JOÃO PAULO ARAGÃO PEREIRA

Método de mitigação contra ataques de Negação de Serviço Distribuídos utilizando Sistemas Multiagentes

São Paulo

2014

JOÃO PAULO ARAGÃO PEREIRA

Método de mitigação contra ataques de Negação de Serviço Distribuídos utilizando Sistemas Multiagentes

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

Área de Concentração: Engenharia de Computação

Orientador: Prof. Dr. Marcos Antonio Simplicio Júnior

São Paulo

2014

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

São Paulo, 16 de Julho de 2014.

Assinatura do autor:

Assinatura do orientador:

Catalogação-na-publicação

Pereira, João Paulo Aragão
Método de mitigação contra ataques de negação de serviços distribuídos utilizando sistemas multiagentes / J.P.A. Pereira. – versão corr. -- São Paulo, 2014.
123 p.

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

1.Sistemas multiagentes 2.Informação (Segurança) I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Computação e Sistemas Digitais II.t.

DEDICATÓRIA

Dedico este trabalho à minha família: mãe, irmão, esposa, filha e sobrinhas. Dedico ao meu professor que ajudou-me de forma fraternal. Dedico a todos os amigos que ajudaram-me nesse trabalho intenso.

AGRADECIMENTOS

Agradeço a Deus primeiramente por ter dado-me a oportunidade de iniciar e finalizar este trabalho.

Agradeço à minha família pelo apoio durante todo esse tempo de estudo e pesquisa.

Agradeço muito o apoio e compreensão do Professor Dr. Marcos Simplicio durante todas as fases de pesquisa e desenvolvimento.

Agradeço ao Luis Gustavo Nardin pelo apoio na pesquisa.

“..., Deus não joga aos dados...”

Albert Einstein

RESUMO

A qualidade do serviço oferecido por Provedores do Serviço de Internet (*Internet Service Provider* - ISPs) depende diretamente da quantidade de recursos disponíveis naquele momento. Nas últimas décadas, essa qualidade tem sido afetada por frequentes e intensos ataques que consomem tais recursos, como é o caso dos ataques de Negação de Serviço Distribuídos (*Distributed Denial of Service* - DDoS). Com o objetivo de tornar a rede dos ISPs mais resiliente aos diferentes tipos de ataques DDoS, foram desenvolvidas técnicas contra tais ataques ao longo dos últimos anos. Com o objetivo de contribuir com a melhoria de tais mecanismos, esta dissertação apresenta um método autônomo reativo para detecção e mitigação de ataques DDoS, utilizando um sistema multiagentes (SMA), em redes de ISPs. A propriedade principal do método proposto é identificar padrões de tráfego característicos de um ataque, como um grande fluxo de pacotes direcionados para um serviço ou equipamento, dentro da rede do ISP. Com os agentes posicionados nas prováveis vítimas e nos pontos da rede com maior fluxo de pacotes, o processo de mitigação inicia-se automaticamente após uma quantidade de pacotes, excedente ao tráfego padrão, passar por qualquer um dos nós monitorados. Como o tráfego entrante na rede do ISP é dinâmico, seja ele legítimo ou malicioso, a utilização de agentes tende a facilitar o processo de definição da rota de ataque, conforme mostram os resultados experimentais obtido com o sistema proposto.

Palavras-chave: DDoS. Detecção e mitigação. Sistema Multiagentes.

ABSTRACT

The quality of service offered by the Internet Service Provider (ISP) depends directly on the amount of resources available at that time. In recent decades, this quality has been affected by the frequent and intense attacks that consume these resources, such as the Distributed Denial of Service (DDoS) attacks. In order to make the ISPs network more resilient to different types of DDoS attacks, techniques have been developed against such attacks over the past few years. Aiming to contribute to the improvement of such mechanisms, this dissertation presents a reactive autonomous method for detecting and mitigating DDoS attacks using a Multi-Agent system (MAS), in networks of ISPs. The main property of the proposed method is to identify characteristic traffic patterns of an attack, such as a large stream of packets directed to a service or equipment within the ISP network. With agents positioned on likely victims and at points of the network with the highest packet stream, the mitigation process starts automatically after a number of packets exceeding the traffic pattern, go through any of the monitored nodes. Since the incoming traffic on the network of any ISP is dynamic, whether legitimate or malicious, the using of agents tends to facilitate the process of defining the route of attack, as shown by the experimental results obtained with the proposed system.

Keywords: DDoS. Detection and mitigation. Multi-agent System.

LISTA DE ILUSTRAÇÕES

LISTA DE GRÁFICOS

LISTA DE TABELAS

LISTA DE ABREVIATURA E SIGLAS

ACL	*Access Control List* (Lista de Controle de Acesso)
AS	*Autonomous System* (Sistema Autônomo)
BDI	*Belief, Desire, Intention* (Crença, Desejo, Intenção)
CPU	*Central Processing Unit* (Unidade de Processamento Central)
DDoS	*Distributed Denial of Service* (Negação de Serviço Distribuído)
DMZ	*Demilitarized Zone* (Perímetro de Rede)
DNS	*Domain Name System* (Sistema de Nomes de Domínios)
DoS	*Denial of Service* (Negação de Serviço)
GN	*Graph Neuron* (Grafo Neurônio)
HSP	*Hosting Service Provider* (Provedor de Serviço de Hospedagem)
HTTP	*HyperText Transfer Protocol* (Protocolo de Transferência de HiperTexto)
ICMP	*Internet Control Message Protocol* (Protocolo de Mensagem de Controle de Internet)
IETF	*Internet Engineering Task Force* (Força Tarefa de Engenharia da Internet)
IGP	*Interior Gateway Protocol* (Protocolo de Gateway Interno)
IOS	*Internetwork Operating System* (Sistemas Operacionais de Inter-rede)
IP	*Internet Protocol* (Protocolo da Internet)
IPv4	*Internet Protocol version 4* (Protocolo da Internet versão 4)
IPv6	*Internet Protocol version 6* (Protocolo da Internet versão 6)

IPFIX	*Internet Protocol Flow Information eXport* (Informação do Fluxo do Protocolo da Internet Exportado)
IPS	*Intrusion Prevention System* (Sistema de Prevenção de Intrusão)
ISP	*Internet Service Provider* (Provedor de Serviço de Internet)
KQML	*Knowledge Query and Manipulation Language* (Linguagem de Manipulação e Consulta de Conhecimento)
LAN	*Local Area Network* (Rede de Área Local)
MAE	Metropolitan Area Exchange (Área Metropolitana de Troca)
NAP	Network Access Point (Ponto de Acesso à Rede)
OSI	*Open Systems Interconnection* (Interconexão de Sistemas Abertos)
OSPF	*Open Shortest Path First* (Primeiro Caminho Aberto Mais Curto)
PPG	Planejamento Parcial Global
PRS	*Procedural Reasoning System* (Sistema de Raciocínio Procedural)
RFC	*Request for Comment* (Requisição de Comentário)
SMA	Sistema Multiagentes
TCP	*Transmission Control Protocol* (Protocolo de Controle de Transmissão)
TMS	*Threat Management System* (Sistema de Gerenciamento de Ameaças)
TTL	*Time To Live* (Tempo de Vida)
UDP	*User Datagram Protocol* (Protocolo de Datagrama de Usuário)
VoIP	*Voice over IP* (Voz sobre IP)
VPN	*Virtual Private Network* (Rede Privada Virtual)

SUMÁRIO

1 INTRODUÇÃO

A sociedade atual está cada vez mais dependente dos serviços da Internet, cuja qualidade depende da velocidade dos enlaces e contínua disponibilidade dos dados oferecidos. Essa dependência é crescente para as relações pessoais e comerciais, de consumo e trabalho, entre outras, principalmente após o acesso aos meios digitais de comunicação ter se tornado tão comum. De fato, usuários e clientes de diversos sistemas são estimulados a realizar a maioria de seus negócios através da Internet, tornando-se raro encontrar computadores, utilizados no âmbito comercial ou acadêmico, que não tenham a capacidade de acesso à rede mundial de computadores.

O acesso à Internet por usuários finais geralmente é realizado indiretamente, por meio de ISPs (*Internet Service Providers* – Provedores do Serviço de Internet), devido ao custo de manter a infraestrutura de rede necessária. ISPs, em sua maioria, são unidades de negócio de operadoras de infraestrutura, como, por exemplo, as empresas telefônicas.

Devido ao consumo de recursos de rede oferecidos ser constante, qualquer indisponibilidade gerada, tanto pelo ISP quanto por qualquer fornecedor de serviço, pode gerar prejuízos incalculáveis. Portanto, a infraestrutura da rede do ISP deve ser robusta e segura para acomodar a crescente demanda de seus clientes e preservar os serviços que são oferecidos através da rede.

Apesar de diversos esforços dos ISPs, nas últimas décadas ataques de Negação de Serviço (*Denial of Service* - DoS) e sua variante distribuída (*Distributed Denial of Service* - DDoS) causaram grandes prejuízos a entidades que dependem

fortemente da Internet em seus negócios. Por esta razão, o ataque DDoS é atualmente considerado uma das ameaças mais sérias para a infraestrutura de serviços críticos de um ISP (ARBOR, 2014). Tais ataques utilizam um grande volume de tráfego, de diferentes origens, para colapsar os serviços oferecidos pelo alvo do ataque. Eles são difíceis de serem bloqueados, devido a fatores como (MIRKOVIC et al., 2004): o grande número de computadores remotos usados no ataque, os quais costumam estar sob diferentes domínios administrativos; a habilidade dos atacantes em usar endereços IP (*Internet Protocol* - Protocolo da Internet) de origem falsa; e a dificuldade para a vítima distinguir o tráfego legítimo do malicioso. Atualmente é difícil encontrar sistemas de defesa contra ataques DDoS, sejam eles comerciais ou ainda em desenvolvimento, que forneçam qualquer garantia da continuidade de serviço para clientes legítimos (MIRKOVIC et al., 2004).

Nesse contexto, é crescente a pesquisa de sistemas distribuídos e concorrentes, sendo que novos paradigmas são criados a fim de estabelecer modelos teóricos que melhor refletem o processo de interação constante entre sistemas computacionais (LUCK; MCBURNEY; GONZALES-PALACIOS, 2006). Muitas destas pesquisas são multidisciplinares, unindo áreas tão diversas como Inteligência Artificial e Redes de Computadores. O atual cenário da vasta exposição tecnológica mostra que as novas aplicações não são mais centralizadas nem distribuídas e gerenciadas por apenas uma organização. Por exemplo, a rede da Internet está dispersa em uma estrutura física de conectividade, com switches, roteadores e enlaces. Estes dispositivos estão organizados em redes conhecidas como Sistemas Autônomos (*Autonomous Systems* - AS), onde cada AS está sob o

controle administrativo de uma mesma organização, como um ISP (WILLINGER; ALDERSON; DOYLE, 2009).

1.1. Motivação

Conforme pesquisa realizada em (ARBOR, 2014), houve uma evolução em volume dos ataques DDoS nos últimos anos: de 10 Gbps (Giga bits por segundo) em 2005 para 309 Gbps em 2013. De todos os participantes dessa pesquisa, realizada em 2014, 40% eram ISPs. A Tabela 1 mostra o percentual das ameaças operacionais mais significativas para ISPs, empresas, instituições de pesquisa e governamentais identificados nessa pesquisa.

Tabela 1 – Resumo das ameaças operacionais mais significativas (ARBOR, 2014).

Instituições	Ameaça operacional mais significativa	%
Tier 1, Tier 2, Tier 3, ISPs, Instituições educacionais de pesquisa, Governos	Ataques DDoS para clientes	64
	Interrupção na infraestrutura por falhas	55
	Ataques DDoS para infraestrutura	46
	Ataques DDoS para serviços	44
	Saturação de banda	44
	Computadores infectados na rede do ISP	34
	Novas vulnerabilidades	6

Os ataques DDoS são difíceis de serem bloqueados. Isso se deve em parte à dificuldade de implementar sistemas de defesa distribuídos em todas as regiões da rede do ISP, desde a borda até a provável vítima, dada a complexidade dessas redes (MIRKOVIC et al., 2004).

Devido ao constante crescimento da frequência dos ataques DDoS contra ISPs, sites de empresas, instituições financeiras, governamentais e de pesquisa, muitas técnicas de detecção e mitigação foram desenvolvidas. Há técnicas de

defesa que são baseadas em equações matemáticas (MANN; KUMAR, 2011), técnicas se baseiam no índice de confiança dos pacotes (BHATTACHARJEE; RAGHAVAN; SANAND, 2011) e também técnicas que utilizam Sistemas Multiagentes (SMA) (WOOLDRIDGE, 2009) para facilitar a detecção e mitigação (JUNEJA; CHAWLA; SINGH, 2009) e (AKYAZI; UYAR, 2008). Estas últimas são de especial interesse para aplicação em ambientes altamente dinâmicos e dispersos em topologias complexas como é o caso das redes dos ISPs e, consequentemente, na Internet. Nesse contexto, um SMA é aplicado para a rede ser controlada e monitorada de forma ubíqua.

É importante ressaltar que o custo da implantação de um sistema de defesa contra ataques DDoS justifica-se também em termos financeiros, em especial quando se compara ao custo de tais soluções com os prejuízos ocasionados por interrupções no fornecimento dos serviços ofertados pelo ISP. Como a maioria dos serviços ofertados por um ISP é cobrada, o cliente que o utiliza para tráfego dos dados de sua empresa poderia tentar atribuir todo o prejuízo decorrente da indisponibilidade do serviço contratado ao ISP.

1.2. Objetivos

A questão chave desta dissertação é, através da combinação de fatores relevantes das áreas de Rede de Computadores e Inteligência Artificial, prover uma solução capaz de bloquear ataques DDoS volumétricos. No cenário alvo é utilizada como base uma área de pesquisa específica da Inteligência Artificial Distribuída, chamada de Sistemas Multiagentes (SMA).

Dessa forma, a meta deste trabalho é projetar, implementar e testar um sistema composto de múltiplos agentes, de modo que a ação local autônoma de cada agente, aliada à cooperação entre eles, permita a detecção, mitigação e bloqueio de ataques DDoS. O SMA desenvolvido, denominado Arquitena, consiste em agentes que representam os equipamentos de rede e se comunicam com os mesmos, sejam eles roteadores, *firewalls*, *switches* ou servidores. O sistema Arquitena tem característica autônoma reativa frente aos diferentes tipos de ataques DDoS.

Esse mecanismo deve ser rápido o suficiente para atuar antes mesmo que qualquer equipamento sob ataque entre em colapso devido ao grande número de pacotes fluindo por ele. O sistema, além de veloz na detecção, não deve onerar excessivamente o processamento dos equipamentos responsáveis pelo roteamento do tráfego durante a mitigação. Assim, o sistema deve apresentar baixas taxas de falso-positivos e falso-negativos.

1.3. Justificativa

Como os ataques de DDoS tornaram-se frequentes e volumosos e permanecem difíceis de serem bloqueados, espera-se um sistema de defesa que possa ser capaz de mitigar e bloquear este tipo de ataque na rede de qualquer ISP. A utilização de um Sistema Multiagentes para mitigar este tipo de ataque pode ser eficaz, pois devido às características do agente como a deliberação nas ações, persistência no cumprimento das metas e sociabilidade de interagir com outros agentes do sistema, pode contribuir com a mitigação.

Devido a rede do ISP ter um grande número de equipamentos e ser um ambiente altamente dinâmico, espera-se que o SMA utilizado possa monitorar todo esse ambiente complexo.

1.4. Métodos

O método adotado no desenvolvimento do presente trabalho é a Pesquisa Aplicada com base na abordagem hipotética-dedutiva, utilizando referências disponíveis na literatura especializada para a definição do problema, especificação de hipóteses e análise de tais hipóteses. A solução proposta é especificada e implementada por meio de um protótipo que permita a validação das hipóteses que serviram como base para o projeto da solução. Considerando o inevitável estado de defesa em que ISPs devem permanecer, requer-se o estudo e desenvolvimento de técnicas anti-intrusão. Essas técnicas, ao mesmo tempo em que não podem onerar o processamento do tráfego de rede, devem proporcionar um sistema de proteção contra os ataques oriundos da rede interna ou externa. Conforme os objetivos propostos, o sistema multiagentes Arquitena foi desenvolvido para detectar e mitigar ataques DDoS, com prévio mapeamento virtual de toda a rede do ISP. Assim, cada agente é construído com base no equipamento a ser monitorado na rede, apresentando as mesmas conexões entre eles.

A primeira fase desse trabalho consistiu no estudo direcionado em quatro etapas. A primeira consistiu na análise das diversas funcionalidades de SMAs, a fim de compreender as vantagens possíveis de sua utilização em uma rede com muitos equipamentos. A segunda etapa concentrou-se no estudo de como um SMA poderia ser aplicado para detectar e bloquear ataques DDoS dentro da rede de um ISP, juntamente com a análise bibliográfica de (MIRKOVIC et al., 2004) e de artigos

referentes a ataques DDoS. A terceira etapa concentrou-se no projeto da solução com o objetivo de detectar e mitigar ataques DDoS usando um SMA, bem como no estudo da plataforma de desenvolvimento de sistemas multiagentes Jason (BORDINI; HÜBNER; WOOLDRIDGE, 2007), que permitisse sua implementação e avaliação em ambiente de simulação com suporte à comunicação entre o agente e o respectivo equipamento de rede por ele monitorado. Finalmente, a quarta etapa consistiu em comparar os resultados alcançados com trabalhos relacionados, a fim de mensurar as vantagens e desvantagens do sistema proposto em relação às soluções existentes.

1.5. Organização do documento

O restante deste documento está organizado em capítulos. No capítulo 2 são descritos os conceitos básicos para um bom entendimento do texto, englobando tanto tópicos de Redes de Computadores como de Inteligência Artificial. No capítulo 3 são discutidos os trabalhos relacionados à presente pesquisa, envolvendo sistemas baseados ou não em múltiplos agentes coordenados. No capítulo 4 é apresentado em detalhes o sistema Arquitena. No capítulo 5 encontram-se os resultados finais e a abordagem de testes adotada para a verificação da eficácia do sistema proposto. Finalmente, no capítulo 6 são apresentadas as conclusões do trabalho e trabalhos futuros com base no sistema proposto.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são detalhados alguns dos tópicos mais relevantes pertinentes à presente pesquisa. Como este é um trabalho multidisciplinar, há subseções envolvendo Inteligência Artificial Distribuída que envolve o conceito de Sistemas Multiagentes, e também Redes de Computadores, mais especificamente Segurança da Informação.

2.1. Agentes e Sistemas Multiagentes

Um agente é um sistema de computador com duas importantes habilidades: decidir de forma autônoma o que precisa fazer para alcançar os seus objetivos; e ser capaz de interagir com outros agentes (WOOLDRIDGE, 2009). Dessa maneira, o comportamento do agente é resultado do conhecimento do ambiente, seja ele limitado ou amplo, da comunicação com outros agentes e da percepção do ambiente.

O agente inserido em um ambiente pode percebê-lo através dos seus sensores. Cada agente deve ter um repositório de possíveis ações que pode executar, mas os atuadores só agem a partir do momento em que a decisão é tomada (BORDINI; HÜBNER; WOOLDRIDGE, 2007). A função de um agente pode variar conforme o ambiente no qual está inserido: ele pode ser considerado um assistente do usuário com algumas funções relevantes ao trabalho que lhe foi designado, monitorar eventos e processos. Algumas propriedades que podem estar presentes em cada agente, ajudam a defini-lo, conforme lista a seguir:

- **Mobilidade:** propriedade básica do agente que o caracteriza na capacidade de ser transportado de um computador a outro.

