecanismo do sistema imunológico garante isoladamente a segurança do corpo. Ao invés disso, esses mecanismos constituem um sistema integrado de defesa em que um grande número de células e moléculas agem cooperativamente, cada qual exercendo uma função especializada.

Diversidade

O sistema imunológico é capaz de identificar uma grande diversidade de antígenos, proporcionando uma reação especializada contra a invasão de diferentes tipos de agentes patogênicos.

Robustez e tolerância a falhas

Nenhuma célula isolada do sistema imunológico é essencial, podendo ser substituída, caso seja infectada ou morta, sem comprometer o funcionamento do sistema. A disponibilidade de células combinada à ausência de um controle hierárquico caracteriza a robustez e a tolerância a falhas do sistema imunológico.

Autonomia

O sistema imunológico, como um todo, não requer um gerenciamento ou manutenção externos. Ele pode, autonomamente, detectar, classificar e eliminar os agentes patogênicos, bem como efetuar a remoção e substituição de suas células danificadas.

Adaptabilidade e memória

O sistema imunológico possui a capacidade de aprendizado na detecção de novos agentes patogênicos, armazenando a habilidade adquirida de reconhecimento em uma memória, conhecida como memória imunológica. A capacidade de reconhecimento especializa-se a cada agente patogênico similar reconhecido, tornando respostas futuras mais eficientes se comparadas às anteriores.

Auto-proteção

Qualquer célula do corpo humano pode ser atacada por um agente patogênico, incluindo as células que compõem o sistema imunológico. Sendo assim, os linfócitos podem proteger o corpo de outros linfócitos comprometidos por uma infecção.

Identificação por meio de comportamento

Em uma infecção, as células contaminadas são identificadas através das moléculas MHC com fragmentos de antígenos que tais células produzem, caracterizando um "comportamento anormal".

Tamanho do exército

O sistema imunológico regula o número de linfócitos de maneira a desenvolver um contra-ataque suficiente para a quantidade de agentes patogênicos.

3.3 Definição de *self* e *nonself* em um sistema computacional

O problema de proteger sistemas computacionais de intrusões pode ser visto, assim como no caso do sistema imunológico, como um problema de distinguir *self* de *nonself* [2, 5]. A definição de *self* em um sistema computacional pode ser feita em termos de padrões de acesso à memória em um *host*, pacotes TCP/IP entrando e saindo de um *host*, comportamento coletivo de uma rede, tráfego através de um roteador, sequências de instruções em um programa, padrões de comportamento de usuários, ou até mesmo padrões de digitação no teclado. Enfim, qualquer mecanismo que permita a identificação de algo estranho, como por exemplo, usuários não-autorizados, código externo na forma de vírus de computador ou *worm*, ações ilegítimas de usuários autorizados, ou dados corrompidos.

Além disso, a definição de *self* deve ser tolerante a mudanças legítimas [2], incluindo edição de arquivos, instalação de novas aplicações, novos usuários, mudanças nos hábitos de um usuário, e atividades rotineiras de um administrador de sistemas.

3.4 Pesquisas em imunologia computacional

Nesta seção são apresentados alguns exemplos de como as idéias de imunologia estão sendo aplicadas no combate a problemas de segurança de redes e sistemas computacionais.

3.4.1 Método de detecção de intrusão

Forrest, Hofmeyr e Somayaji apresentam em [4] um método de detecção de intrusão, baseado no sistema imunológico humano. A definição de *self* é feita em termos de sequências curtas de chamadas ao sistema executadas por processos privilegiados, provendo uma assinatura compacta do comportamento normal dos processos [4].

O método é baseado na construção de uma base de dados do comportamento normal de cada programa de interesse. Cada base de dados é específica para uma determinada arquitetura, versão e configuração de software, política de administração e padrões de uso. Depois de construída, a base de dados é utilizada para monitorar o comportamento do programa relacionado. As sequências de chamadas ao sistema da base de dados formam o conjunto de padrões normais do processo, e as sequências não encontradas na base de dados indicam anomalias no comportamento do processo.

O método proposto em [4] apresenta dois estágios. No primeiro estágio, são extraídos traços de comportamento normal dos processos e construídas as bases de dados. No segundo estágio, as sequências de chamadas ao sistema executadas pelos processos são comparadas com os padrões armazenados em suas bases de dados. Se uma sequência não é encontrada na base de dados do processo, é reportado um *mismatch*, representando uma possível invasão.