- **Comunicabilidade:** propriedade pela qual o agente tem a capacidade de trocar mensagens com outros agentes, provendo as informações pertinentes sobre o ambiente no qual o agente está inserido (FRANKLIN; GRAESSER, 1996).
- **Representatividade:** determina se o agente tem uma noção robusta sobre o contexto no qual está inserido, podendo ser vista como uma extensão dos atributos, desejos, crenças e propósitos da entidade representada (WOOLDRIDGE, 2009).
- **Cooperação:** capacidade dos agentes, que compartilham o mesmo ambiente competindo por recursos escassos, em alcançar as metas a eles delegadas com eficiência e de forma cooperativa (LITHGOW-SMITH; TAMMA; WOOLDRIDGE, 2011).
- **Aprendizagem:** um agente pode acumular conhecimento baseado em experiências anteriores e, consequentemente, executar uma tarefa com maior eficiência do que em execuções prévias (HAMMOUD; MAAMRI; SAHNOUN, 2011).
- **Planejamento:** o agente tem a habilidade de discernir entre diferentes ações, dentro de uma biblioteca de ações possíveis, com o propósito de alcançar suas metas (WEERDT; MORS; WITTEVEEN, 2005).

Embora nesta subseção sejam descritas as principais características, funções e propriedades dos agentes, há uma definição separada de agentes inteligentes, já que estes últimos apresentam características adicionais. Especificamente, a inteligência em um agente pode habilitá-lo a lidar efetivamente com ambiguidades durante a tomada de uma decisão. A seguir são detalhadas as características dos

agentes inteligentes (PADGHAM; WINIKOFF, 2006) que são utilizadas para desenvolvimento do sistema Arquitena:

- **Situado:** o agente existe em um ambiente dinâmico e imprevisível.
- **Autônomo:** o agente é independente e toma suas próprias decisões.
- **Reativo:** o agente é capaz de responder, em tempo hábil, às mudanças em seu ambiente.
- **Proativo:** o agente persegue persistentemente suas metas, mesmo que haja falhas no decorrer das tentativas de alcançá-las.
- **Robusto:** o agente é capaz de se recuperar de falhas**.**
- **Social:** o agente tem a habilidade de interagir com outros agentes.

A proatividade e a reatividade devem ser balanceadas nas características de um agente, a fim dessas características não entrarem em conflito. Se o agente é muito reativo, então ele está constantemente ajustando seus planos e não alcança suas metas. Contudo, se o agente não é suficientemente reativo, então ele desperdiça tempo tentando seguir planos que não são relevantes (PADGHAM; WINIKOFF, 2006).

2.1.1. Modelo Crença-Desejo-Intenção

O modelo Crença-Desejo-Intenção (*Belief, Desire, Intention* - BDI) (RAO; GEORGEFF, 1992) refere-se a programas de computador como se eles tivessem estados mentais. Um sistema BDI tem analogia com as crenças, desejos e intenções que os humanos tem. As crenças são informações que o agente tem do ambiente no qual está inserido. Os desejos são todos os possíveis estados de coisas que o agente poderia realizar. Ter o desejo não implica que o agente irá agir para atingi-lo.

As intenções são os estados de coisas em que o agente tem decidido trabalhar. Caso um agente tenha escolhido um dos possíveis desejos, esse desejo especificamente é uma intenção (WOOLDRIDGE, 2009). O Raciocínio Prático é o modo pelo qual o agente alcança suas metas utilizando suas crenças, desejos e intenções dividindo-se em duas atividades (WOOLDRIDGE, 2009):

- **Deliberação:** O processo de deliberação resulta em um agente adotar uma intenção. Espera-se que o agente realize alguma tentativa de alcançar a intenção escolhida, e persista nesta tarefa. Assim, caso uma intenção tenha sido escolhida, o raciocínio prático futuro fica restrito. As intenções estão fortemente relacionadas com as crenças, pois o agente pode acreditar que sua intenção pode ser alcançada com sucesso ou falha.
- **Raciocínio meios fins ou planejamento:** É o processo de como o agente alcança a intenção que tenha escolhido utilizando os meios disponíveis, isto é, as ações que o agente pode executar no ambiente.

2.1.2. AgentSpeak (L)

AgentSpeak(L) (RAO, 1996) é uma linguagem de programação baseada em uma linguagem de primeira ordem com eventos e ações, definindo o comportamento do agente, isto é, a interação dele com o ambiente. Através dessa linguagem há uma formalização dos agentes BDI. Dessa forma, pode ser considerada como uma abstração do sistema BDI implementado, conhecido como Sistema de Raciocínio Procedural (*Procedural Reasoning System* - PRS) (RAO, 1996). As crenças, desejos e intenções do agente não são explicitamente representados como fórmulas. Ao invés disso, os programadores podem atribuir essas noções para o agente com

AgentSpeak(L). O estado atual do agente, que é o modelo dele próprio, de seu ambiente e de outros agentes, pode ser considerado como seu estado de crenças atuais. Os desejos podem ser considerados os estados que o agente quer alcançar através dos estímulos internos e externos. As intenções são programas adotados para satisfazer os estímulos. Em resumo, um agente pode ser especificado por um conjunto de bases de crenças e um conjunto de planos.

A linguagem interpretada pelo Jason (BORDINI; HÜBNER; WOOLDRIDGE, 2007) é uma extensão do AgentSpeak(L), que é baseada na arquitetura BDI. Ela implementa a semântica operacional dessa linguagem e fornece a plataforma para o desenvolvimento de Sistemas Multiagentes, com recursos customizáveis. A arquitetura de agentes utilizada é o PRS. Assim, os componentes da arquitetura são a base de crenças e as metas dos agentes que são alcançadas pela execução dos planos. Com a utilização do Jason, embora a abordagem de comunicação seja baseada do Ato de Fala (AUSTIN, 1962) e não no fornecimento de uma longa lista de rótulos performativos, podem-se escolher rótulos de uma pequena lista disponível, derivada do KQML (*Knowledge Query and Manipulation Language* - Linguagem de Manipulação e Consulta de Conhecimento) (FININ; LABROU; MAYFIELD, 1995). A vantagem de se utilizar essa abordagem é evitar confusão entre os rótulos.

2.1.3. Ambientes de Sistemas Multiagentes

Basicamente, um Sistema Multiagentes é um sistema computacional em que diversos agentes inteligentes interagem entre si e com o ambiente para atingir alguma meta. O ambiente de um SMA pode ser classificado como (WOOLDRIDGE, 2009):

- **Acessível x Inacessível:** Um ambiente acessível é aquele em que os agentes podem obter informações completas e atualizadas sobre o estado do ambiente. Ao contrário, nos ambientes inacessíveis os agentes tem informações limitadas do ambiente no qual estão inseridos.
- **Determinístico x não determinístico:** Um ambiente determinístico é aquele em que qualquer ação do agente tem um efeito garantido, não havendo incerteza sobre o estado do ambiente após a execução da ação. O ambiente não determinístico captura o fato dos agentes terem uma limitada esfera de influência e o resultado da execução das ações não ser unicamente determinado. Nesse caso, uma reação em cadeia pode ser disparada como resultado da execução de uma única ação.
- **Estático x dinâmico:** Um ambiente estático é aquele que permanece inalterado, exceto pela execução das ações do agente. Em contraste, o ambiente dinâmico é aquele em que há outros processos operando além do SMA, e por isso as mudanças dos caminhos superam o controle do agente.
- **Discreto x contínuo:** Um ambiente é discreto se o número de ações e percepções é finito e fixo, ao contrário do ambiente contínuo.

2.2. Segurança de Redes e Ataques de DDoS

Ataques DDoS são voltados especificamente a afetar a disponibilidade de dados e serviços de uma rede. Assim, eles diferem consideravelmente de crimes cibernéticos voltados à subtração ou exploração de dados, desconfiguração de equipamentos, alteração de páginas Web ou até mesmo invasões em instituições

financeiras para transferir dinheiro de maneira ilícita (MIRKOVIC et al., 2004), embora possa ser utilizado para acobertar tais atividades maliciosas.

Para um atacante conseguir executar um ataque DDoS, ele precisa ou explorar vulnerabilidades inerentes ao alvo ou enviar um grande número de mensagens semelhantes às legítimas para o alvo. No primeiro tipo de ataque, as mensagens podem, por exemplo, fazer a aplicação da vítima entrar em um laço infinito, causando lentidão ou parada total do serviço. Tais vulnerabilidades podem ser exploradas nos sistemas operacionais, protocolos de rede e programas de aplicação (MIRKOVIC et al., 2004). O segundo tipo de ataque trabalha enviando um grande número de mensagens que consomem os recursos do alvo. Caso essas mensagens sejam complexas de serem interpretadas pela vítima, o tempo necessário para processamento pode ser longo e dessa forma, novas mensagens (muitas delas de usuários legítimos) não podem ser tratadas. Pode-se também atacar a interface de rede da vítima, excedendo sua capacidade total. Por exemplo, se uma interface de rede pode tratar apenas 1Gbps, o atacante pode enviar um tráfego de pacotes IP de 10Gbps. Outra maneira de efetuar o ataque DDoS é utilizando máquinas comprometidas e a partir destas máquinas efetuar ataques volumétricos com endereços IPs forjados. Assim, o atacante realiza uma intrusão em massa instalando ferramentas para obter acesso às máquinas. Após essa fase, pode instalar softwares DDoS em cada uma das máquinas comprometidas. A última fase é lançar ataques de inundações com origens distintas, muitas vezes forjadas.

Qualquer que seja o tipo de ataque DDoS, o prejuízo é sentido não apenas pela vítima mas também por outros usuários da rede que estejam compartilhando o

mesmo ISP que está sob ataque DDoS. Alguns exemplos de prejuízos causados por ataques DDoS incluem (MIRKOVIC et al., 2004):

- Perda direta de renda para empresas que usam principal ou exclusivamente a Internet para execução de suas vendas.
- Perda direta de renda para novos sites e mecanismos de pesquisa que dependem de visitas de usuários a páginas Web.
- Parada de fornecimento de informações cruciais para usuários de empresas públicas de água, esgoto ou luz.
- Perda de credibilidade para ISPs que oferecem serviços de DNS *Domain Name System* (Sistema de Nomes de Domínios), pois todos os navegadores Web e muitas outras aplicações dependem desse serviço.
- Perda de credibilidade de instituições financeiras, que por um momento mantiveram-se indisponíveis para oferecer seus serviços pela Internet.

Há pelo menos quatro motivos comuns que motivam o engajamento de um indivíduo ou um grupo para iniciar um ataque DDoS (CERT, 2014):

- **Demonstração de poder:** mostrar para uma empresa que seu sistema de defesa é vulnerável e, em consequência dessa fragilidade, uma extorsão pode ser gerada pelo atacante.
- **Motivação comercial:** uma empresa pode tornar inacessível o site de uma concorrente ou contratar um atacante para fazê-lo.
- **Prestígio:** um atacante pode demonstrar a seus pares que é capaz de tirar do ar o serviço de uma entidade.

- **Motivação ideológica:** negar o serviço HTTP (*HyperText Transfer Protocol* - Protocolo de Transferência de HiperTexto), por exemplo, de uma entidade que divulgue conteúdo contrário à opinião do atacante.

Segundo a taxonomia dos ataques DDoS (RIOREY, 2014), dos 25 tipos de ataques, 15 são ataques de inundação de pacotes:

1. *SYN Flood* (baseado em TCP) – servidor vítima recebe altas taxas de requisições SYN contendo endereços IPs forjados.
2. *SYN-ACK Flood* (baseado em TCP) – servidor vítima recebe altas taxas de requisições SYN-ACK contendo endereços IPs forjados.
3. *ACK & PUSH ACK Flood* (baseado em TCP) – a vítima recebe pacotes ACK forjados em uma alta taxa de envio que pertencem a qualquer sessão dentro da lista de conexão do servidor.
4. *Fragmented ACK* (baseado em TCP) – usa 1500 bytes de tamanho de pacote para consumir grande quantidade da banda.
5. *RST* ou *FIN Flood* (baseado em TCP) – um servidor vítima recebe pacotes RST ou FIN forjados em uma alta taxa de envio que não pertencem a qualquer sessão dentro da base de dados do servidor.
6. *Synonymous IP* (baseado em TCP) - a vítima recebe pacotes TCP-SYN forjados em uma alta taxa de envio que tem a informação da vítima especificada como ambos: IP de origem e de destino.
7. *UDP Flood* (baseado em UDP (*User Datagram Protocol* - Protocolo de Datagrama de Usuário)) - servidor vítima recebe pacotes UDP forjados em altas taxas e com uma grande extensão de IPs de origem.

8. *UDP Fragmentation* (baseado em UDP) – variação do *UDP flood* sendo que o atacante usa grandes e poucos pacotes (1500 bytes).
9. *DNS Flood* (baseado em UDP) – um servidor DNS vítima recebe pacotes de requisição DNS válidos, mas forjados em uma alta taxa de uma grande extensão de IPs de origem.
10. *VoIP Flood* (baseado em UDP) – servidor *VoIP* recebe pacotes *VoIP* em uma alta taxa de uma grande extensão de IPs de origem.
11. *Media Data Flood* (baseado em UDP) – variação de *UDP flood*, mas pacotes tem a forma de qualquer dado de mídia.
12. *Non-Spoofed UDP Flood* (baseado em UDP) - servidor vítima recebe pacotes UDP não forjados em altas taxas.
13. *ICMP Flood* (baseado em ICMP (*Internet Control Message Protocol* (Protocolo de Mensagem de Controle de Internet))) - servidor vítima recebe pacotes ICMP forjados em altas taxas e com uma grande extensão de IPs de origem.
14. *ICMP Fragmentation* (baseado em ICMP) – variação do *ICMP fl*ood sendo que a vítima recebe grandes pacotes ICMP fragmentados (1500 bytes) e esses pacotes não podem ser remontados.
15. *Ping Flood* (baseado em ICMP) - servidor vítima recebe ping (*ICMP echo requests*) forjados em altas taxas e com uma grande extensão de IPs de origem.

2.2.1. Estratégias comuns de defesa contra ataques DDoS

Um ataque DDoS pode ser difícil de ser detectado e depende de alguns fatores como o local onde o sistema foi implantado, a extensão da implantação e a

velocidade desejada de detecção. Em geral, a sensibilidade e precisão de um detector de ataque se deterioram proporcionalmente à sua distância da vítima, isto é, quanto mais perto da vítima maior a precisão obtida.

Segundo a pesquisa em (ARBOR, 2014), de todos os ataques monitorados no mundo em 2013, 61% foram ataques DDoS volumétricos, sendo esta ainda a forma mais comum de ataque. Contudo, há pelo menos 25 tipos de ataques DDoS (RIOREY, 2014) classificados por protocolos. Esta grande variedade de ataques levou à criação de diversos mecanismos destinados à defesa contra os mesmos, sendo que as estratégias mais utilizadas são baseadas em: assinatura, anomalia e mau comportamento (MIRKOVIC et al., 2004). A detecção baseada em assinatura constrói uma base de dados com as características observadas, e todos os pacotes de entrada são comparados com essa base. A detecção por anomalia tenta modelar o tráfego legítimo e sinaliza caso o tráfego corrente viole o modelado. Por exemplo, um comportamento anômalo pode ser detectado pelo fato de um tráfego alcançar 100 Mbps ao invés dos usuais 10 Mbps. Já a terceira técnica consiste na modelagem do que pode ser considerado um “mau comportamento”, e posterior observação de sua ocorrência. Esse modelo pode ter como componentes a assinatura de um ataque e definição do que constitui um tráfego legítimo.

Como parte de um mecanismo de defesa, a resposta ao ataque deve ser rápida. Sendo assim, há basicamente três maneiras para melhorar a situação de usuários legítimos e mitigar o efeito do ataque DDoS. A primeira consiste no policiamento do tráfego, usando um filtro para impedir a passagem de pacotes considerados suspeitos, ou aplicando uma taxa limite aceitável de pacotes suspeitos (MIRKOVIC et al., 2004). A segunda forma se baseia em construir a trajetória

reversa do ataque DDoS, com o objetivo de identificar o atacante que controla as máquinas realizadoras do ataque. A terceira técnica se resume a incluir mecanismos para verificar a legitimidade do cliente e, dessa forma, fornecer um serviço aceitável durante um ataque DDoS.

O processo de defesa contra um ataque DDoS pode consistir de alguns mecanismos bem comuns para ISPs (MIRKOVIC et al., 2004):

- **Estudo prognóstico de vulnerabilidades:** manter atualizados os softwares e sistemas operacionais de roteadores, *switches*, *firewalls* e IPS (*Intrusion Prevention Systems* - Sistemas de Prevenção de Intrusão).
- **Filtro de pacotes maliciosos:** ter no mínimo um sistema para contabilizar ou inspecionar pacotes, a fim de efetuar o reconhecimento de pacotes que podem causar algum dano.
- **Validação da origem:** tentar validar o cliente que efetua a requisição de um determinado serviço, através de uma lista de usuários legítimos ou por meio da tentativa de obter resposta desse possível cliente a requisições de rede.
- **Prova de trabalho:** abordagem para proteger protocolos assimétricos, isto é, que consomem mais recursos no lado do servidor do que no lado do cliente. Consiste em inserir um passo a mais no lado do cliente a fim de gastar recursos suficientes para sua comunicação antes do servidor gastar seus próprios recursos.
- **Ocultação:** proteger o servidor atrás de *firewalls*, não divulgando abertamente a localização do servidor que provê o serviço. Em casos

extremos, apenas clientes autenticados são autorizados a utilizar o serviço.

- **Super-provisionamento:** abordagem simples, porém cara, que consiste na instalação de mais recursos no sistema de forma a atender tanto as requisições reais como as maliciosas. O mais comum é ter mais largura de banda de entrada e um conjunto de servidores atrás de um balanceador de carga.

Apesar desses mecanismos ajudarem a reduzir o efeito de ataques DDoS, há alguns desafios a serem enfrentados por sistemas de defesa contra tais ataques (MIRKOVIC et al., 2004):

- Fácil disponibilidade de ferramentas para efetuar ataques DDoS.
- Similaridade do tráfego malicioso com o legítimo.
- Possível modificação do endereço IP de origem.
- Alto volume de tráfego.
- Numerosas máquinas comprometidas no mundo que podem ser usadas para ataques DDoS.
- Topologia da Internet com pontos centrais desprotegidos.

Quando um ataque DDoS ocorre, muitas empresas e ISPs podem ter a falsa sensação de segurança devido à existência de soluções de IPS ou *firewalls* implantados em suas redes. Entretanto, como nos últimos anos os ataques DDoS tem sido dominados por ataques volumosos e sofisticados, se faz necessário implantar um sistema inteligente de mitigação DDoS, a fim de garantir a disponibilidade dos dados e serviços (ARBOR, 2014).

2.2.2. ISPs e ASs

Como mencionado no capítulo 1, a Internet é constituída por inúmeros Sistemas Autônomos (*Autonomous Systems* - ASs). A definição clássica de um AS é de um conjunto de roteadores sob uma única administração, usando um IGP (*Interior Gateway Protocol* - Protocolo de *Gateway* Interno) e, métricas em comum para determinar como rotear pacotes em seu interior e utilizando um Protocolo de Roteamento inter-AS para determinar como rotear pacotes para outros ASs. É comum para um único AS utilizar muitos IGPs e conjuntos diferentes de métricas. Mesmo nessa condição, a administração de um AS aparece para outros ASs como tendo um único plano de roteamento interno e apresenta um desenho consistente dos destinos que são alcançáveis através dele.

Um ISP, além de fornecer acesso à Internet, pode oferecer serviços como voz sobre IP (VoIP (*Voice over IP* - Voz sobre IP)), VPN (*Virtual Private Network* – Rede Privada Virtual), armazenamento de sites, entre outros. Por isso, um ISP pode ser considerado a maior unidade de negócio das operadoras de infraestrutura.

Segundo a Requisição de Comentário (*Request for Comment* - RFC) de número 3013 – Procedimentos e Serviços Recomendados para Segurança de ISP (RFC 3013, 2014), alguns mecanismos de segurança são recomendados para que um ISP evite ataques DoS, como:

- **Filtro de rotas:** ISPs devem filtrar os anúncios de roteamento que aprendem, a fim de reduzir os riscos de colocar excessiva carga no roteamento em parte da rede.