3.4.2 Algoritmo distribuído de detecção de mudanças

Este algoritmo, também apresentado em [4], é baseado no processo de seleção negativa do sistema imunológico. Segundo o algoritmo, um conjunto de dados considerado *self* é monitorado contra mudanças ilegítimas, como segue.

1. Gere um conjunto de detectores que falham em reconhecer o *self*;
2. Use os detectores para monitorar os dados protegidos;
3. Sempre que um detector for ativado (reconhecer algo), uma mudança deve ter ocorrido em um dado protegido, e a localização da mudança é conhecida.

O *self* é definido como um conjunto de *strings* de mesmo tamanho, geradas pela segmentação dos dados protegidos em *substrings* de igual comprimento. Cada detector é também uma *string* de tamanho igual ao *self*. O reconhecimento por parte dos detectores é modelado como o "casamento" entre pares de *strings*. Um "casamento" perfeito entre duas *strings* de mesmo tamanho indica que os símbolos são idênticos em cada posição da *string*. Entretanto, o algoritmo utiliza um tipo especial de "casamento", chamado r-bits contíguos (mais plausível com o modelo imunológico). Desse modo, o algoritmo busca por r posições contíguas, cujos símbolos são idênticos, em um par de *strings*. Os detectores são gerados aleatoriamente, e passam pelo processo de seleção negativa, eliminando aqueles que "casam" com alguma *string self*.

3.4.3 Sistema imunológico artificial

Hofmeyr e Forrest apresentam em [7] um sistema imunológico artificial, que detecta conexões TCP não usuais com a rede protegida pelo sistema. A definição de *self* é feita em termos de conexões TCP, determinadas pelos endereços IP de origem e destino, e pela porta TCP de serviço. Cada conexão é representada por uma *string* de 49 bits, sendo que o *self* é o conjunto das conexões que normalmente são observadas na rede.

O sistema apresenta um único tipo de detector, que combina propriedades de várias células do sistema imunológico, podendo assumir diferentes estados. Cada detector é representado por uma *string* de 49 bits e um conjunto de estados. A geração dos detectores é aleatória, passando pelo processo de seleção negativa, e o reconhecimento se dá pelo "casamento" de r-bits contíguos.

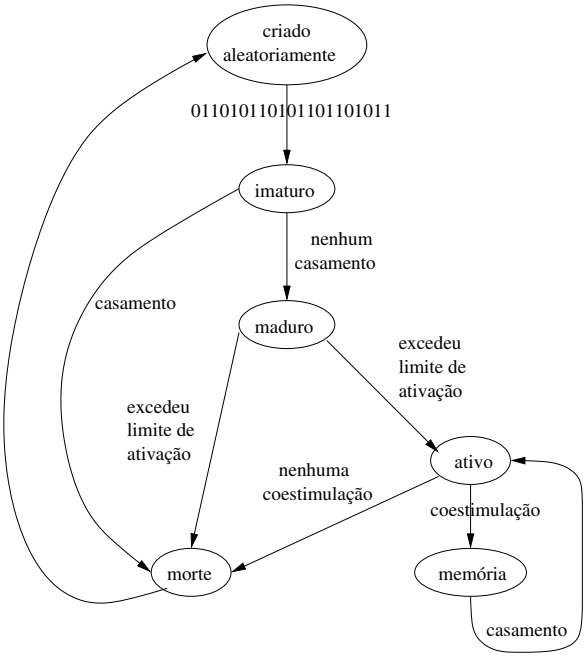

Figura 2: Ciclo de vida de um detector.

A figura 2 resume o ciclo de vida de um detector. Um detector é inicialmente criado aleatoriamente e, então, permanece imaturo por um certo período de tempo. Se o detector "casar" com alguma *string* de conexão uma única vez, enquanto imaturo, ele é substituído por um novo (também gerado aleatoriamente). Se o detector sobreviver ao processo de seleção negativa, ele existirá por um tempo finito. Ao final de sua vida ele é substituído por um novo detector, a menos que tenha ultrapassado um limite de "casamentos" e tenha se tornado um detector de

memória. No estágio maduro, o detector é ativado se exceder o limite de "casamentos" para ativação. Se o detector estiver ativo e não receber coestimulação (do administrador de segurança, por exemplo), ele morre. Entretanto, se houver coestimulação o detector torna-se um detector de memória, com um tempo de vida indefinido. Um detector de memória necessita de apenas um "casamento" para ser ativado.