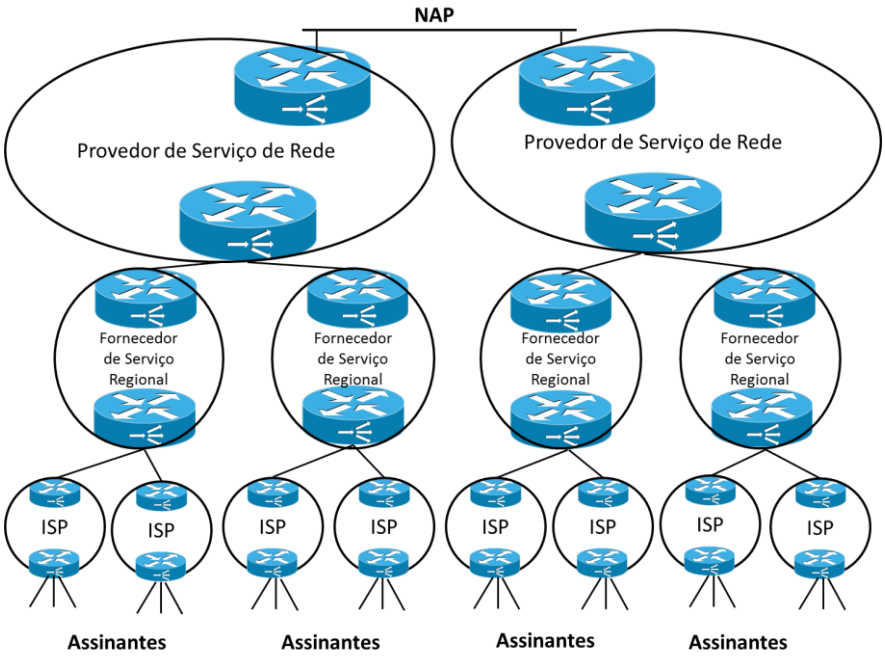
- ***Broadcast* dirigido:** Os roteadores conectados devem ser configurados para não permitir *broadcasts* dirigidos, pois muitos ataques DoS podem fazer uso da replicação de pacotes enviados para todos os nós conectados em uma mesma sub-rede.

Entretanto, na RFC 3013 não existem mecanismos de segurança focados somente em ataques DDoS, sejam eles oriundos de vulnerabilidades de protocolos ou de inundação de pacotes.

Embora a Internet esteja crescendo fora do padrão de apenas um *backbone*, ela ainda mantém uma estrutura hierárquica. A figura 1 é uma macro topologia com pequenos ISPs, chamados de IPS locais, conectados aos ISPs regionais. Os ISPs regionais cobrem áreas como estados de um país ou vários estados adjacentes. Os ISPs maiores tem os ISPs regionais como clientes e cobrem áreas nacionais, e são chamados de Provedores de Serviço de Rede ou Área Metropolitana de Troca (*Metropolitan Area Exchange* - MAE).

Os ISPs que estão no mesmo nível geralmente não utilizam um ao outro para transmitir suas mensagens. Através de um NAP (*Network Access Point* - Ponto de Acesso à Rede) ou MAE, um ISP pode conectar-se com outro ISP em arranjos ponto a ponto. Os NAPs são os elementos chaves do b*ackbone* da Internet, pois as conexões dentro deles determinam como o tráfego da Internet é roteado. Por esta razão, são considerados os maiores pontos de congestionamento da Internet (DOYLE; CARROLL, 2005).

Figura 1 – Arquitetura da Internet (DOYLE; CARROLL, 2005).



3 TRABALHOS RELACIONADOS

Esta seção descreve trabalhos relevantes que estão envolvidos com sistemas de defesa contra ataques DDoS. Durante a descrição, são consideradas separadamente as técnicas baseadas em SMA daquelas que não o são. Complementar a essa discussão, na subseção 3.3 são descritos alguns equipamentos utilizados comercialmente para mitigação de ataques DDoS.

3.1. Detecção e mitigação de ataques DDoS, sem utilização de agentes

Nesta seção são discutidas soluções que utilizam estatística, recursos de rede, entre outras técnicas, porém não são baseadas em SMA.

3.1.1. DefCOM

DefCOM (*Defensive Cooperation Overlay Mesh* – Malha de Sobreposição para Cooperação Defensiva) (MIRKOVIC; ROBINSON; REIHER, 2003) é um sistema de defesa que consiste de nós (roteadores) em uma rede ponto a ponto, os quais se comunicam para alcançar uma defesa cooperativa e dinâmica. Nesse sistema há três tipos de nós:

- **Gerador de alertas:** detecta ataque e informa outros nós. Implantados na borda ou na vítima.
- **Classificador:** distingue tráfego legítimo do malicioso e encaminha os pacotes legítimos com uma marca, além de marcar pacotes suspeitos e limitar a sua taxa de transmissão. Implantados na parte central da rede.
- **Limitador de taxa:** limita a taxa de todo o tráfego para a vítima e dá maior prioridade para o tráfego legítimo. Implantados na parte central da rede.

No DefCOM, os nós desenvolvem as ações distributivamente com uma detecção mais precisa no lado da vítima, uma classificação do tráfego no lado da origem e limitação da taxa na parte central da rede. A operação do DefCOM depende de uma implantação nos roteadores centrais da rede. Desta forma, um ponto negativo dessa solução é que o comprometimento de qualquer nó envolvido no sistema de detecção por um atacante pode afetar o funcionamento de todo o processo. Além disso, esse arcabouço não contém um mecanismo de resposta aos ataques DDoS: ele foi projetado para ser acoplado a sistemas de defesa existentes e, assim, facilitar a ação colaborativa.

Quando comparada ao sistema Arquitena, uma semelhança importante é que o nó que gera o alerta de ataque DDoS propaga o sinal para outros nós conectados diretamente. Por outro lado, o sistema Arquitena utiliza um sistema de classificação baseado na quantidade de pacotes e não efetua qualquer marcação de pacotes. Os agentes monitoram a maioria dos equipamentos e dinamicamente os utilizam para calcular os limites a serem definidos, sendo esses os classificadores dos ataques DDoS. Além disso, o Arquitena contempla a detecção e também o processo de mitigação contra ataques DDoS. Outra vantagem do sistema Arquitena é que ele utiliza o processamento do equipamento de rede apenas para a coleta de dados. Assim, a maior parte do processamento relacionado à mitigação é realizado pelos agentes do sistema, enquanto o DefCOM utiliza os equipamentos de rede para realizar a classificação e detecção.

3.1.2. Path Attestation Scheme

O *Path Attestation Scheme* (Esquema de Comprovação de Caminho) (BHATTACHARJEE; RAGHAVAN; SANAND, 2011) é um mecanismo para mitigar

ataques de inundação e fornecer serviço legítimo para os usuários reais. A estratégia baseia-se no acréscimo de um campo no cabeçalho do pacote: o caminho do pacote entre a origem e o destino. A ideia é que pacotes subsequentes ao inicial provavelmente seguem o mesmo caminho entre a origem e o destino, permitindo diferenciar pacotes legítimos ou maliciosos desde que o atacante não seja capaz de modificar esse campo.

O índice de confiança para discernimento é contabilizado através da quantidade de desvios, com relação ao caminho mais antigo, de cada pacote. Os pacotes com menos desvios a eles associados entram em filas separadas, com maior grau de confiança, e são priorizados. Dessa forma, os pacotes com menor grau de confiança, que provavelmente são parte de um ataque de inundação, são eliminados mais facilmente, reduzindo a quantidade de pacotes maliciosos que deixam o roteador em caso de congestionamento do mesmo.

Um das principais desvantagens nesse sistema é o fato do roteador ser o equipamento que executa a ação de classificação, onerando o seu processamento além da sua função básica de roteamento. Além disso, esse esquema depende da estabilidade dos caminhos nas topologias da Internet. Entretanto, estudos mostram tendências inesperadas de instabilidade no roteamento e anomalias nas trocas de informações de roteamento inter-domínios (LABOVITZ; MALAN; JAHANIAN, 1999). Outro ponto de limitação do sistema *Path Attestation* é que ele é efetivo somente contra ataques de inundação com endereços IP de origem forjada, já que em ataques com origens legítimas (e.g., realizados por meio de uma rede zumbi) os pacotes seguem o mesmo caminho conforme esperado de um tráfego legítimo. Como a maioria dos ataques DDoS são baseados em volumetria (RIOREY, 2014), o

sistema Arquitena tem um abrangência maior na defesa, sendo a origem forjada ou não. Ao contrário do Arquitena, esse sistema se limita a detecção de ataques DDoS e não há qualquer técnica de renovação do valor do fluxo médio de pacotes na rede.

3.1.3. Mecanismo baseado em equações matemáticas

A abordagem descrita em (MANN; KUMAR, 2011), baseada em equações matemáticas, se destaca por descobrir a quantidade de pacotes maliciosos no fluxo juntamente com os pacotes legítimos. Esse número é então utilizado em conjunto com um algoritmo denominado HCI-MPR (*Hoop Count Inspection with Malicious Probability Rate* - Inspeção de Contagem de Saltos com Taxa de Probabilidade Maliciosa), a fim de mitigar os pacotes maliciosos. Esse algoritmo se baseia na contagem dos saltos entre a origem e destino: se a contagem atual de saltos é diferente da armazenada, o pacote é descartado. Conforme testes realizados pelos autores, o método alcançou aproximadamente 95% de taxa de detecção, além de diminuir o tempo de computação para tal ação.

Como o HCI-MPR usa o TTL (*Time To Life* - Tempo de Vida) para executar o descarte do pacote comparando com o TTL anterior, qualquer alteração de roteamento na rede pode invalidar o dado de comparação, induzindo o modelo a falhas na detecção. O HCI-MPR tem uma técnica para efetuar a contagem dos pacotes maliciosos e não realiza qualquer fase de mitigação após a detecção. O sistema Arquitena não contabiliza os pacotes maliciosos, e se baseia na renovação do valor do fluxo médio de pacotes (volumetria). Adicionalmente, além do processo de detecção efetua a mitigação e bloqueio do tráfego malicioso.

3.1.4. Abordagem baseada em estatística para ISPs

A abordagem baseada em estatística descrita em (GUPTA; MISRA; JOSHI, 2008) é composta de dois módulos: detecção e classificação do tráfego. O método de detecção consiste em monitorar as mudanças abruptas da propagação de tráfego dentro da rede de um ISP. As principais métricas utilizadas são o volume e o fluxo e, assim, pode-se determinar qual o modelo de tráfego normal. O tráfego é medido constantemente, sendo que para verificar o volume monitora-se a quantidade de bytes e, para o fluxo, monitora-se a quantidade de pacotes.

Após a detecção, os limites são definidos para determinar a classificação daquele determinado pacote, através de ferramentas estatísticas que analisam e medem as causas dos defeitos, tal como o Seis-Sigmas (KWAK; ANBARI, 2006). Dessa forma, pode-se inferir qual o fluxo malicioso e qual o tráfego legítimo. Com a definição dos limites entre faixas possíveis como, por exemplo, normal, suspeito e malicioso, o tráfego pode ser bloqueado nos roteadores de borda. De acordo com os testes realizados pelos proponentes dessa técnica, os resultados foram efetivos alcançando uma taxa de detecção em torno de 98%.

Uma desvantagem desse sistema é que a classificação de anomalias de tráfego e a definição do limite de tráfego são baseadas em cálculos estatísticos, e não efetua verificação constante do tráfego. Dessa forma, o processo de detecção fica desatualizado. O sistema Arquitena se baseia apenas na volumetria do tráfego. Uma vantagem do sistema proposto é ter fases pós-detecção e apresentar uma técnica de renovação do valor do fluxo médio de pacotes.

3.2. Detecção e mitigação de ataques DDoS, com utilização de agentes

Nesta seção, são discutidas soluções que se baseiam na utilização de arcabouços com agentes ou sistemas multiagentes para detecção e/ou mitigação de ataques DDoS.

3.2.1. Defesa contra DDoS utilizando Aprendizado por Reforço Cooperativo

O método descrito em (MALIALIS; KUDENKO, 2013) consiste em utilizar agentes que interagem com o ambiente, no caso uma rede, e aprendem por reforço, ou seja, é retornado um valor numérico na forma de recompensa ou punição. O aprendizado usa um Processo de Decisão Markoviano como modelo matemático, e de acordo com a atualização do estado do ambiente somado à recompensa ou punição, as próximas ações do agente são delegadas. Os agentes, com a função de aprendizado, são instalados nos locais da rede mais próximos das vítimas ou mais próximos da borda. Cada roteador, com o seu respectivo agente, aplica um limiar probabilístico de tráfego e cada agente tem a mesma função de recompensa de forma a receber recompensas ou punições idênticas. A quantidade de agentes com essa função de aprendizado pode depender do tamanho da estrutura da rede do ISP, por exemplo, sendo importante o seu correto dimensionamento para prover uma taxa de detecção adequada sem onerar demasiadamente a rede. O conceito de cooperação entre os agentes é empregado a fim de tornar mais preciso à detecção de ataques e discernimento entre tráfego legítimo e malicioso.

Assim como o Arquitena, há diferentes funções para os agentes, mas com a mesma meta de mitigar o ataque DDoS. A grande desvantagem dessa técnica em

relação ao Arquitena é que a renovação do valor do fluxo de pacotes é realizado somente na fase inicial de treino. Como resultado, ele não consegue distinguir tráfego malicioso quando o seu volume está próximo do valor normal.

3.2.2. Arcabouço de defesa adaptativa e cooperativa contra ataques à Internet

O arcabouço proposto em (KOTENKO; ULANOV, 2007) se baseia em conceitos comuns de um Sistema Multiagentes para defesa contra ataques DDoS: diferentes designações para cada time de agentes; a coordenação das ações dos agentes, comunicação seletiva e monitoramento é baseada em uma plano geral. Nesse SMA, o agente que identifica o ataque solicita aos outros agentes do sistema informações que supostamente servem como base para discernimento do ataque e todas as mensagens entre agentes ocorrem usando o protocolo TCP (*Transmission Control Protocol* - Protocolo de Controle de Transmissão).

Como os agentes desse arcabouço utilizam sensores para coletar dados do tráfego normal é possível identificar anomalias no tráfego corrente. Assim, o sistema de defesa se baseia na detecção com aprendizado constante do tráfego e utiliza a probabilidade para aplicar limites de tráfego. Esse arcabouço conta também com uma base de dados de clientes legítimos a fim de selecionar o tráfego prioritário. Diferentes modos de defesa são utilizados com base na cooperação entre os agentes, sendo o melhor resultado alcançado pelos autores com a completa cooperação entre os agentes. Essa técnica é semelhante ao Arquitena, pois utiliza filtros para bloquear os ataques DDoS, aprende com a variação de tráfego e utiliza a cooperação entre os agentes para facilitar todo o processo de mitigação.

Porém, uma das vantagens do sistema Arquitena é que alguns agentes são ativados somente quando um ataque DDoS está em ação. Caso contrário, os agentes permanecem inativos, sem realizar troca de mensagens ou solicitar coleta de dados no agente externo, não consumindo processamento do equipamento de rede ou do servidor onde os agentes são executados.

3.2.3. Mecanismo com multiagentes para reconhecimento de padrões e detecção de ataques DDoS

O sistema proposto em (BAIG; SALAH, 2009) é composto de múltiplos agentes com a única função de detectar ataques DDoS dentro de uma rede de produção. Através do reconhecimento contínuo de padrão pelos agentes, a taxa de assertividade na detecção dos ataques DDoS é alta. O mecanismo consiste de um SMA para detecção, no padrão distribuído das propriedades de reconhecimento do algoritmo Grafo Neurônio (*Graph Neuron* – GN) (BAQER; KHAN; BAIG, 2005). Como há muitos agentes, a tolerância a falhas é maior, mesmo que aumente a possibilidade de se ter agentes comprometidos pelo ataque DDoS e a sobrecarga de comunicação entre os agentes.

Os agentes detectores tem em sua memória limites para determinação do tráfego normal. Cada vítima tem diferentes padrões de fluxo definidos como um limite, e esses padrões são distribuídos para os agentes. A estrutura GN é análoga a um grafo dirigido, com múltiplos agentes. Esses agentes são como os vértices do grafo e as conexões entre os agentes são as arestas. O algoritmo requer que os agentes armazenem novos padrões observados, e requer também que os mesmos agentes verifiquem padrões gravados na memória. Cada agente desenvolve o mesmo conjunto de operações requeridas para a detecção do ataque DDoS. Os

agentes detectores são configurados previamente com conhecimento da vizinhança de agentes, a fim de facilitar a construção dos padrões. O principal resultado desse método é que, com uma maior quantidade de agentes detectores sendo utilizada, ele permite melhorar consideravelmente a taxa de falso-positivo e falso-negativo, conforme mostrado na tabela 2. Outros resultados foram alcançados, através de simulações, como a média de detecção que ficou próxima de 88% com intensidade maior de tráfego.

Tabela 2 – Correlação entre número de agentes e taxas de detecção. Adaptado de (BAIG; SALAH, 2009).

Número de agentes	**Falso-positivo (%)**		**Falso-negativo (%)**	
	Baixa taxa entre chegadas	**Alta taxa entre chegadas**	**Baixa taxa entre chegadas**	**Alta taxa entre chegadas**
128	**3.11**	**5.28**	**22.8**	**38.72**
1024	**0.15**	**1.55**	**1.1**	**11.37**

De maneira similar ao sistema Arquitena, os agentes desse arcabouço criam um padrão de limite de pacotes em cada equipamento e gradualmente o tornam mais preciso, a fim de identificar o tráfego anômalo. Entretanto, esse SMA é focado apenas em detectar ataques DDoS, não havendo mitigação do mesmo. Além disso, uma desvantagem desse SMA é que a troca de mensagens entre os agentes é constante. Desta forma, dependendo da quantidade de agentes empregados, pode-se ter um uso elevado de recursos no equipamento que executa o SMA. No Arquitena, pode-se ativar/desativar agentes específicos dependendo da intensidade do ataque DDoS, reduzindo o uso de recursos.

3.2.4. Detecção distribuída usando agentes móveis contra ataques DDoS

O Sistema de Detecção de Intrusão Distribuído (*Distributed Intrusion Detection Systems* - DIDS) (AKYAZI; UYAR, 2008) é um método que correlaciona dados de intrusão, coletados em diversos servidores da rede, utilizando agentes móveis para facilitar o processo de detecção. O principal objetivo do sistema é diminuir a carga de trabalho da rede. Para isso são utilizados quatro tipos de agentes:

- **Agente principal (AP):** localizado na parte isolada da rede
- **Agente Estático (AE):** localizado nos servidores
- **Agente Móvel (AM):** pode ser criado pelo AP ou AE
- **Agente Alarme:** localizado na parte isolada da rede

Nesse método, quando uma ou duas mensagens, seja legítimo ou de um atacante, com o mesmo endereço IP são enviadas de diferentes servidores, o Agente Principal cria um Agente Móvel e envia-o para os servidores donde as mensagens se originaram. Após o Agente Móvel analisar os dados gravados nas hipotéticas vítimas, ele verifica se a origem do endereço IP é a mesma. Em caso positivo, o Agente Móvel envia uma mensagem para o Agente Alarme que comunica o gerenciador do DIDS.

Nesse sistema não há um processo de mitigação após a detecção, de maneira oposta ao sistema Arquitena que efetua a mitigação do ataque até o bloqueio. O tempo de detecção aumenta de acordo com a quantidade de verificações dos endereços IP de origem que os agentes móveis efetuam nos

servidores. O sistema Arquitena tende a diminuir o tempo de detecção com a renovação constante do fluxo médio de pacotes, tornando mais assertiva a mitigação de um ataque DDoS.

4 SOLUÇÃO PROPOSTA: SISTEMA ARQUITENA

Esta seção descreve o sistema Arquitena, detalhando suas funcionalidades (subseções 4.1 a 4.7), arquitetura (subseção 4.8) e também as principais vantagens e limitações do sistema (subseção 4.9 e 4.10). Adicionalmente, na subseção 4.11 são discutidos os requisitos de hardware e software para implantação do sistema, enquanto na subseção 4.12 é detalhado o ambiente de simulação Delos utilizado para sua análise experimental.

4.1. Requisitos do sistema

A seguir está uma descrição dos requisitos funcionais e não-funcionais do sistema Arquitena:

- **Requisitos Funcionais do sistema**: permitir detecção de ataques DDoS; permitir bloqueio de ataques DDoS; permitir mitigação de ataques DDoS; sistema de detecção deve se adaptar ao padrão de tráfego da rede; preservar tráfego legítimo.
- **Requisitos Não-Funcionais do sistema**: minimizar o número de agentes ativos enquanto o sistema não estiver sob ataque; não deve onerar o processamento dos equipamentos de rede; minimizar processamento do sistema de análise; minimizar a troca de mensagens entre os agentes na fase de detecção.

4.2. Descrição

O sistema Arquitena utiliza o conceito de sistemas multiagentes para fornecer segurança para a rede de um ISP, algo que parece apropriado considerando que este é um ambiente de fluxo altamente dinâmico e com uma vasta quantidade de

nós interligados. Conforme subseção 2.6, características como inacessível, dinâmico, não determinístico e contínuo são inerentes ao ambiente onde o sistema proposto está inserido. O sistema multiagentes trabalha de forma cooperativa utilizando o modelo BDI, apresentado na seção 2.3. Com essa estrutura, pode-se monitorar a rede do ISP, detectar ataques DDoS baseados na inundação de pacotes, mitigar o ataque gradualmente e bloquear dinamicamente as entradas da rede do ISP, se necessário.

Para um sistema de defesa ser efetivo, ele precisa assegurar algumas características essenciais para os dados e serviços, tais como disponibilidade, integridade e confidencialidade. O sistema Arquitena tem como objetivo garantir essas características por meio de agentes implantados em várias localidades da rede do ISP, como a borda da rede, a parte central da rede e a camada de serviço, onde podem estar os equipamentos alvos de um hipotético ataque. Dessa forma, o sistema tende a ter uma resposta mais rápida a um ataque DDoS, mantendo assim o serviço oferecido disponível e o tráfego legítimo preservado.

A fim de garantir um sistema de defesa ubíquo, o Arquitena tem um mapa virtual da rede real do ISP, com todos os equipamentos e conexões pertinentes. Este mapa virtual fica em um ambiente de simulação para multiagentes chamado Delos.

O sistema proposto tem como característica importante a habilidade de comunicação entre todos os agentes que o compõem. A cooperação entre os agentes é uma característica importante utilizada de modo que está implícita nos planos do sistema. O sistema proposto não é puramente reativo, pois os agentes, além de responderem continuamente às alterações do ambiente, tentam alcançar sistematicamente suas metas com procedimentos complexos de ação. Uma base de

crenças é utilizada por cada agente, a fim de armazenar informações necessárias para alcançar as metas individuais e as metas do sistema.

4.2.1. Agentes

O sistema Arquitena tem quatro tipos de agentes:

- **Externo:** analisador/coletor do tráfego de rede. Esse agente está sempre ativo. Ele envia as estatísticas do fluxo de pacotes que trafegam por cada equipamento de rede, que é monitorado por um agente ativo. As informações são sempre requisitadas pelos equipamentos que os agentes ativos monitoram, descritos a seguir. O agente externo é necessário pela necessidade do sistema Arquitena coletar dados estatísticos de cada equipamento monitorado pelos agentes. A comunicação entre o equipamento de rede e o agente externo pode ser realizada via protocolos de gerenciamento.

- **Internos:** Todos os agentes internos tem em sua base de crenças, no mínimo as informações: endereço IP do equipamento monitorado; tabela de adjacências (populada no momento inicial de ativação do sistema Arquitena); valor do fluxo médio de pacotes (**FPmed**) para aquele determinado equipamento, de cada equipamento alvo; valor do fluxo máximo de pacotes (**FPmax**) para aquele determinado equipamento. Apesar desta propriedade em comum, os agentes internos apresentam algumas características que permitem classificá-lo em 3 funções distintas, descritas a seguir:

- **Agente Minos:** são responsáveis pelo monitoramento dos equipamentos alvos que hospedam os serviços oferecidos pelo ISP como, por exemplo, DNS e WEB. O agente, nesse caso, pode mapear qualquer equipamento que forneça o respectivo serviço, como um servidor ou balanceador de carga. Dependendo da topologia desenhada pelo ISP, um firewall que permeia a DMZ (*Demilitarized Zone* - Perímetro de Rede) para a LAN (*Local Area Network* – Rede de Área Local) pode ser monitorado permanentemente devido ao seu estado de criticidade; dessa forma, os agentes ficam sempre ativos. Cada agente pode apresentar 12 ou 24 contadores, divididos por intervalo de tempo (**tcont**) de 2 ou 1 hora, respectivamente. Esses contadores são relacionados ao fluxo de pacotes que são direcionados para o possível alvo em certo período do dia, e são inseridos como crenças na base de cada agente. Cabe notar que um número diferente de contadores poderia ser utilizado, provendo maior ou menor granularidade do processo de detecção, mas os intervalos definidos são necessários para delimitar os horários específicos de utilização da rede. Por exemplo, o tráfego de um ISP pode ter um pico maior próximo das 12h quando comparado com o tráfego das 23h. Como resultado, o agente Minos precisa de menos contadores, pois ele é o sumidouro do tráfego e possível alvo de ataques. Isso o difere dos agentes Éacos e Adamantos na quantidade de contadores, que além de dependerem do ciclo diário, dependem da quantidade de serviços oferecidos que podem ser considerados potenciais alvos. Os outros agentes internos devem ter o monitoramento de qualquer tráfego

destinado para os potenciais alvos, assim a quantidade de contadores é maior, pois em cada **tcont** há um valor de **FPmed** diferente.

- **Agente Éacos:** são responsáveis pelo monitoramento dos equipamentos que estão na parte central da rede como roteadores e firewalls, desde a primeira linha de vizinhança dos servidores até a borda da rede do ISP, com exceção dos nós com maior fluxo de pacotes passante. Podem ser ativados por outros agentes, caso a rede esteja sob um hipotético ataque DDoS, ou aleatoriamente pelo ambiente. Agentes deste tipo podem apresentar de 12 a 24 contadores para cada equipamento alvo da rede do ISP, divididos pelo intervalo **tcont**, e outro contador mestre para sumarizar cada conjunto de contadores. Esse agente pode estar em dois estados diferentes: ativo ou inativo. O estado **ativo** refere-se a quando o agente está monitorando algum equipamento de rede. Nesse estado, o agente mantém a comunicação com o equipamento, a fim de requisitar atualizações. O agente utiliza a base de crenças, com as informações do equipamento que monitora, para efetuar as ações pertinentes à interação com o ambiente e com os outros agentes. Já no estado **inativo**, o agente não está monitorando um equipamento de rede. Nesse estado, ele não mantém comunicação com o equipamento, mas mantém as informações pertinentes do equipamento em sua base de crenças.
- **Agente Adamantos:** são responsáveis pelo monitoramento dos equipamentos da parte central da rede com maior fluxo de pacotes como roteadores centrais, firewalls e balanceadores, desde a primeira

linha de vizinhança dos servidores até os equipamentos de borda da rede do ISP. Esse agente está sempre ativo. Agentes deste tipo podem apresentar de 12 a 24 contadores para cada equipamento alvo da rede do ISP, divididos pelo intervalo **tcont**, e outro contador mestre para sumarizar cada conjunto de contadores. Esses agentes tem em sua base de crenças uma informação adicional, além das informações padrão: se ele pertence à borda ou não.

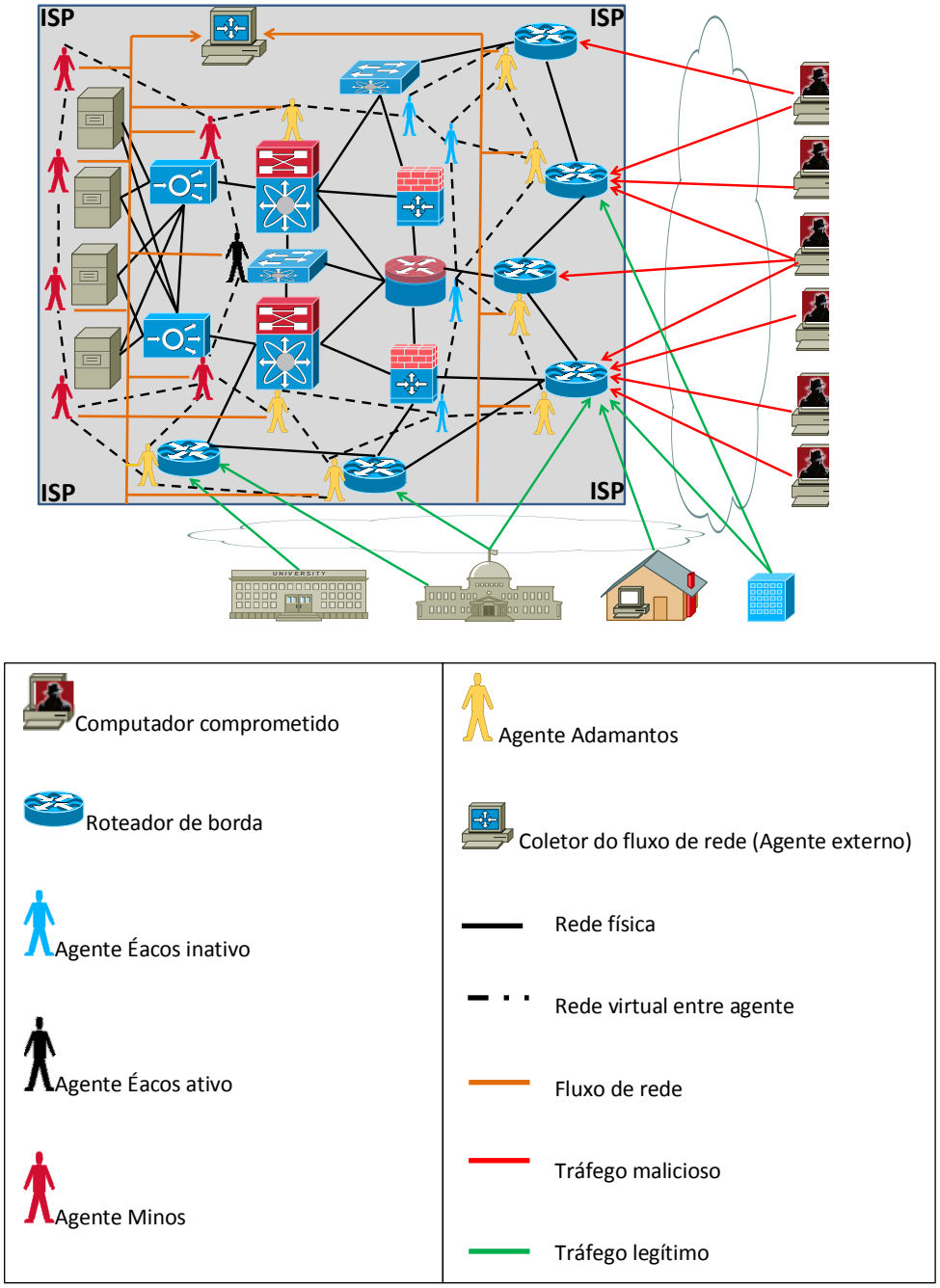
A Tabela 3 resume as características de cada agente do sistema Arquitena.

Tabela 3 – Características dos agentes do sistema Arquitena.

Agente	Função	Estado	Contadores	Local de atuação
Externo	Coleta e envio de dados	Sempre ativo	Não há	Toda rede do ISP
Interno: Minos	Monitoramento, detecção, mitigação	Sempre ativo	1 x tcont	Camada de serviço do ISP (vítimas)
Interno: Adamantos	Monitoramento, detecção, mitigação	Sempre ativo	1 x tcont + sumarizador	Borda da rede e equipamentos críticos do ISP
Interno: Éacos	Monitoramento, detecção, mitigação	Ativo ou inativo	1 x tcont + sumarizador	Centralidade da rede do ISP

A figura 2 ilustra a distribuição dos agentes, no sistema Arquitena, considerando um cenário em que a rede está sob um ataque DDoS. Nesse exemplo, os roteadores de borda recebem tanto o tráfego malicioso quanto o legítimo, e a distribuição dos agentes na rede do ISP é realizada de acordo com a criticidade do equipamento na topologia.

Figura 2 – Distribuição dos agentes Arquitena na rede do ISP.



4.2.2. Ações dos agentes internos

A operação dos agentes depende fortemente do valor médio do fluxo de pacotes por segundo na rede, **FPmed**, o qual é expresso em pacotes por segundo (pps). Para os agentes Minos e Adamantos, um novo valor do **FPmed** é calculado a cada 1 ou 2 horas, durante os ciclos diários, devido a estarem sempre ativos. A cada

período **tcont**, o valor do **FPmed** é adicionado ao contador referente àquele intervalo, e um novo valor é gerado pela média aritmética do último valor do **FPmed** gerado com o valor do **FPmed** armazenado no contador. O novo valor do **FPmed** segue esta fórmula: **FPmed = (FPmed(tcont) + FPmed(t)) / 2,** conforme código abaixo. Esse pequeno código tem em todos os agentes do sistema, para cada horário.

```
?counter(Z); .print(Z);
if(Z == 5){-+normal_traffic(false); -+counter(0);!!wake_Eacos;}
         else {-+normal_traffic(true)};
if (Y < F){(NA = ((Y+F12)/2)); .print(NA); -+fpav12(NA);};!verify_flow1.
```

Essa nova média é adicionada à base de crenças do agente, e esse valor é utilizado como parâmetro até um novo valor ser gerado pela próxima coleta. Se a rede do ISP estiver sob ataque, os agentes Éacos ativos tem o mesmo comportamento dos agentes Minos e Adamantos, em relação ao ciclo de coleta de dados. Assim, os equipamentos recebem os dados ininterruptamente do agente externo, com a granularidade especificada por **tcont**, pois o monitoramento está ativo.

Caso a rede do ISP esteja com tráfego normal, todos os agentes Éacos podem permanecer inativos, não monitorando qualquer equipamento da rede, mas recebem um novo valor do **FPmed**, de acordo com o ciclo contínuo de ativação entre todos os agentes Éacos inativos. Nesse ciclo, o intervalo entre a ativação de cada agente depende da quantidade de agentes Éacos inativos na rede. O objetivo desse ciclo é não sobrecarregar o agente externo com muitas requisições simultâneas e desnecessárias naquele momento. O procedimento da coleta dos dados no intervalo de inatividade é diferente do momento que o agente esteja ativo, pois no primeiro

caso o equipamento monitorado não recebe os dados do agente externo. A figura 3 mostra um exemplo com quatro agentes Éacos, em um ciclo contínuo entre agentes inativos e assim, cada agente Éacos é ativado pelo sistema, fica ativo e monitorando o equipamento da rede pelo intervalo definido por **tcont**, por exemplo, duas horas.

Os agentes Adamantos e Éacos apresentam contadores específicos com o valor médio de pacotes que deveriam passar por aquele equipamento, e com o destino de cada um dos equipamentos hospedeiros do serviço. A quantidade de contadores em cada agente Éacos ou Adamantos depende da quantidade de equipamentos alvos do ISP, conforme equação abaixo:

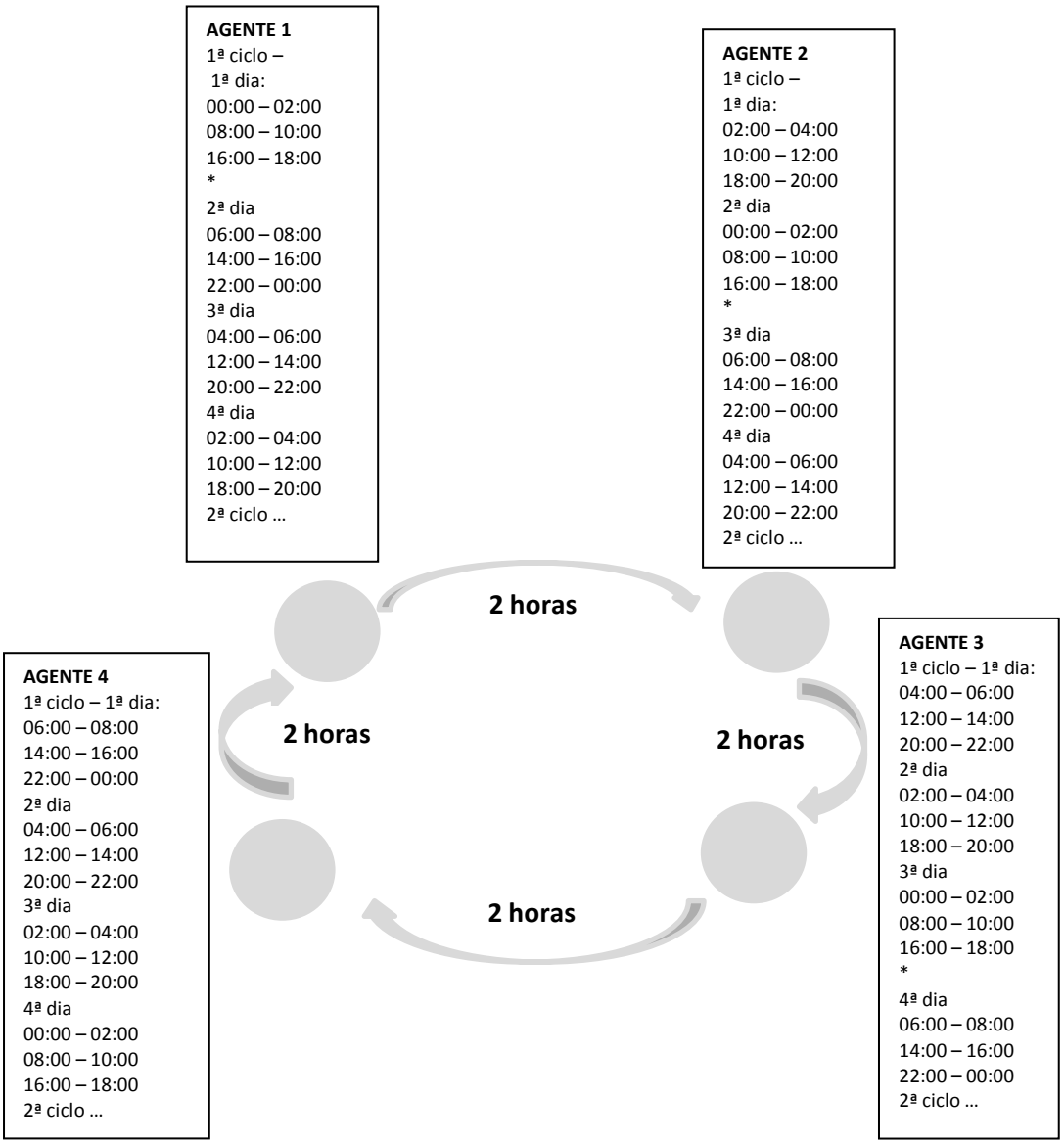
Quantidade de contadores = (12 intervalos x equipamentos alvos) + agentes mestre sumarizadores

O ciclo de ativação dos agentes Éacos é codificado conforme abaixo, sendo que para cada agente foi determinado o específico horário para se tornar ativo.

```
+!wakeup : .time(H,N,S) & (H == 00) & (N == 00) & (S == 00)
 <- .print("leader to do wake up Eacos hour 1");
    ?lista(NL4);
    .nth(0, NL4, W);
    .send(W, askOne, online(X), online(X));
    .my_name(Name);
    if (X == inactive) {
         .send(W, achieve, statusOnline_Fromleader(Name));};
.wait(2000);
!!start1.

+!sleep : .time(H,N,S) & (H == 01) & (N == 59) & (S == 00)
 <- .print("leader to do sleep Eacos hour 1");
    .my_name(MyName);//sempre leader
    ?lista(NL4);
    .nth(0, NL4, W);
    .send(W, askOne, normaltraffic(Traffic), normaltraffic(Traffic));
    .send(W, askOne, online(Status), online(Status));
    .send(W, askOne, wakeUp(Name), wakeUp(Name));
    if ((Traffic == true) & (Status == active) & (Name == MyName)) {
                   .send(W, achieve, statusOffline(Name));};
```

```
      .wait(2000); !!start2.

+!wakeup: true <- !!start1.
+!start1: true <- !!sleep.

+!sleep: true <- !!start2.
      +!start2: true <- !!wakeup.
```

Figura 3 – Ciclo de ativação dos agentes Éacos inativos.