4 Proposta de um sistema de segurança imunológico

Esta seção apresenta a modelagem inicial de um sistema de segurança imunológico, desenvolvida pelos autores deste artigo como parte de um projeto de pesquisa na área de imunologia computacional. O projeto, conduzido pelos autores deste trabalho, visa o desenvolvimento de um sistema de segurança imunológico que seja capaz de detectar anomalias, elaborar um plano de resposta especializado e efetuar o contra-ataque. Agregando, ainda, a capacidade de aprendizado e adaptação do sistema imunológico, podendo reagir a ataques desconhecidos. A modelagem inicial de tal sistema de segurança é apresentada como segue.

4.1 Modelo estrutural

O sistema de segurança imunológico proposto obedece a mesma organização estrutural do sistema imunológico humano, sendo composto pelos sistemas inato e adaptativo. Além da estrutura, a equivalência estende-se às funcionalidades de cada componente, de modo que o funcionamento global do sistema de defesa humano é reproduzido. A figura 3 mostra a estrutura do sistema de segurança, ilustrando os componentes e suas interações. A descrição de cada componente é apresentada como segue.

4.1.1 Sistema inato

O sistema inato é responsável pela detecção de condições anormais e extração de informações que permitam a identificação do ataque, por parte do sistema adaptativo. Embora ainda não se tenha conhecimento do tipo de ataque, o sistema inato executa medidas genéricas com o intuito de retardá-lo. Tal sistema é composto pelos seguintes componentes.

Detector

O detector é o componente equivalente, em funcionamento e estrutura, ao fagócito do sistema imunológico humano, sendo responsável por monitorar o sistema em busca de anomalias.

Para a detecção de condições anormais, o detector utiliza um conjunto de padrões que representam

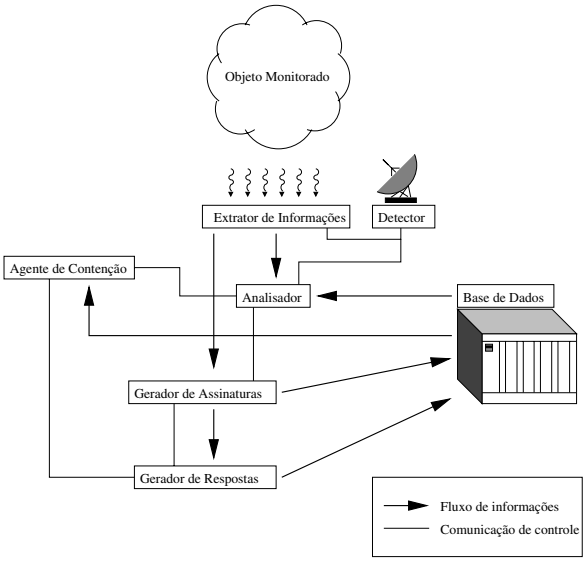

Figura 3: Modelo estrutural do sistema de segurança imunológico.

tais condições, chamados de receptores. Os receptores podem ser implementados, por exemplo, como sequências de chamadas ao sistema ou padrões de pacotes que trafegam na rede. Quando um padrão observado assemelha-se a um receptor, ocorre um *match*, indicando a detecção de uma condição anormal. Como consequência, o detector alerta o restante do sistema de segurança, dando início ao processo de defesa.

Algumas abordagens podem ser consideradas para a geração dos receptores, podendo ser destacadas as seguintes:

- Geração aleatória de padrões. São considerados receptores válidos, os padrões que não ocasionam um *match* com algum padrão de condição normal (processo de seleção negativa do sistema imunológico humano);
- Geração de padrões através da permutação de receptores conhecidos. Igualmente à abordagem anterior, os padrões gerados só são considerados receptores válidos caso não ocasionem um *match* com algum padrão de condição normal;
- Geração de receptores de acordo com assinaturas conhecidas de ataques.

Extrator de informações

O extrator de informações atua junto ao detector, armazenando informações que serão utilizadas pelo sistema adaptativo no processo de identificação do ataque.

4.1.2 Sistema adaptativo

O sistema adaptativo é responsável pela identificação precisa do ataque, através das informações

fornecidadas pelo sistema inato. Uma vez identificado o ataque, o sistema adaptativo executa as medidas de contenção específicas. Caso o ataque não seja conhecido, as informações recebidas são processadas com o intuito de gerar uma caracterização do ataque e responder da maneira mais eficaz possível. O sistema adaptativo é composto pelos seguintes componentes.

Base de dados

É responsável por armazenar as assinaturas dos ataques, relacionando-as com as respectivas medidas de contenção. A base de dados do sistema de segurança, modelada na seção 4.3, implementa a memória imunológica do corpo humano. Sua alimentação pode ser efetuada das duas maneiras seguintes:

- Alimentação pelo administrador do sistema de segurança, permitindo a inserção de assinaturas de ataques já identificadas e medidas de contenção já testadas. Esta abordagem faz uma analogia ao processo de vacinação no tratamento de doenças conhecidas;
- Alimentação pelo próprio sistema de segurança, quando este gera automaticamente a assinatura de um ataque desconhecido e identifica o conjunto de medidas de contenção mais eficazes. Esta segunda abordagem é empregada no processo de aprendizagem do sistema de segurança imunológico.

Analisador

O analisador é o componente encarregado da identificação dos ataques, utilizando-se das informações fornecidas pelo extrator e de buscas na base de dados do sistema.

Gerador de assinaturas

No caso de não haver alguma assinatura na base de dados que identifique o ataque, o gerador de assinaturas é invocado para filtrar as informações fornecidas pelo extrator e identificar padrões que caracterizam o ataque.

Agente de contenção

Este componente é invocado, assim que a identificação do ataque é feita, para executar as medidas de contenção relacionadas ao ataque.

Gerador de respostas

Quando um ataque novo é identificado, ainda não existe uma resposta específica para o mesmo. Nesse caso, o gerador de respostas encarrega-se da criação de um conjunto de medidas capazes de responder, com certa especialidade, ao ataque. A resposta criada é uma seleção feita em um conjunto de medidas genéricas preestabelecidas, com base na análise da assinatura do ataque.

4.2 Modelo da assinatura dos ataques

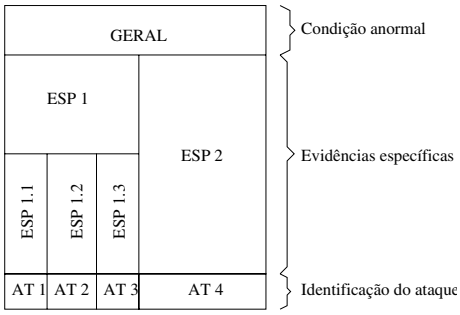

Figura 4: Modelo da assinatura dos ataques.

A assinatura do ataque é composta por um conjunto de informações organizadas segundo o modelo hierárquico ilustrado na figura 4. O topo da hierarquia contém informações que possibilitam a detecção de uma condição anormal, sem necessariamente caracterizar um tipo de ataque específico. Tal detecção é feita pelo sistema inato, responsável pela monitoração de uma grande quantidade de dados de maneira não especializada. Devido ao grande processamento, as informações do topo da hierarquia devem ser simples e abrangentes, permitindo a caracterização de quaisquer condições anormais causadas pelos ataques.

À medida em que a hierarquia é percorrida, as informações tornam-se mais específicas, possibilitando a identificação precisa do ataque. Tais informações podem ser quaisquer evidências deixadas pelo ataque, como, por exemplo, registros em arquivos de log, alterações em determinados arquivos ou execução de determinados programas.

A modelagem da assinatura como uma estrutura hierárquica leva em consideração o fato de que vários ataques podem ter características em comum. Dessa forma, o prefixo de uma assinatura pode ser compartilhado por vários ataques, como ilustrado na figura 5.

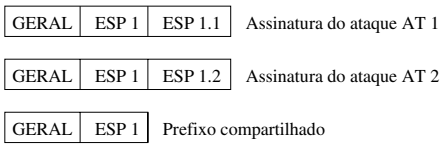

Figura 5: Exemplo do compartilhamento de prefixo de assinatura.