Nesse contexto, qualquer agente, seja Minos, Adamantos ou Éacos, coleta e aperfeiçoa gradualmente o valor do **FPmed** que passa por aquele equipamento naquele intervalo de tempo. Esse processo tende a se tornar mais preciso ao longo

do tempo, garantindo que o valor do **FPmed** fique mais próximo do realmente apresentado pela rede. Assim, o processo de detecção dos ataques DDoS tende a desenvolver um menor grau de variação quando esses contadores se estabilizam, assumindo que a rede apresente um comportamento razoavelmente regular.

Existem duas situações em que um volume de tráfego monitorado é considerado anormal pelos agentes. O primeiro é a extrapolação do valor de **FPmed**, sendo que dentro de um período específico de monitoramento, caso o valor do **FPmed** seja ultrapassado, o tráfego excedente naquele equipamento não é mais considerado normal. O segundo é o quando o valor estabelecido de margem de segurança (**cns**=90% do **FPmax**) é ultrapassado, assim cada um dos valores do **FPmed** aprendidos por cada agente deve ser sempre menor do que o **FPmax** para aquele equipamento monitorado. Portanto, a soma dos valores do **FPmed** de cada contador também deve ser menor do que o **FPmax**, e esta regra tem como objetivo evitar um colapso do equipamento perante um ataque DDoS.

Como os agentes Minos cuidam exclusivamente de potenciais alvos, eles devem conhecer os valores médios de pacotes que deveriam chegar naquele equipamento hospedeiro do serviço oferecido. Os pacotes que chegam tem como destino o endereço IP de um servidor, um balanceador de carga ou um firewall.

As figuras 4, 5 e 6 ilustram o fluxograma que mostra o processo de ativação, monitoramento, detecção e bloqueio dos agentes Minos, Adamantos e Éacos, respectivamente, após 24 horas de implantação do sistema Arquitena.

Figura 4 – Fluxograma de ações do agente Minos.

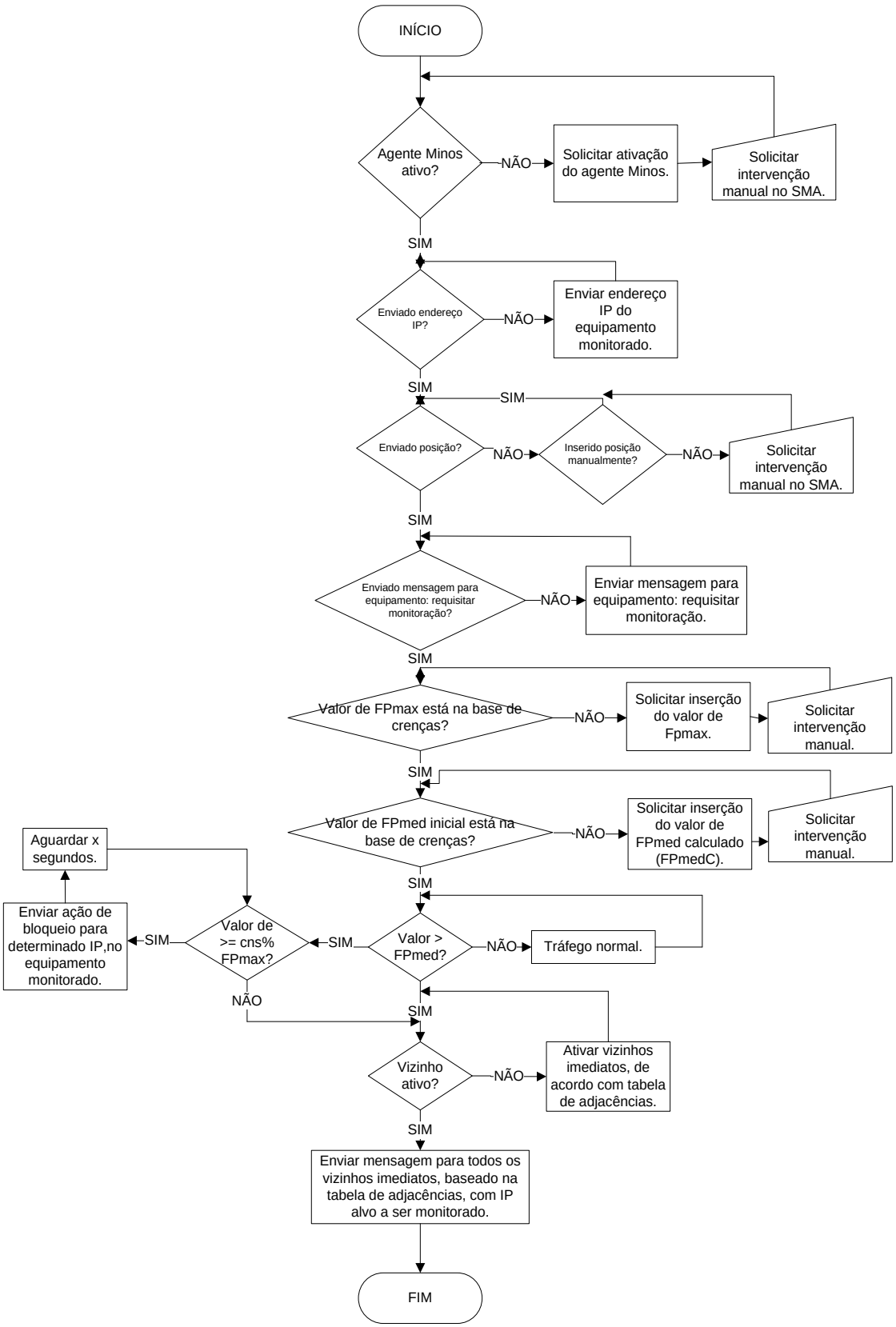


Figura 5 – Fluxograma de ações do agente Adamantos.

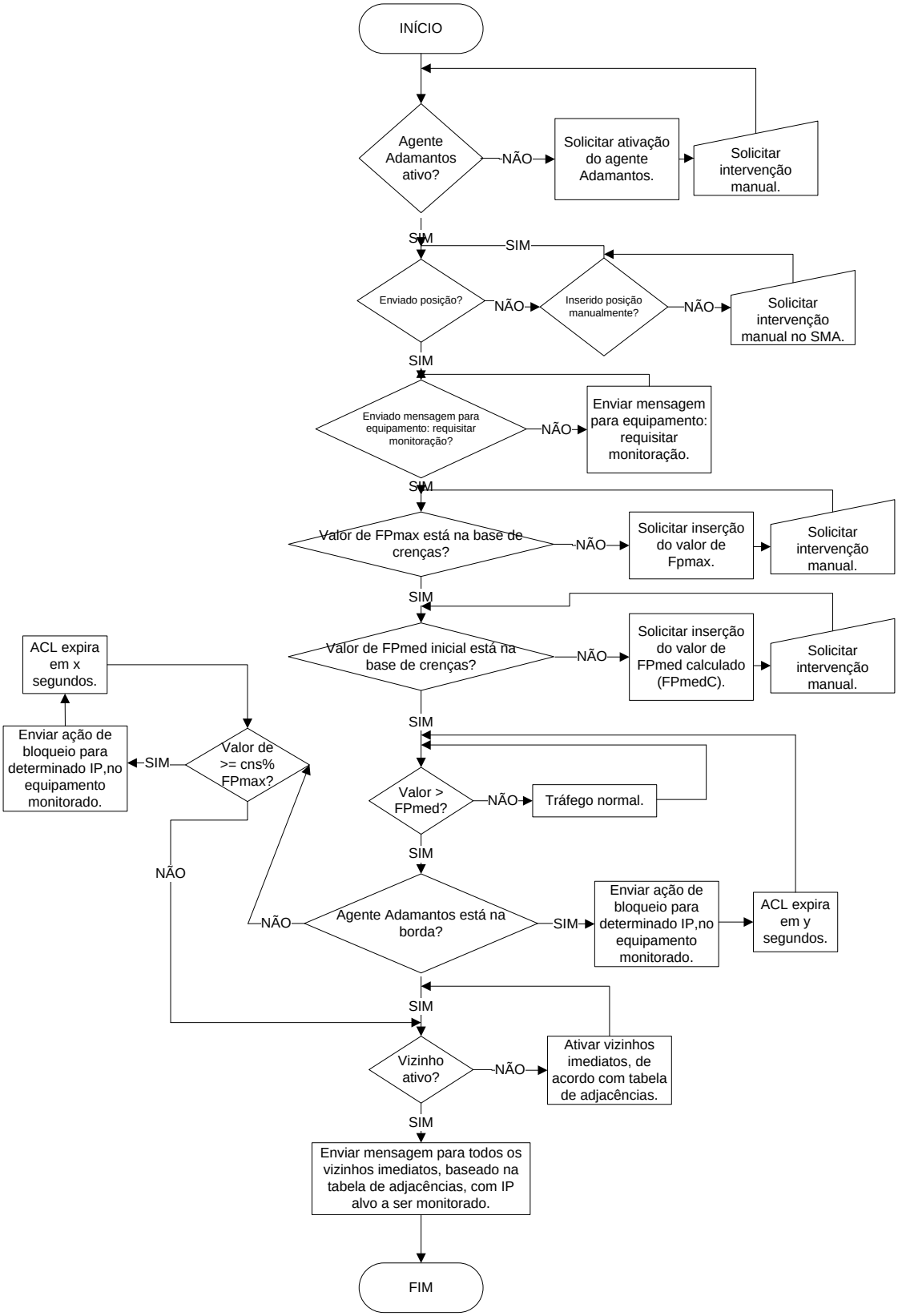
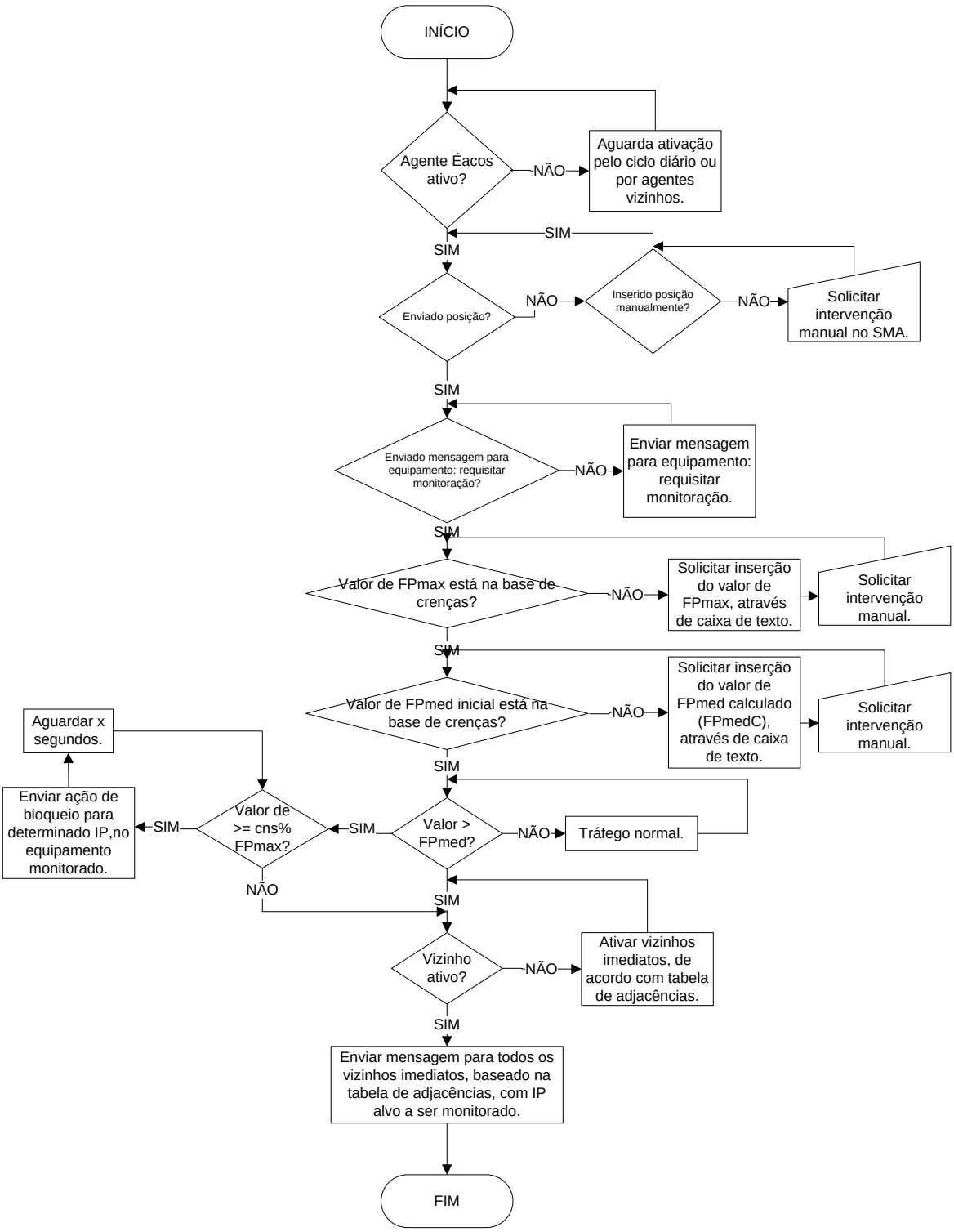


Figura 6 – Fluxograma de ações do agente Éacos.



4.3. Inicialização do sistema Arquitena

No momento de ativação do sistema Arquitena, todos os agentes Minos informam o endereço IP dos equipamentos por eles monitorados para todos os outros agentes do sistema, a fim de introduzir as informações nos contadores de cada agente Éacos ou Adamantos. A partir desse momento, cada agente Éacos ou Adamantos pode criar a quantidade correta de contadores.

A localização de cada agente é inserida manualmente em sua base de crenças, de acordo com a topologia do ISP, criando uma topologia virtual. No momento de inicialização do sistema Arquitena, todos os agentes internos trocam informações de localização entre si a fim de compor a tabela de adjacências, isto é, determinar quais são seus vizinhos diretos.

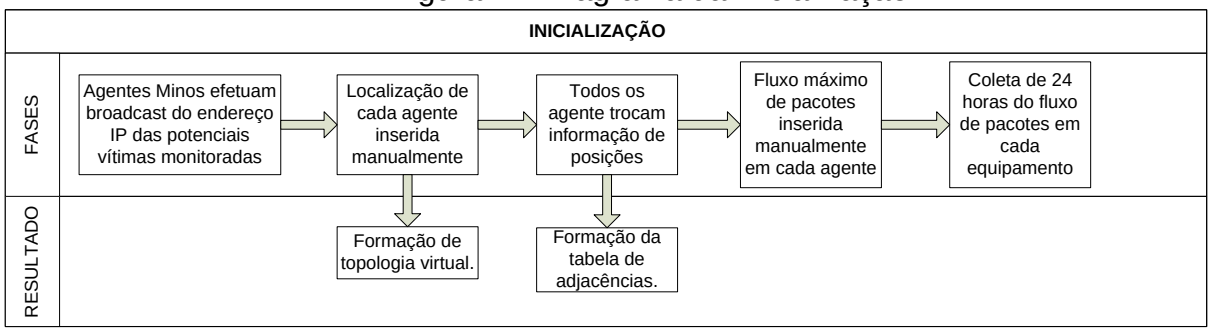
A fim de facilitar o processo inicial de detecção e determinar um valor base **FPmed**, todos os agentes permanecem ativos 24 horas após a implantação do sistema Arquitena. Alternativamente, o valor do **FPmed** e **FPmax** podem ser inseridos manualmente pelo próprio administrador do ISP na implantação do Arquitena, caso essa estatística já tenha sido gerada. Qualquer que seja a abordagem adotada, os agentes aprendem gradualmente a quantidade de pacotes que passa por aquele equipamento, naquele específico intervalo de tempo, pois coletam no equipamento um novo valor e efetuam a média aritmética entre o novo e o antigo valor na base de crenças, conforme seção 4.2.2. Já o valor do **FPmax**, que também é expresso em pacotes por segundo (pps) é disponibilizado pelo fabricante do equipamento ou com testes realizados pelo próprio ISP.

As 24 horas iniciais são essenciais para tornar o processo de mitigação mais preciso, pois os agentes iniciam a coleta e posterior renovação da quantidade de pacotes que tramitam por aquele equipamento. Essa renovação do valor do fluxo de pacotes consiste nos valores de:

- **FPmed** de cada alvo em potencial.
- **FPmed** sumarizado do equipamento específico.

O diagrama da figura 7 mostra por fases a etapa de inicialização do sistema Arquitena.

Figura 7 – Diagrama da inicialização.



4.4. Monitoramento no sistema Arquitena

A função primária dos agentes é monitorar a rede a fim de detectar algum tráfego anômalo. Como cada agente ativo monitora cada equipamento da rede do ISP através de comunicação direta, o equipamento por delegação do agente passa seus valores referentes ao limite **FPmed** e a amostragem para o agente externo. A informação que o agente passa para o equipamento é o endereço IP da potencial vítima a ser protegida contra eventuais ataques DDoS.

Com estas informações do endereço IP, **FPmed** e amostragem, o agente externo pode assertivamente coletar o fluxo do tráfego a ser analisado e enviar mensagens direcionadas corretamente para cada equipamento, pois este pode estar na rota de um ataque DDoS. Dessa forma, os contadores tem as informações mais atuais do fluxo de pacotes naquele equipamento monitorado.

Os agentes ativos sempre estão monitorando os equipamentos de rede, como os agentes Adamantos e os agentes Minos, e outros agentes podem estar ativos por um determinado período, dependendo da granularidade do ataque, como os agentes Éacos. Outra possibilidade de se ter agentes Éacos momentaneamente ativos, mesmo que a rede não esteja sob um ataque é quando tais agentes estão no processo cíclico de ativação dos agentes para atualização de sua base de crenças.

O monitoramento de cada equipamento, quando ativada, consiste da comunicação unidirecional de cada agente com o equipamento, bem como dos equipamentos com o agente externo implantado na rede do ISP. O agente externo envia três informações para os equipamentos: alertas de que o valor do **FPmed** foi ultrapassado, quantidade de pacotes naquele intervalo e fluxo de pacotes. Para os agentes Minos, Adamantos e Éacos ativos o intervalo entre as mensagens recebidas pelos equipamentos é definido pelo contador **tmen**, o qual pode assumir valores da ordem de segundos (e.g., 1 a 3), entre as mensagens. O agente externo recebe informações de cada equipamento da rede, sendo que os agentes do Arquitena definem quais equipamentos devem enviar seus dados. Os agentes próprios do agente externo, implantados em cada equipamento, executam a tarefa de coleta.

Todos os equipamentos, tais como servidores, roteadores, *switches*, *firewalls* e balanceadores de carga podem ser monitorados através desse sistema. A

definição de quais deles são monitorados em certo intervalo de tempo depende de regras pré-estabelecidas, com as seguintes recomendações:

- Agente ativado por agentes vizinhos: possível ataque DDoS;
- Equipamento é considerado uma potencial vítima de ataque DDoS, como servidor, *firewall* ou balanceador de carga;
- Equipamento está na borda da borda da rede do ISP;
- Equipamento é identificado pelo administrador do ISP como tendo com muitas conexões na parte central da sua rede.

Para não onerar o processamento do computador hospedeiro, no qual o Arquitena é executado, nem dos equipamentos da rede e do agente externo, a amostragem é definida pelo contador **tamo**, que é o valor da amostragem de pacotes analisado pelo agente externo. Por exemplo, suponha que **tamo**=1000. Nesse caso, a cada 1000 pacotes apenas 1 é analisado. Se, adicionalmente, faz-se **tmen**=1 segundo, cada agente ativo recebe do agente externo a informação dos pacotes analisados nesse intervalo. É importante notar, entretanto, que cada equipamento tem um fluxo característico de pacotes: em 1 segundo um servidor pode receber 2000 pacotes enquanto um roteador recebe 50 mil pacotes, por exemplo. Portanto, o tempo de resposta caso os limites estabelecidos sejam ultrapassados pode ser diferente para cada agente. O valor da amostragem determina, então, o quão rápido o sistema detecta o ataque DDoS devendo ser ajustado para que isto seja realizado antes que os recursos de hardware dos equipamentos se esgotem. Quanto menor o valor definido da amostragem, maior é o processamento referente a essa análise, mas em compensação a verificação de resultados positivos para ataques DDoS torna-se mais precisa.

No processo de monitoramento, se o valor do **FPmed** for ultrapassado por certa quantidade de vezes em um determinado intervalo de tempo, os agentes ativos acionam os seus vizinhos com uma mensagem *multicast* enviada a todos os seus vizinhos imediatos. Todos os agentes utilizam os contadores, mas apenas os agentes ativos os incrementam a cada mensagem enviada pelo agente externo ao equipamento, informando que o limite foi excedido. É necessário que o agente externo envie a mensagem com a quantidade de pacotes para os equipamentos, assim os agentes Éacos e Adamantos podem executar a soma de todos os contadores, a fim de não se aproximar do valor de segurança **FPmax**.

Considere, por exemplo, que um servidor tem média de tráfego de 40 pacotes por segundo (40 pps), e que o agente pode ativar todos os seus vizinhos em até 5 segundos, se a média for ultrapassada 3 vezes. Nesse caso, ocorre a sequência de eventos mostrada na Tabela 4. Nesse exemplo, no terceiro segundo o agente ativo envia uma mensagem *multicast* para seus vizinhos imediatos e ativa-os.

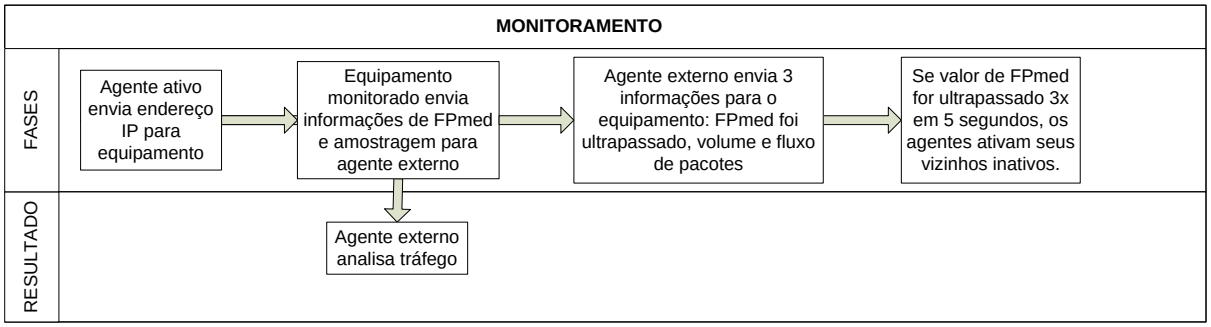
Tabela 4 – Modelo padrão de detecção e ativação de vizinhos.

Tempo	Quantidade de pacotes	Pacotes analisados	Pacotes direcionados para o alvo	FPmed excedido
1º segundo	100000	100	50	Sim
2º segundo	200000	200	70	Sim
3º segundo	**150000**	**150**	**45**	**Sim**
4º segundo	100000	100	38	Não
5º segundo	120000	120	42	Sim

Com esse sistema gradual de monitoramento, há um menor consumo de memória e processamento tanto para o agente externo quanto para os

equipamentos da rede. Uma razão é que, por exemplo, apenas os equipamentos que estão na rota do tráfego malicioso são monitorados em um ataque DDoS. Além disso, caso o ataque não perdure, os agentes Éacos envolvidos param o monitoramento e posterior mitigação e, assim, retornam ao estado de hibernação temporária. O diagrama da figura 8 mostra por fases a etapa de monitoramento do sistema Arquitena.

Figura 8 – Diagrama do monitoramento.



4.5. Detecção no sistema Arquitena

Com o sistema de detecção utilizado no Arquitena, o agente ativo verifica a quantidade de pacotes direcionados para a vítima que excede o **FPmed** presente em sua base de crenças. Caso o tráfego em excesso perdure sobre aquele equipamento, a verificação conjunta entre o agente interno e o agente externo deve diagnosticar se os pacotes em questão tem como destino a possível vítima. Nesse caso, o agente pode ativar todos seus agentes vizinhos inativos, e assim sucessivamente até a borda da rede. Esses agentes vizinhos consequentemente estabelecem comunicação direta com o equipamento de rede e estes diretamente com o agente externo.

De acordo com esse processo de reação em cadeia a informação original do endereço IP da hipotética vítima provém do próprio agente que ativa seus vizinhos.

Essa informação é passada de agente a agente, e no momento da ativação de qualquer agente, uma mensagem com essa informação (endereço IP) é passada para o agente externo, através do equipamento de rede.

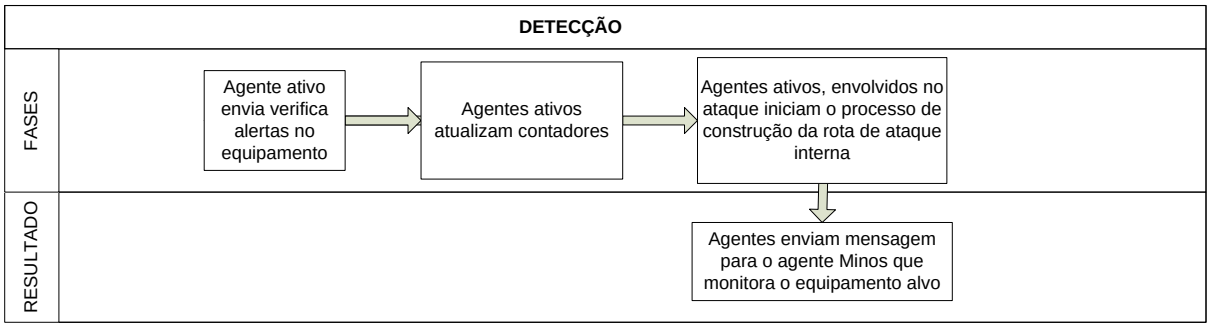
Se, por exemplo, em cinco segundos a média do tráfego não for excedida por três vezes, o agente pode tornar-se inativo, caso naquele intervalo ele não seja um dos agentes Éacos ativado pelo ambiente, dentro do ciclo contínuo de ativação. Como um equipamento da rede pode estar na rota de um ou mais ataques DDoS para diferentes vítimas, o agente tem a capacidade de ativar seus contadores internos proporcionalmente à quantidade de ataques concomitantes. Os agentes tem a capacidade de discernir as mensagens enviadas pelo agente externo ao equipamento, associando cada mensagem com o seu contador específico. Um agente, antes mesmo de ativar os seus vizinhos, deve verificar se o agente vizinho está ativo, pois, neste caso, não é necessário ativá-lo.

O processo de detecção é recorrente em toda a rede do ISP até sua borda. Caso um agente Minos detecte um ataque, os agentes Éacos e Adamantos são empregados para construir a rota de ataque desde a borda da rede até os equipamentos hospedeiros dos serviços disponíveis. Quando todos os agentes estiverem ativos entre o agente Minos e o agente que enviou a mensagem, seja ele central ou de borda, o agente Minos inicia o processo de construção da rota de ataque. Sendo assim, toda vez que um agente Minos tiver seu contador interno excedido ou receber uma mensagem de outro agente sobre o fluxo excedido, ele começa a construir a rota de ataque. Essa rota é constituída do caminho desde a posição de qualquer agente Minos até o ponto mais próximo alcançável do atacante, isto é, a borda da rede do ISP.

Caso um agente Éacos torne-se ativo e detecte um ataque ou um agente Adamantos detecte um ataque DDoS, esse agente deve enviar uma mensagem (sua posição) para o agente Minos responsável por monitorar aquele equipamento específico. Os agentes Adamantos devem executar o mesmo procedimento dos agentes Éacos ativos, pois, embora sempre ativos, nem sempre se encontram na rota de ataques. Quando o último agente Adamantos for acionado na borda, a rota de ataque fica completa. Essas rotas de ataque são importantes para avaliar padrões de ataque, para um determinado equipamento alvo da rede do ISP.

O ataque DDoS é considerado completamente detectado após a rota de ataque estar estabelecida. Nesse momento, as métricas que podem identificar a eficácia do sistema Arquitena, como o número de falso-positivos ou de falso-negativos, podem ser avaliados. O diagrama da figura 9 mostra por fases a etapa de detecção do sistema Arquitena.

Figura 9 – Diagrama de detecção.



4.6. Bloqueio do sistema Arquitena

A função primária do processo de bloqueio é evitar o consumo excessivo dos recursos de um determinado equipamento durante um ataque DDoS. Neste caso, o bloqueio do tráfego para um determinado alvo pode ser total através de regras aplicadas nos equipamentos de uma mesma camada, por exemplo, a borda ou a

parte central. A outra função é tentar bloquear o tráfego malicioso e preservar o tráfego legítimo. Isso é possível pelo posicionamento estratégico dos agentes ativos durante as fases de monitoramento e detecção.

O sistema de bloqueio é utilizado principalmente nos equipamentos de borda da rede de um ISP. Sendo assim, o agente Adamantos, por deliberação, pode bloquear o roteamento de um endereço IP específico, na entrada da rede do ISP. No sistema Arquitena, isto é realizado por meio de uma ACL (*Access Control List* - Lista de Controle de Acesso). Cada ACL é aplicada no equipamento de rede.

Todos os agentes internos são identificados pelos equipamentos que monitoram, sendo que estes estão em localidades diferentes da rede, e também pelos valores de seus contadores. Cada agente armazena um valor diferente de **FPmed**, **FPmax** e uma valor do nível de segurança, determinado **cns**. Esse valor é a referência para o agente enviar a mensagem com a ACL para o equipamento de rede monitorado, caso o valor seja ultrapassado. O valor de **cns** equivale a 90% do valor de **FPmax**. Outro valor utilizado na fase do bloqueio é o tempo de perduração das ACLs (e.g., 3 segundos). Dessa forma, após o término do intervalo determinado, a ACL pode ser excluída do equipamento se o tráfego for estabilizado, isto é, seja menor do que o **cns**. Após o agente detectar um tráfego anômalo, ele permanece monitorando o equipamento para verificar se o tráfego não se aproxima do **FPmax**.

A partir do momento em que o agente começa a coletar estatísticas do equipamento, ele está habilitado a verificar se o tráfego sumarizado, entre todos os contadores, está próximo do valor de segurança pré-estabelecido (% do **FPmax** = **cns**). Esse valor depende de qual modelo de equipamento ele está monitorando.

Como cada agente tem os contadores com os respectivos endereços IPs dos equipamentos alvo, ele pode direcionar um plano de ação para o bloqueio momentâneo do endereço IP da potencial vítima. A ação de bloqueio consiste na criação de uma ACL para descartar pacotes com o endereço IP da potencial vítima, antes mesmo que passe pelo equipamento de rede. Como o agente monitora continuamente o equipamento, essa mesma ACL pode ser reaplicada repetidamente, na mesma granularidade de tempo, caso o ataque persista. Os agentes que podem participar da fase de bloqueio são os seguintes:

- **Agentes Minos:** extremidade interna da rede
- **Agentes Adamantos na parte central da rede**
- **Agentes Adamantos na borda da rede:** esse agente não precisa esperar que o tráfego alcance algum percentual crítico do equipamento; ao contrário, ele proativamente efetua o bloqueio assim que o tráfego anômalo for identificado, isto é, caso o valor do tráfego seja maior do que **FPmed**.
- **Agentes Éacos na parte central da rede**

A fase de bloqueio contempla também o desbloqueio do tráfego destinado para a potencial vítima, nos equipamentos de qualquer extremidade da rede (e.g., *firewall*, e servidores), e nos equipamentos da parte central da mesma (e.g., roteadores, balanceadores de carga ou *firewalls)*.

Uma vantagem do sistema de bloqueio é que a aplicação e retirada de ACLs é dinâmica, conforme a frequência de ataques. Dessa forma, a velocidade de inclusão e exclusão das regras de bloqueio (ACLs) não seria possível de ser alcançada manualmente por um administrador de rede do ISP, enquanto a

característica distribuída dos agentes facilita esse processo. Esse escopo dinâmico do bloqueio tende a preservar o tráfego legítimo, por 2 motivos:

- O bloqueio consiste somente de intervalos pertinentes ao ataque, de modo que a regra de bloqueio é excluída imediatamente após o fluxo de pacotes tornar-se menor que o valor de **cns** ou quando o intervalo de bloqueio da ACL expirar (e.g., após 3 segundos).
- O bloqueio pode ser realizado o mais longe possível da vítima, possivelmente somente nas bordas da rede. Neste último caso, o tempo de bloqueio deve ser mais curto do que nos demais equipamentos. Esse mecanismo tende a preservar o tráfego legítimo, ao contrário de bloqueios exclusivos nas proximidades do alvo.

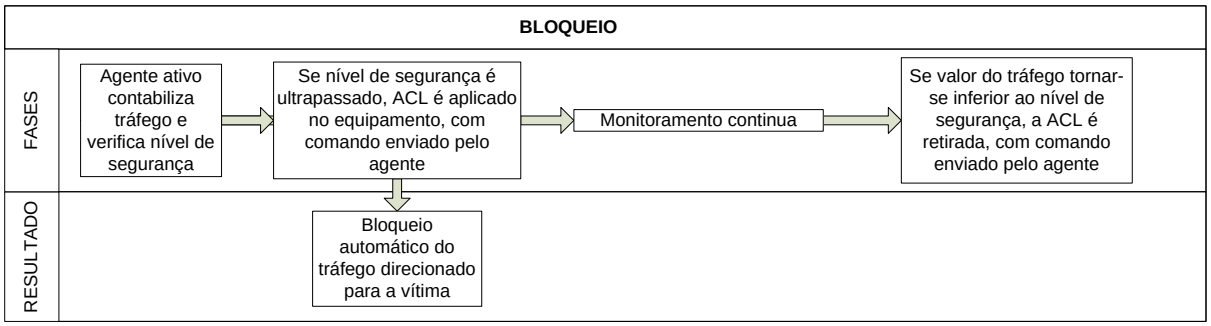
Durante o intervalo de bloqueio, outros equipamentos conectados podem receber uma carga alta de tráfego devido ao redirecionamento do tráfego. Contudo, o processo de bloqueio atua também nos equipamentos que passaram naquele momento a receber uma carga maior de tráfego, tornando o processo de bloqueio intermitente, mas coerente com o aumento anômalo do tráfego.

O sistema de bloqueio pode em um determinado momento bloquear todo o tráfego destinado para um equipamento alvo, e caso não haja redundância, o serviço pode ficar inacessível para os usuários reais e também para os maliciosos. Isto pode ocorrer se alguns equipamentos da região central ou da borda da rede estiverem com as mesmas ACLs simultaneamente, formando uma camada de bloqueio, de modo que o tráfego não percole mais entre a borda e o serviço. Nessas situações, os agentes internos percebem o ambiente e respondem em tempo hábil, a fim de alcançarem os objetivos delegados. Outra característica pertinente dessa fase

de bloqueio é a habilidade social, pois os agentes conseguem o bloqueio do ataque através do compartilhamento de informações, isto é, a cooperação entre si.

Todos agentes Éacos e Adamantos tem diversos valores de **FPmed** em sua base de conhecimento, devido aos horários de pico de tráfego, inerentes a infraestrutura de rede de qualquer ISP. Um plano de ação está associado a cada **FPmed** e para o **FPmax**, e a tomada de decisão de bloqueio além de depender do valor do cns. O diagrama da figura 10 mostra por fases a etapa de bloqueio do sistema Arquitena.

Figura 10 – Diagrama de bloqueio.



4.7. Mitigação do sistema Arquitena

O processo de mitigação engloba desde o processo de monitoramento até o possível bloqueio do tráfego malicioso. No momento em que o primeiro agente identifica tráfego anômalo, todos os vizinhos da hipotética vítima são ativados, quaisquer que sejam os equipamentos correspondentes, iniciando o processo em cadeia. A mitigação consiste em minimizar os efeitos do ataque DDoS, antes mesmo que o bloqueio total do tráfego seja realizado nas extremidades da rede, como a borda ou os hospedeiros dos serviços.

Embora o objetivo de ataques DDoS seja consumir ao máximo os recursos de um servidor ou de um site, nem todos eles são idênticos, havendo uma sutil diferença na forma como estes ataques são perpetrados. Mesmo assim, acredita-se que esta estratégia de mitigação adotada pelo sistema Arquitena seja capaz de minimizar os resultados dos ataques de inundação, mais comuns da camada rede, e também indiretamente os ataques de aplicação.

Planos definidos previamente nos agentes definem qual atitude eles devem executar nos equipamentos específicos. A fim de bloquear qualquer tipo de indisponibilidade nos serviços, o processo de mitigação atua em cada equipamento quando os agentes estão ativos.

4.8. Comunicação entre os agentes do sistema Arquitena

A comunicação entre os agentes do sistema Arquitena se baseia na Teoria do Ato de Fala. As mensagens, por serem enunciados performativos, não se submetem ao critério de verificabilidade, isto é, são enunciados que realizam uma ação. É importante observar que o simples fato de proferir um enunciado performativo não garante sua realização. As circunstâncias devem ser adequadas para que a mensagem tenha um efeito positivo no receptor. A linguagem de mensagens no nível do conhecimento utilizada é a KQML.

No início de cada ciclo de raciocínio, os agentes verificam por mensagens que podem ter recebido de outros agentes. Qualquer mensagem recebida tem a seguinte estrutura: **<enviador, illoc_força, conteúdo>**

O campo **enviador** é o nome do agente, identificado no sistema Arquitena, que enviou a mensagem. Todos os agentes no sistema Arquitena são identificados da seguinte maneira:

- **Minos:** numeração sequencial variando até a quantidade máxima de agentes. Exemplo: Minos1, Minos2, etc.
- **Éacos:** numeração sequencial variando até a quantidade máxima de agentes. Exemplo: Éacos1, Éacos2, etc.
- **Adamantos:** numeração sequencial variando até a quantidade máxima de agentes. Exemplo: Adamantos1, Adamantos 2, etc.

O campo **illoc_força** (força ilocucionária) denota a intenção do enviado, enquanto que o campo **conteúdo** é um termo *AgentSpeak* que varia de acordo com a força ilocucionária (BORDINI; HÜBNER; WOOLDRIDGE, 2007). Segue abaixo exemplos de ações utilizadas:

```
+!initial4 : true <-.my_name(Me);.broadcast(tell, Me);+myPosition(a1).
```

A linha de código acima tem a ação **`broadcast`** que envia mensagem para todos os agentes registrados no sistema. A força ilocucionária **`tell`** implicita a ação de informar ao outro agente do conteúdo, nesse caso o nome do agente enviador (**`Me`**).

4.8.1. Possibilidades de comunicação

Quando o sistema Arquitena é ativado, cada agente questiona o ambiente sobre sua posição e guarda essa informação em sua base de crenças. Com a informação de sua posição e endereço IP do equipamento que monitora, cada

agente interno envia uma mensagem para seus vizinhos imediatos, inclusive os agentes Minos. Entretanto, o agente Minos que receber uma mensagem *multicast* de outro agente Minos deve ignorar tal mensagem, pois ele não tem a necessidade de saber a vizinhança das outras potenciais vítimas, e sim dos agentes vizinhos que possam impedir o fluxo de pacotes direcionados para ele. Adicionalmente, os agentes Minos enviam as informações de sua posição e o endereço IP para todos os agentes do sistema Arquitena. O objetivo dessa mensagem é inserir as informações pertinentes dos possíveis alvos de ataque na base de crenças, bem como compor a base de contadores em cada agente. Novamente, caso o agente Minos seja o receptor desta mensagem de *broadcast*, ele deve descartar essa informação.

Conforme discutido anteriormente, até 24 horas depois da ativação do sistema Arquitena, todos os agentes internos permanecem ativos, sendo que os equipamentos de rede correspondentes devem coletar dados e fornecer os dados ao agente externo. Após o término do primeiro ciclo de 24h, todos os agentes Éacos tornam-se inativos após enviar uma mensagem para o equipamento por eles monitorado, a fim de parar o envio de dados para o agente externo. Em paralelo à mensagem enviada para o agente externo, uma automensagem deve ser enviada a fim de tornar o agente Éacos inativo. Nesse momento, os agentes Éacos tornam-se inativos até entrarem no ciclo de ativação. Os outros agentes Adamantos e Minos continuam com o processo contínuo de monitoramento e, portanto, permanecem ativos.