4.3 Modelo da base de dados

A base de dados do sistema de segurança imunológico constitui-se de duas componentes relacionadas entre si. Uma base de assinaturas de ataques e uma base de respostas, como ilustrado na figura 6.

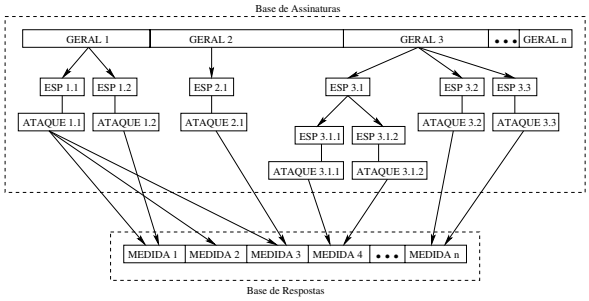

Figura 6: Modelo da base de dados do sistema de segurança imunológico.

A escolha da estrutura dessas componentes é fundamental, tendo em vista a sua implicação direta no funcionamento geral do sistema de segurança. Essa implicação vem do fato de todo o sistema adaptativo interagir com a base de dados constantemente.

A base de assinaturas pode ser implementada através de uma estrutura de dados não linear, como, por exemplo, a árvore de prefixos [11], que se adequa ao modelo de assinatura proposto. As assinaturas são organizadas por uma lista indexada pelo topo da hierarquia, de modo que ataques com prefixos em comum compartilham a mesma entrada da lista. Cada entrada da lista especializa-se até a caracterização específica dos ataques.

Uma vez identificado um ataque, algumas abordagens podem ser consideradas para relacioná-lo às medidas de contenção. Uma abordagem é fazer com que o ataque referencie diversas medidas da base de respostas, de modo que cada medida realiza uma única tarefa e o conjunto referenciado forma a resposta ao ataque. Um exemplo desse esquema é ilustrado na figura 6, onde o ataque 1.1 tem como resposta as medidas 1, 2 e 3. Outra abordagem considera cada entrada da lista de respostas como um conjunto de tarefas. Esse conjunto representa a medida de contenção completa, de modo que cada ataque referencia um único elemento da base de respostas, como o ataque 3.2 da figura 6.

Apesar de parecer conveniente, a estrutura da base de dados apresentada é inadequada para modelar ataques onde não é possível estabelecer uma relação única de ordem entre alguns eventos. Nesse caso, um mesmo ataque pode ter diversas representações, onde cada representação difere das demais apenas pela permutação de alguns eventos, como ilustrado na figura 7.

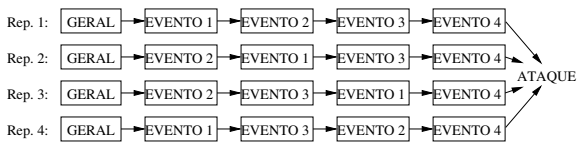

Figura 7: Exemplo de ataque hipotético com quatro representações distintas. Os eventos 1, 2 e 3 não possuem uma ordenação única.

A figura 8 apresenta a base de assinaturas composta pelas quatro representações do ataque exemplificado na figura 7. Nessa base de assinaturas, uma grande quantidade de dados é armazenada para representar um ataque, considerando que a maior parte desses dados aparece duplicada. Visando reduzir essa duplicação, outra organização para a base de assinaturas pode ser considerada.

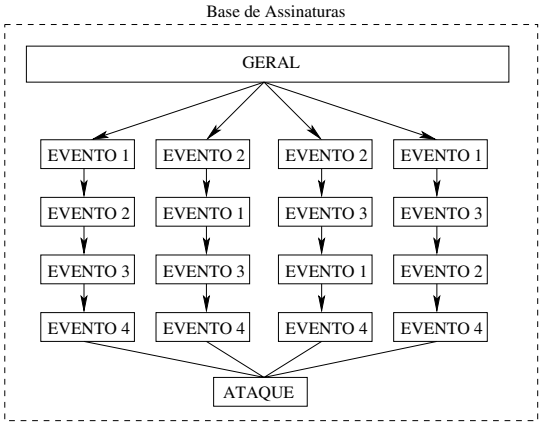

Figura 8: Base de assinaturas para o exemplo da figura 7. Os eventos de 1 a 4 representam as camadas específicas.