Outra possibilidade de comunicação ocorre entre qualquer agente e o equipamento que ele está monitorando, pois em algum momento o agente pode enviar uma mensagem de bloqueio para um endereço IP específico. Após o término

do tempo pré-determinado, o agente pode se comunicar novamente, caso o tráfego tenha retornado ao normal, para retirar a ACL do equipamento. A tabela 5 tem as possíveis mensagens trocadas entre os agentes.

Tabela 5 – Possíveis mensagens trocadas.

Enviador	Receptor	Código	Função
Todos agentes	Todos agentes	`.send(W,askOne,online(X),online(X))`	Verifica se agente está ativo
Todos agentes	Todos agentes	`.send(W,achieve,statusOnline_Fromleader(Name))`	Muda estado do agente para ativo
Todos agentes	Todos agentes	`.send(W,askOne,normaltraffic(Traffic, normaltraffic(Traffic))`	Verifica estado da variável Traffic
Todos agentes	Todos agentes	`.send(W,askOne,wakeUp(Name),wakeUp(Name))`	Verifica estado da variável Name
Todos agentes	Todos agentes	`.send(W,achieve, statusOffline(Name))`	Muda estado do agente para inativo
Todos agentes	equipamento	`arquitena.ativaRoteador("IP_address", 23, "user", "password")`	Comunicação com o equipamento pela porta 23

4.9. Arquitetura de implantação

Os múltiplos agentes podem estar dispostos de três maneiras, conforme explanação a seguir:

1) **Plataforma distribuída:** Cada agente é alocado diretamente nos equipamentos e servidores da rede do ISP. Deste modo, todo o processamento referente ao agente e sua comunicação é de

responsabilidade do equipamento. A comunicação entre os agentes do sistema acontece através da infraestrutura de rede já existente. A escalabilidade neste caso é maior quando se compara com a disposição centralizada dos agentes, porém os equipamentos devem estar configurados ou ter hardware compatível para processar o software específico do sistema Arquitena. Este modo de alocação dos agentes é uma possibilidade que pode tornar-se especialmente factível com a utilização de tecnologias de redes definidas por software (MCKEOWN et al., 2008), em que se tem maior controle sobre as funcionalidades que são executadas no equipamento.

2) **Plataforma centralizada:** Todos os agentes são alocados em uma estrutura centralizada de hardware (disco, CPU e memória). Neste cenário, nenhum processamento referente aos agentes é executado nos equipamentos monitorados. A comunicação entre os agentes do sistema é mais simples quando comparada ao cenário distribuído, devido ao compartilhamento do mesmo hardware. Assim, a única comunicação externa ocorre com o equipamento monitorado. O limite da escalabilidade é, então, definido pelos recursos do servidor no qual o sistema é executado. Nesse caso, o sistema criado é um ambiente virtual de toda a rede monitorada, com as mesmas conexões e equipamentos, sendo que cada equipamento tem um agente responsável pelo monitoramento.

3) **Plataforma centralizada, mas com hardware distribuído:** A única diferença deste cenário quando comparados ao de plataforma centralizada é que a estrutura de hardware é distribuída entre diversos computadores. Nesse caso,

a escalabilidade é maior do que a plataforma centralizada, mas há uma carga adicional relativa à comunicação entre os agentes, pois os mesmos estão sendo executados em computadores diferentes.

Dentre essas abordagens, há pelo menos duas infraestruturas diferentes para executar um SMA:

- Jason (processamento centralizado em único servidor) (BORDINI; HÜBNER; WOOLDRIDGE, 2007)
- Jade (processamento distribuído entre servidores) (JADE, 2014)

Em seu estado atual, o sistema Arquitena se apoia na abordagem centralizada de hardware para executar as simulações no ambiente Delos. Esta abordagem é adotada porque, assim, o ambiente está disponível para tratar as requisições dos agentes concorrentemente, sem onerar o processamento do computador com a comunicação entre os agentes. O estudo realizado em (FERNANDÉZ et al., 2010) mostra que o custo de comunicação pode onerar o processamento quando comparado com a estrutura centralizada. Deste modo, o sistema Arquitena e o ambiente Delos foram desenvolvidos em Jason.

Cabe notar que, como o Jason é uma extensão da linguagem AgentSpeak(L), está implícito na adoção desta linguagem a utilização do modelo BDI, a linguagem KQML de comunicação e o sistema de raciocínio PRS. A arquitetura de coordenação dos agentes está baseada no PPG (Planejamento Parcial Global) que consiste no ato dos agentes alinharem suas submetas, de modo a não conflitarem acidentalmente na execução dos planos (WOOLDRIDGE, 2009).

4.10. Vantagens do sistema Arquitena

Nessa seção estão descritas as funcionalidades relevantes do sistema Arquitena que podem ser ressaltadas como vantajosas. Além destas, há outras vantagens que estão destacadas nas seções 5.6.1 e 5.6.2.

Primeiramente, devido à característica esparsa do Arquitena, ele pode ser considerado um sistema de defesa ubíquo, pois contempla sua implantação em todos os pontos da rede do ISP. Desta forma, ele tem a vantagem no fato de monitorar a região próxima dos potenciais alvos, que é o melhor ponto para saber quando um ataque se inicia (MIRKOVIC et al., 2006). Ao mesmo tempo, o sistema evita problemas relativos a essa localização, como o alvo estar em uma posição da rede no qual não consiga desempenhar ações que requerem complexa análise e diferenciação do tráfego legítimo. Além disso, o ataque DDoS por definição pode sobrecarregar por completo a vítima, ao menos que o mecanismo de defesa possa suportar tal volume de tráfego.

O sistema de defesa proposto desfruta também das vantagens pertinentes à localização próxima da origem do ataque, onde os fluxos de ataque DDoS tendem a não ser altamente agregados. Isso permite que mais processamento seja devotado para detecção do ataque, ao invés de dedicação exclusiva à mitigação do ataque DDoS. Outro ponto a ser destacado é que, quanto mais próximo da origem, mais fácil de separar o tráfego legítimo do malicioso (MIRKOVIC et al., 2006).

O Arquitena pode detectar ataques ou tráfego anômalo interno oriundo da própria rede do ISP, por ter agentes implantados na região central da rede. Esses agentes monitoram os equipamentos com o maior fluxo de pacotes, normalmente

chamados de "*core routers*" bem como equipamentos adjacentes. O sistema de defesa Arquitena, para aproveitar essas características, apresenta diversos agentes deliberativos que monitoram os equipamentos de rede de forma cooperativa.

Paralelamente ao processo de mitigação, o Arquitena pode criar rotas de rotas de ataque dentro da rede do ISP. Com as rotas de ataques definidas, um aprovisionamento pode ser determinado em relação ao sistema de defesa implantado. Portanto, a comparação entre as rotas de ataque pode determinar o padrão das rotas internas, e também permitir estudos para tornar a rede do ISP mais resiliente.

Nesse contexto, o sistema Arquitena tem o potencial de detectar e bloquear ataques de inundação de pacotes rapidamente. Como os protocolos comuns utilizados, por exemplo, Netflow e OSPF, são de classe comum à maioria dos fornecedores de equipamentos de rede, como a Cisco ou Juniper, o sistema Arquitena poderia ser aplicado em conjunto com esses equipamentos.

4.11. Limitações do sistema Arquitena

O Arquitena não foi projetado para mitigar diretamente ataques de vulnerabilidade, ou seja, que trabalham enviando mensagens especialmente construídas para tirar vantagem de uma aplicação alvo vulnerável. Esses ataques estão relacionados com a camada de aplicação e tentativas de consumir as tabelas de estados presentes em muitos equipamentos de infraestrutura. Assim, a atuação do Arquitena é complementar à de soluções que visam prevenir a exploração de tais vulnerabilidades, como *firewalls*.

A escalabilidade da quantidade de equipamentos monitorados pelos agentes depende da capacidade de processamento e quantidade de memória presentes no único servidor que executa o Arquitena. Caso a estrutura de rede do ISP cresça, esse servidor também deve ser escalável a fim de manter o sistema Arquitena funcional.

4.12. Requisitos de hardware e software para implantação do sistema Arquitena

Para o sistema Arquitena trabalhar corretamente são necessários os seguintes elementos:

- **Protocolos de rede aberta** para gerar registros do fluxo de pacotes no equipamento, qualquer que seja ele (e.g., roteador, firewall, balanceador de carga ou switch). Os protocolos mais comuns são: Netflow (NETFLOW, 2014) (Cisco), IPFIX (*Internet Protocol Flow Information eXport* - Informação do Fluxo do Protocolo da Internet Exportado) (IPFIX, 2013) (padrão IETF (*Internet Engineering Task Force* - Força Tarefa de Engenharia da Internet)), JFlow (JFLOW, 2013) (Juniper). Os equipamentos devem ser capazes de habilitar tal protocolo e, assim, se comunicar com o coletor.
- **Um coletor de dados** para receber e analisar os registros enviados pelos equipamentos de rede. Assim, ele consegue enviar estatísticas de acordo com a amostragem requisitada.
- **Servidor** capaz de executar as simulações do ambiente Delos e todos os agentes pertencentes ao sistema Arquitena, o que requer que as linguagens de programação Jason e Java estejam disponíveis. Esse

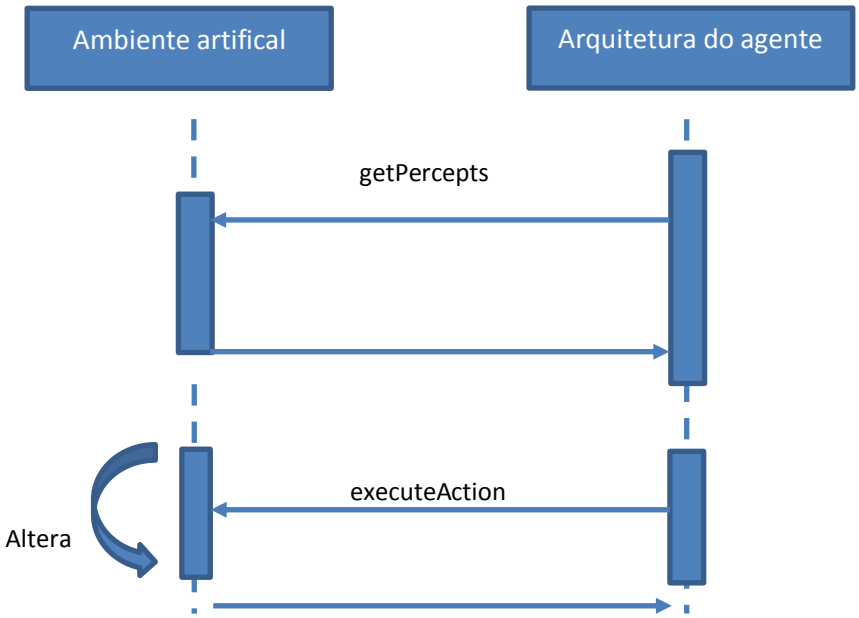
mesmo servidor pode executar o coletor de dados dos equipamentos de rede. O servidor deve ter acesso a toda a rede do ISP, pois cada agente do sistema Arquitena eventualmente envia comandos para o equipamento monitorado.

4.13. Ambiente de simulação Delos

O conceito chave de agentes autônomos é que eles estão situados em um ambiente (BORDINI; HÜBNER; WOOLDRIDGE, 2007). No Arquitena, o ambiente é compartilhado por múltiplos agentes, de modo que a ação de um agente tende a interferir nas ações dos outros agentes. Assim, para análise do Arquitena foi criado um modelo computacional para simular a rede real do ISP que se deseja monitorar, simulando os aspectos dinâmicos do ambiente real. Os agentes implementados operam nesse ambiente simulado, denominado Delos, agindo sobre ele e percebendo suas propriedades. No processo de simulação no ambiente Delos, cada plano selecionado corresponde a uma ação no ambiente real. No sistema Arquitena, cada agente deve interagir com o ambiente (veja figura 11) e, possivelmente, diretamente com outros agentes através da comunicação baseada no Ato da Fala.

O ambiente virtual Delos deve ser previamente montado usando as informações oriundas do próprio ISP, com o posicionamento e interconexões atuais dos equipamentos que compõem sua rede.

Figura 11 – Interação entre ambiente e um agente (BORDINI; HÜBNER; WOOLDRIDGE, 2007).



O ambiente Delos pode ser constituído pela combinação dos seguintes componentes:

1. **Agente Minos**
2. **Agente Éacos**
3. **Agente Adamantos**
4. **Equipamento Roteador:** equipamento que representa efetivamente um equipamento de rede do tipo roteador. Esse equipamento é monitorado por um agente Éacos ou Adamantos, caso ele faça parte da rota de ataque.
5. **Equipamento *Firewall*:** equipamento que representa efetivamente um equipamento de segurança do tipo *firewall*. Esse equipamento é monitorado por um agente Éacos, Minos ou Adamantos, caso ele faça parte da rota de ataque.

6. **Equipamento Servidor:** equipamento que representa efetivamente um equipamento do tipo servidor. Esse equipamento sempre é monitorado por um agente Minos, mesmo antes de um hipotético ataque DDoS.
7. **Equipamento Balanceador de carga:** equipamento que representa efetivamente um equipamento do tipo balanceador de carga. Esse equipamento sempre é monitorado por um agente Minos, mesmo antes de um ataque DDoS.

Um exemplo das posições adjacentes é mostrado na tabela 6, que associa uma posição do ambiente ao endereço IP.

Tabela 6 – Exemplo de tabela de adjacências de cada agente.

Endereço IP	Posição na grade (ambiente Delos)
192.168.10.32	A1
X	A2
X	A3
192.168.10.33	B1
192.168.10.34	C1
192.168.10.49	C2

5 ANÁLISE EXPERIMENTAL

Com o objetivo de avaliar a eficácia do sistema Arquitena na detecção e mitigação de ataques DDoS, a solução foi implementada e testada sob diferentes aspectos. Nesta seção são descritos os protocolos alvos dos testes, os recursos utilizados nos mesmos e as métricas definidas. Na subseção 5.6, são mostrados os testes finais do sistema Arquitena e a comparação com outras técnicas de defesa citadas nesta dissertação.

5.1. Alvos

A determinação de valores de amostragem adequados deve ser feita de forma a permitir um balanceamento entre a eficácia do processo de detecção e o consumo dos recursos nos equipamentos. Equipamentos exclusivos para detecção e mitigação, como o Peakflow (PEAKFLOW, 2014), tem como métrica a amostragem de 1000 pacotes, de modo que este valor é também adotado no sistema Arquitena.

Durante os testes realizados, foram utilizados servidores WEB no papel das vítimas perante um ataque DDoS controlado usando pacotes HTTP. Esta escolha é motivada pela pesquisa realizada pela Arbor (ARBOR, 2014), na qual foi relatado que os alvos mais comuns de ataques na camada de aplicação são HTTP e DNS: percentual de ataque alcança 82% e 77%, respectivamente.

5.2. Recursos

Os elementos utilizados nos testes com o sistema Arquitena e que constituem o conjunto para criação de uma cenário de teste são listados a seguir. Esse cenário é baseado no ambiente descrito em *Benchmark for DDoS Evaluation* (MIRKOVIC et

al., 2006), artigo que descreve quais são os elementos mínimos para uma boa referência de teste com ataques DDoS:

- recursos para criar um ataque DDoS (tráfego malicioso)
  - cenários típicos de ataques
- recursos para criar tráfego legítimo, a fim de competir com o tráfego malicioso
- topologias diferentes de rede

O método de análise se baseia nos ataques que trabalham com o envio de um grande número de pacotes para um determinado alvo, devido ao sistema Arquitena focar na detecção e mitigação de ataques baseados na inundação de pacotes, excrescente ao tráfego normal.

Em todas as simulações, a estrutura centralizada foi utilizada. Todos os agentes estavam no mesmo servidor e não foram executados nos equipamentos de rede monitorados.

Para os testes iniciais foram utilizados os softwares de emulação Dynamips (DYNAMIPS, 2014) e o software Dynagen (DYNAGEN, 2014) para configurar e emular os roteadores, *firewalls* e *switches* e executar simultaneamente várias instâncias de roteadores virtuais. Para visualização da topologia emulada foi utilizado o software GNS3 (GNS3, 2014). Nos roteadores emulados pelo Dynamips/Dynagen, o protocolo de rede aberta Netflow foi habilitado para gerar registros Netflow e exportá-los para o coletor Netflow (agente externo). Na tabela 7 é dada a descrição dos roteadores emulados.

Tabela 7 – Descrição dos roteadores emulados.

Roteador	**Modelo**	**IOS (*Internetwork Operating System* (Sistemas Operacionais de Inter-rede)**	**Quantidade**	**Simulação**
Cisco	7200	c7200-a3jk9s-mz.122.23.bin	12	Local

Os roteadores emulados foram configurados com o protocolo de roteamento OSPF (*Open Shortest Path First* - Primeiro Caminho Mais Curto Aberto), para estabelecer a comunicação entre eles. Foi utilizado um computador físico para desenvolvimento do sistema Arquitena e criação do ambiente virtual de simulação Delos. Um servidor virtual foi utilizado para emulação da rede e instalação do coletor Netflow. Dois servidores WEB foram instalados para receberem o tráfego em máquinas virtuais. Outros 10 servidores virtuais (tabela 8), com 10 interfaces de rede (1GB) cada um, foram utilizados para gerar tráfego legítimo e malicioso. Em todos os roteadores emulados, foi configurado o Cisco IOS EEM (*Embedded Event Manager* – Gerenciador de Eventos Embutido). Com essa funcionalidade ativa, os comandos nos roteadores como o bloqueio de interfaces específicas, ACLs, habilitação do Netflow e Cache Flow, ficam prontos para serem executados. Na Tabela 9 são listadas as características de hardware dos computadores que foram utilizados nas simulações.

Para efetuar os testes de estresse, utilizou-se a ferramenta T50 (T50, 2013), capaz de injetar pacotes TCP, UDP e ICMP. Como o T50 só tem binários para sistemas operacionais Linux ou Unix, dez máquinas virtuais foram instaladas com o sistema operacional Ubuntu. Através do T50, foi possível determinar a quantidade de

pacotes injetados na rede, o que facilitou a determinação do percentual de falso-positivos, falso-negativos, o cálculo do fluxo de pacotes individualmente e o tempo de resposta do sistema Arquitena, nas simulações propostas.

Tabela 8 – Descrição dos computadores usados nos testes.

Estado	Sistema instalado	CPU	Memória RAM	Quantidade	Sistema Operacional	Interfaces de rede 1GB
Físico	Arquitena e Delos	IntelCore i7-2630QM 8x2.0GHz	8GB	1	Windows 7 SP1 64 bits	1
Virtual	Emulação, Netflow, coletor Netflow e servidores WEB	4 vCPUs	6GB	1	Windows 7 SP1 32 bits	1
Virtual	Injetor de tráfego T50	2 vCPUs	2GB	10	Ubuntu 12.04 64 bits	10

Para os testes dos limites de hardware de cada dispositivo, devem ser conhecidos os seguintes limites operacionais a fim de evitar que os dispositivos colapsem durante o ataque DDoS controlado: a vazão máxima; o número máximo de sessões concorrentes; o número de conexões por segundo; e a quantidade de pacotes por segundo. Os testes avaliaram se o sistema Arquitena é efetivo em detectar e bloquear o ataque DDoS, antes mesmo do colapso do servidor (vítima) e dos equipamentos centrais e da borda da rede.

Os recursos que foram utilizados para as simulações e testes do sistema Arquitena são listados a seguir:

<u>Software:</u>

1) 10 máquinas virtuais com sistema operacional Ubuntu versão12.04.