Uma nova organização para a base de assinaturas, mais genérica que a anterior, possibilita que a relação de ordem entre alguns eventos seja desconsiderada. A figura 9 mostra a representação de um ataque composto por quatro eventos nessa nova organização. Nesse ataque, não existe uma ordenação estabelecida entre os eventos 1, 2 e 3, e o evento 4 só ocorre após os três primeiros.

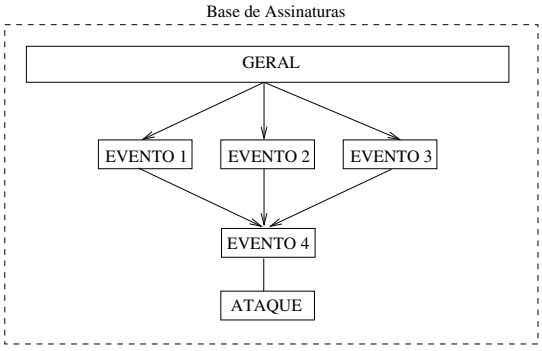

Figura 9: Outra organização para a base de assinaturas.

A assinatura mostrada na figura 9 é mais simples que a assinatura apresentada na figura 8, pelo fato de não possuir dados duplicados. Além disso, representa duas sequências adicionais[2] de eventos não encontradas na figura 7, o que torna as duas assinaturas diferentes. De fato, mesmo utilizando o

[2] As sequências adicionais são: (GERAL, EVENTO 3, EVENTO 2, EVENTO 1, EVENTO 4) e (GERAL, EVENTO 3, EVENTO 1, EVENTO 2, EVENTO 4).

novo esquema, a base de assinaturas pode apresentar dados duplicados, como no esquema anterior.

4.4 Modelo de funcionamento

O funcionamento do sistema de segurança proposto, ilustrado na figura 10, pode ser dividido em três fases, assim como ocorre na resposta imunológica humana. Tais fases são apresentadas como segue.

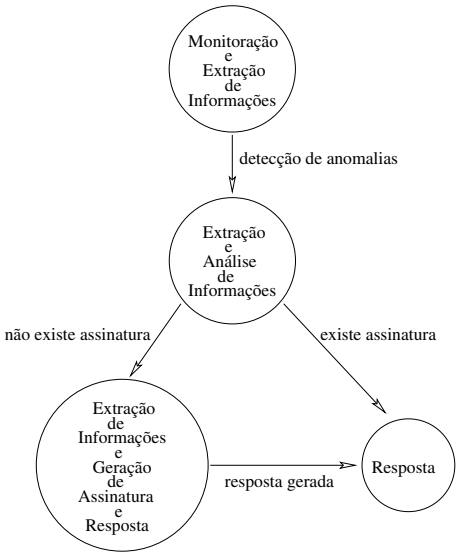

Figura 10: Modelo de funcionamento do sistema de segurança imunológico.

Fase de detecção

Quando o detector identifica alguma condição anormal no objeto monitorado[3], ele ativa o analisador, para iniciar a identificação do ataque, e aciona o extrator, que passa a fornecer informações ao sistema adaptativo.

Fase de ativação do sistema adaptativo

O analisador filtra as informações providas pelo extrator e busca por uma assinatura correspondente na base de dados. Se existir, ele ativa o agente de contenção para executar a resposta. Caso não haja uma assinatura, o analisador ativa o gerador de assinaturas para criá-la, com base nas informações obtidas pelo extrator. Em seguida, o gerador de assinaturas invoca o gerador de respostas e passa a lhe fornecer dados mais específicos, usados na elaboração das medidas de contenção. Quando os geradores concluem suas tarefas, a base de dados é atualizada e o agente de contenção é acionado, pelo gerador de respostas, para combater o ataque.

Fase de contra-ataque

O agente de conteção busca, na base de dados, as medidas relacionadas ao ataque e as executa.

[3] O objeto monitorado pode ser, por exemplo, um processo ou uma conexão de rede.

5 Conclusão

A analogia entre segurança de computadores e imunologia constitui uma rica fonte de inspiração para o desenvolvimento de novos mecanismos de defesa, sejam algoritmos e técnicas de detecção de intrusão, políticas de segurança mais atentas à existência de falhas, ou sistemas completos de segurança. Um sistema de defesa, modelado segundo o sistema imunológico, certamente teria uma noção mais sofisticada de identidade [3], o que a maioria dos administradores de redes e sistemas não possui (poucos são os que conhecem o comportamento usual de seu sistema ou rede de computadores).

Apesar de existirem vários trabalhos relacionados à analogia, muito ainda pode ser explorado, como por exemplo: os mecanismos de comunicação entre as células do sistema imunológico, o processo de inflamação, a idéia por trás das vacinas, as primeiras reações do sistema imunológico (febre e cansaço, por exemplo), a integração entre as três fases da resposta imunológica, o mecanismo que permite a identificação precisa do agente patogênico, a atuação das proteínas complementares, o processo Darwiniano de aperfeiçoamento dos receptores e o papel dos órgãos do sistema linfático.

Embora a analogia seja valorosa, existem alguns aspectos que precisam ser devidamente analisados. Alguns deles são apresentados como segue.

- O sistema imunológico garante a sobrevivência de um indivíduo, mas o termo "sobrevivência", em um sistema computacional, possui um sentido mais amplo. Garantir a "sobrevivência" de um sistema computacional geralmente implica em garantir confidencialidade, integridade, disponibilidade, responsabilidade e corretude [6].

- Algumas soluções biológicas podem não ser diretamente aplicáveis a sistemas computacionais. É importante não ignorar soluções não-biológicas, que talvez sejam mais apropriadas, como por exemplo, o uso de criptografia.

- É possível, através da analogia, herdar características indesejáveis do sistema imunológico (aspectos falhos ou suposições inapropriadas).

- O sistema imunológico não é capaz de reconhecer todos os antígenos *nonself* em um determinado instante. A diversidade é garantida por um processo aleatório na geração dos receptores.

- O processo de seleção negativa não é de todo perfeito. As células imaturas não são expostas a todas as proteínas *self*, podendo ocasionar doenças autoimunes.

- A cobertura provida pelo sistema imunológico é incompleta, ou seja, ele é vulnerável a agentes patogênicos particulares. Porém, como ca-

da indivíduo tem um sistema imunológico peculiar, nem todos são vulneráveis aos mesmos patógenos em um mesmo grau. Essa diversidade de sistemas imunológicos garante a sobrevivência da população como um todo.

Referências

[1] Cohen, F. (1987). *Computer viruses.* Computers & Security, 6:22-35.

[2] Forrest, S., Perelson, A. S., Allen, L. and Cherukuri, R. (1994). *Self-nonself discrimination in a computer.* In Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy Los Alamos, CA: IEEE Computer Society Press.

[3] Kephart, J. O. (1994). *A biologically inspired immune system for computers.* In Artificial Life IV: MIT Press.

[4] Forrest, S., Hofmeyr, S. A. and Somayaji, A. (1997). *Computer immunology.* Communications of the ACM, 40(10), 88-96.

[5] Forrest, S., Hofmeyr, S. A. and Somayaji, A. (1996). *A sense of self for UNIX processes.* In Proceedings of the 1996 IEEE Symposium on Research in Security and Privacy Los Alamitos, CA: IEEE Computer Society Press.

[6] Somayaji, A., Hofmeyr, S. A. and Forrest, S. (1997). *Principles of a computer immune system.* In Proceedings of the Second New Security Paradigms Workshop.

[7] Hofmeyr, S. A. and Forrest, S. *Immunity by Design: An Artificial Immune System.*

[8] Venema, W. (1996). *Murphy's law and computer security.* Presented at the Sixth USENIX Security Symposium (San Jose, July 1996). Disponível *online* em (agosto de 2001): ftp://ftp.porcupine.org/pub/security/murphy.ps.gz.

[9] Roitt, I., Brostoff, J. and Male, D. (1996). *Immunology*, 4th Edition. Mosby.

[10] Garfinkel, S. and Spafford, G. (1996). *Practical Unix and Internet Security*, 2nd Edition. O'Reilly and Associates, Inc.

[11] Szwarcfiter, J. e Markenzon, L. (1994). *Estruturas de Dados e Seus Algoritmos.* Segunda Edição. Editora LCT.