2) Ambiente de Desenvolvimento Integrado: Eclipse versão Indigo.
3) Linguagem de programação: Jason versão 1.3.6.
4) Linguagem de programação: Java versão SDK 7
5) Netflow: protocolo habilitado nos equipamentos da rede emulada
6) Programa de análise de rede e captura de pacotes: Winpcap versão 4.1.2
7) Servidor WEB: Apache versão 9.2
8) Software de comunicação com o Hypervisor Dynamips: Dynagen versão 0.11.0
9) Software de virtualização: VMware vSphere 5.1
10) Software de visualização das topologias de rede emuladas: GNS3 versão 0.8.3
11) Software emulador: Dynamips versão 0.2.8
12) Software injetor de pacotes: T50 versão 5.3
13) Analisador Netflow 9.0
14) Cisco *IOS Embedded Event Manager* (EEM) 4.0

Hardware:

1) 1 notebook

5.3. Rede emulada

A rede de testes utilizada foi composta de roteadores emulados, servidores WEB, *switches* virtuais, injetor de pacotes e coletor/analisador de tráfego, conforme descrito na seção 5.2. Em específico, os testes serviram para verificar a quantidade de pacotes que alcança cada potencial vítima (servidores), e se o **FPmed** calculado estava de acordo com o tráfego real em cada servidor.

A fim de determinar quais equipamentos de rede seriam monitorados constantemente, além das hipotéticas vítimas, foi utilizada uma medida de centralidade como *betweenness* (WHITE; BORGATTI, 1994) (intermediação), tendo como parâmetro a conectividade de cada nó na rede. Através desse parâmetro, pode-se determinar quais dos nós em uma rede tem maior fluxo de pacotes passante. Os nós com maior valor na intermediação são também monitorados constantemente. Esses valores calculados foram testados e comparados com um tráfego de testes, a fim de verificar a real necessidade de se ter mais elementos monitorados dentro da rede do ISP, além das hipotéticas vítimas.

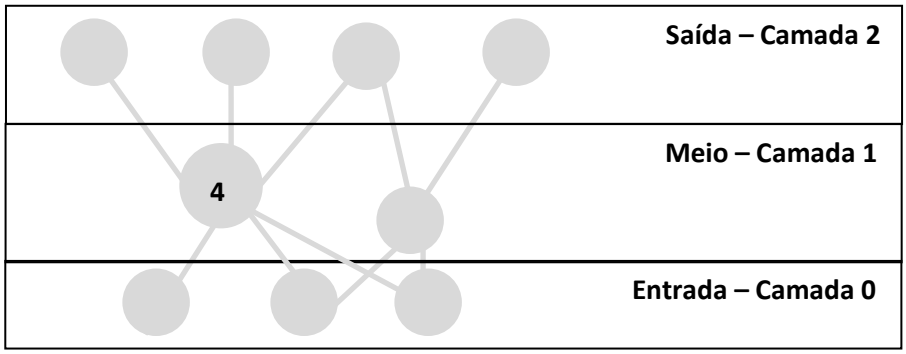
O cálculo de intermediação de certo nó **n**, pode ser realizado a cada 2 nós **s** e **t** conectados ao nó **n**, dado pela seguinte expressão (WHITE; BORGATTI, 1994):

$$\mathbf{Ci(n) = \Sigma_{s \neq t \neq n} (\sigma st\ (n) / \sigma st)}$$

sendo **σst** o número total de caminhos mais curtos do nó **s** para o nó **t**, e **σst** (n) o número desses caminhos que passam através de **n**.

Outro método para se comparar os testes realizados foi o cálculo do fluxo médio (**FPmedC**). Assim, a rede de testes foi dividida em camadas, sendo que cada camada teve como base de distância o número de salto(s) a partir dos nós de entrada. Dessa forma, a borda da rede está na camada 0 e o nível das potenciais vítimas está na camada **N-1**, sendo **N** o número total de camadas ou níveis. A Figura 12 ilustra os níveis em uma rede de exemplo.

Figura 12 – Exemplo de rede em três camadas.



Nesse cálculo, os parâmetros a serem usados são: o fluxo de entrada de pacotes e conectividade de entrada em cada nó da rede. Considerando X como o fluxo total de pacotes que entram na camada zero e sucessivamente nas outras camadas, o fluxo do nó 4, por exemplo, é dado pela expressão

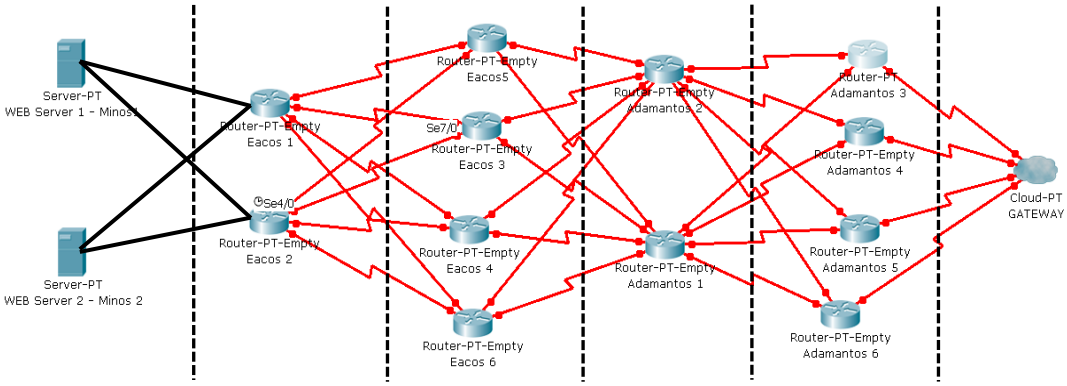
PFmedC4 = (número de entradas no nó 4 / número de entradas na camada 1) * X

Portanto, tem-se que **PFmedC4 = 3X/5**.

Cabe ressaltar o valor de **FPmedC** (média do fluxo de pacotes calculado) deve sempre ser menor do que o **FPmax**, estipulado pelo fabricante do equipamento.

Seguindo essa definição de centralidade, foi definida a topologia de testes mostrada na Figura 13, com apenas 2 roteadores na centralidade da rede sendo monitorados pelos agentes Adamantos 1 e 2.

Figura 13 – Topologia de testes.



A estrututura da rede de um ISP geralmente é dividida em três camadas, sendo uma de acesso, outra de distribuição e a do centro da rede onde se concentra os equipamentos de rede mais robustos. Há também a camada de serviços com os equipamentos fornecedores de serviço e a camada da borda, onde todo o tráfego legítimo e malicioso entra. A quantidade de equipamentos utilizada nessa topologia de testes não é fiel a de um ISP, mas dado à arquitetura de ISPs e salvo as devidas proporções, a topologia é uma boa referência. Assim como em DefCOM (MIRKOVIC; ROBINSON; REIHER, 2003), a topologia de testes também foi dividida em 3 níveis. Mas em (GUPTA; MISRA; JOSHI, 2008) foi utilizado uma topologia estrela com 1 servidor, 4 roteadores e 4 computadores atacantes. Em (MALIALIS; KUDENKO, 2013) foi utilizado uma topologia com 20 nós, e de forma dicotômica um nó da rede recebe o tráfego e distribui pelas suas 2 interfaces para os outros 10 nós de cada lado. No arcabouço de defesa Cooperativa (KOTENKO; ULANOV, 2007) utilizou-se 10 roteadores e 10 servidores clientes para realizar os testes.

As medidas de centralidade e o cálculo médio de tráfego para essa rede são mostrados na Tabela 9. Para o cálculo de tráfego médio foi utilizado um milhão pacotes por segundo (pps) na entrada do roteador central (gateway).

Tabela 9 – Valores de centralidade e cálculo médio de tráfego.

AGENTE	Centralidade	PFmedC (mil pps)
Adamantos 1	0,25	500
Adamantos 2	0,25	500
Adamantos 3	0,0054	250
Adamantos 4	0,0054	250
Adamantos 5	0,0054	250
Adamantos 6	0,0054	250
Éacos 1	0,021	500
Éacos 2	0,021	500
Éacos 3	0,095	250
Éacos 4	0,095	250
Éacos 5	0,095	250
Éacos 6	0,095	250

5.4. Métricas

Durante as simulações, as seguintes métricas foram avaliadas:

- Porcentagem de detecção (verdadeiramente positivos) – **PDET**.
- Tempo levado pelo sistema para detectar um ataque DDoS, baseado nos patamares previamente definidos - **TDAD**.
- Tempo levado pelo sistema para bloquear um ataque DDoS - **TABD**.
- Porcentagem de falso-positivos, em tempo determinado - **PFPO**.
- Porcentagem de falso-negativos, em tempo determinado - **PFNE**.

Além das métricas definidas, verificou-se se o **FPmed** calculado (**FPmedC**), para cada nó da rede, é factível com a realidade de um ataque DDoS.

5.5. Fases dos testes

Para realizar esses testes, os seguintes passos foram seguidos:

1) Determinar a estrutura de rede:

a. Número de nós

b. Conexões de entrada e saída

c. Protocolo de roteamento: OSPF

2) Simular e obter o fluxo típico de pacotes para cada nó

a. Conhecer o fluxo de pacotes

Considerando a implantação de um sistema real de defesa na rede de qualquer ISP, a coleta de dados para medir o fluxo de pacotes em cada nó é obrigatória. Porém, como para esses testes não foi realizado a medição do fluxo de pacotes de um ISP real, foram utilizados três métodos para avaliação dos resultados obtidos:

1) Medir o fluxo de pacotes em cada nó da rede emulada, antes mesmo da ativação do sistema Arquitena, e o intervalo de tempo durante o qual o serviço fornecido fica indisponível após o início do ataque DDoS controlado. Esses valores de tempo são então comparados com os obtidos após a ativação do sistema Arquitena, caso o ataque consiga desativar o serviço fornecido.
2) Comparar os valores de falso-positivos, falso-negativos e porcentagem de detecção, obtidos pelo Arquitena com aqueles de outros métodos descritos neste texto.

3) Verificar se há diferença no tempo de detecção e bloqueio quando se comparam os cenários (1) em que todos os agentes permanecem ativos e (2) em que os agentes Éacos permanecem inativos até que sejam ativados durante o processo de detecção.

5.6. Resultados obtidos

A fase de simulações foi dividida em diferentes formas de ataque, em termos da quantidade e localização dos atacantes e separados também pelos protocolos TCP e UDP. Em todas as simulações, foi necessário habilitar o Cisco IOS EEM nos roteadores emulados. Dessa forma, quando um agente se comunica com o roteador, ele pode capturar o cache do fluxo no próprio equipamento e habilitar os filtros. Uma ação interna foi criada para cada agente, em Java, para executar tais comandos, acessando-o através da porta 23 (Telnet).

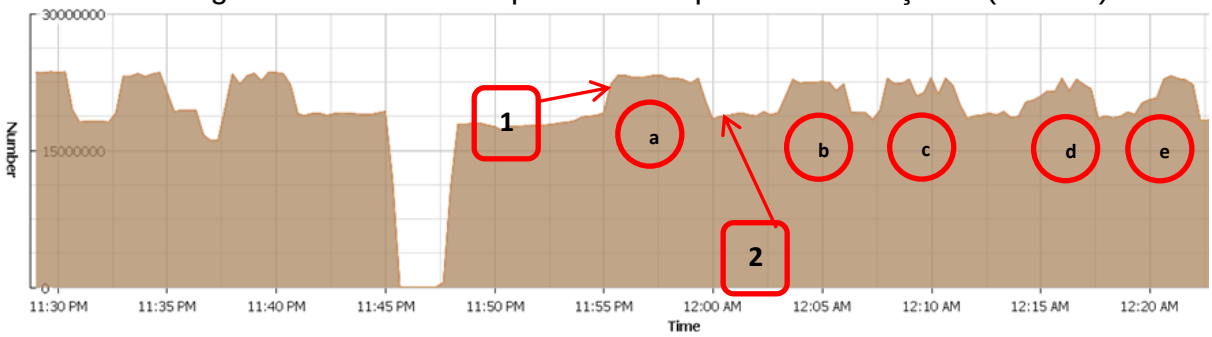
Outra premissa dos testes foi que, nos roteadores de borda, apenas o tráfego real era permitido, filtrando-se endereços IPs desconhecidos. Assim, apenas clientes do mesmo AS podem usufruir dos recursos disponíveis. O protocolo de roteamento OSPF foi configurado nos 12 roteadores emulados, de modo a estabelecer comunicação entre todos eles. Além disso, o tráfego gerado foi configurado de tal forma que, em todos os roteadores, sempre havia um fluxo ativo de tráfego legítimo. As características de cada simulação são descritas na tabela 10.

Tabela 10 – Especificações das simulações realizadas.

Teste ID	Quantidade de injetores (legítimo)	Quantidade de injetores (malicioso)	Tráfego entrante legítimo	Tráfego entrante malicioso	Repetições
1	1	10	Adamantos3	Adamantos5	10
2	1	10	Adamantos5	Adamantos5	10
3	1	20	Adamantos3	Adamantos5 (10) e 6 (10)	10
4	1	20	Adamantos5	Adamantos5	10
5	5	50	Adamantos3 (5)	Adamantos4 (30), 6(10) e 5 (10)	10
6	5	50	Adamantos3 (2), 4 (1), 6 (1) e 5 (1)	Adamantos3 (30), 5 (10) e 4 (10)	10
7	1	90	Adamantos3	Adamantos5	10
8	1	90	Adamantos5	Adamantos5 (40), 6 (30), 3 (10) e 4 (10)	10

Nos testes 1, 3, 5 e 7, o comportamento foi similar, pois o tráfego legítimo que foi injetado no roteador de borda Adamantos era exclusivo, enquanto que o tráfego malicioso foi injetado nos outros roteadores. Como resultado, o tráfego não foi mesclado nos roteadores de borda e o bloqueio do tráfego malicioso foi facilmente executado após exceder o padrão do equipamento. Após 10 repetições dos testes, o tráfego médio resultante apresentou o comportamento mostrado na figura 10. Com relação ao volume do teste 1, o pico do tráfego no roteador Adamantos5 foi de 380 mil pacotes por segundo (pps) (ponto 1 da Figura 14). Após o bloqueio, o tráfego voltou ao fluxo normal com 312 mil pps (ponto 2 da Figura 14).

Figura 14 – Gráfico de picos de ataques na simulação 1 (volume).



A Figura 15 mostra os mesmos picos de ataque da Figura 14, mas com o viés do fluxo de pacotes no roteador Adamantos5. O tráfego nos roteadores Adamantos3, 4 e 6, não foi afetado. Os valores obtidos, nos testes 1, 3, 5 e 7 (Tabela 10), para as métricas discutidas na seção 5.4 são os mostrados na Tabela 11.

As figuras 14 e 15 mostram o tráfego de pacotes acumulado a cada minuto no agente Adamantos5, durante o ataque DDoS controlado. Nesse ataque, o sistema Arquitena manteve um padrão de tempo no bloqueio e desbloqueio do tráfego no equipamento de borda. Com 10 injetores de tráfego malicioso, os valores dos picos de tráfego foram similares, e houve apenas divergência na duração do tempo de permanência do ataque controlado DDoS, pois por exemplo, o pico (a) ficou 5 minutos enquanto que o pico (e) permaneceu durante 2,5 min. Nesse caso, o sistema Arquitena foi mais rápido para detectar o ataque e bloqueou o tráfego no equipamento monitorado pelo agente Adamantos5.

Figura 15 – Gráfico de picos de ataques na simulação 1 (fluxo).

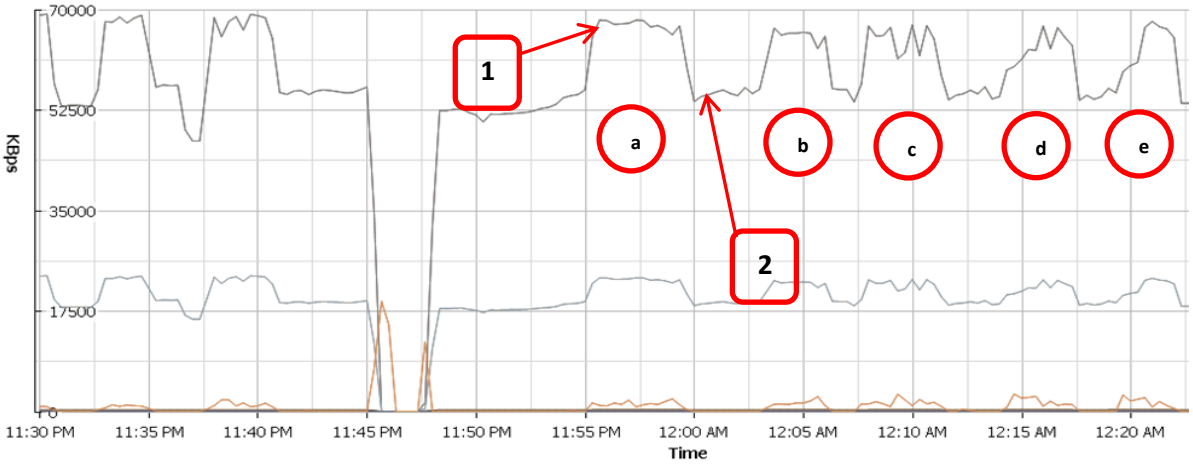


Tabela 11 – Resultados da simulação1.

Métrica	Calculado	Medido (TCP/UDP)	Medido (TCP/UDP) com fluxo de ativação	Agente/Sistema
FPmed	250 mil	312 mil	312 mil	Adamantos5
PDET	xx	70%	70%	Arquitena
TDAD	xx	5 seg	5 seg	Arquitena
TABD	xx	a – 4 min b – 2,5 min c – 3 min d – 2 min e – 2 min	a – 4 min b – 2,5 min c – 3 min d – 2 min e – 2 min	Arquitena
PFPO	xx	1,5%	1,5%	Arquitena
PFNE	xx	11%	11%	Arquitena
Tráfego legítimo	xx	100%	100%	Adamantos3

No teste 1, os seguintes agentes sofreram as ações a seguir:

- Adamantos1 – tráfego de ataque distribuído
- Adamantos2 – tráfego de ataque distribuído
- Adamantos3 – tráfego normal

- Adamantos4 e 6 – sem tráfego
- Adamantos5 – bloqueio intermitente
- Éacos1 e 2 – ficaram ativos por acionamento do Minos1
- Éacos3, 4, 5 e 6 - tráfego de ataque distribuído (não ficaram ativos)
- Minos1 – tráfego excedido (acionou vizinhos imediatos)

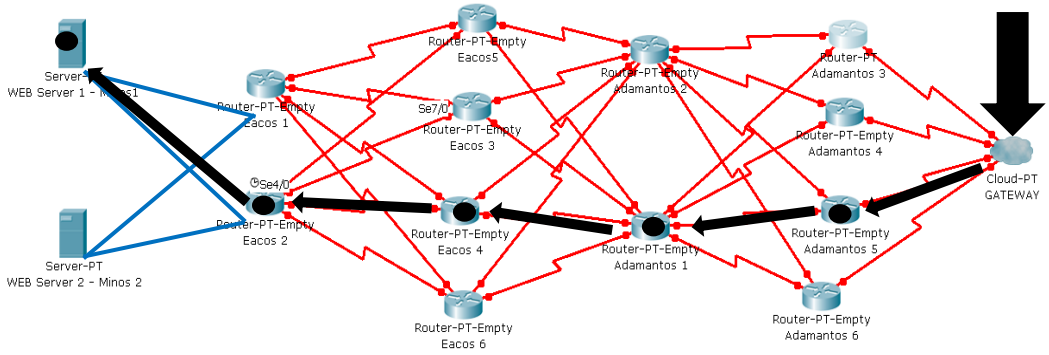
Nos testes 1, 3, 5 e 7 foi realizado medições da utilização de CPU depois do início do ataque. Antes da implantação do sistema Arquitena a porcentagem de CPU tinha a média de aproximadamente 10%, sem tráfego, e após a implantação do Sistema Arquitena a taxa de processamento subiu apenas 2% na fase de inicialização, quando o número de mensagens é maior. A tabela 12 mostra a taxa de utilização de CPU nas fases de ataque e bloqueio, no agente Adamantos5, de acordo com as figuras 14 e 15.

Tabela 12 – Utilização de CPU (testes ímpares).

Posição no tempo	**Comando no roteador: #*show processes***
1	CPU utilization for five seconds: 93%; one minute: 91%; five minutes: 90%
2	CPU utilization for five seconds: 50%; one minute: 48%; five minutes: 47%
a	CPU utilization for five seconds: 90%; one minute: 88%; five minutes: 87%
b	CPU utilization for five seconds: 92%; one minute: 89%; five minutes: 88%
c	CPU utilization for five seconds: 90%; one minute: 86%; five minutes: 85%
d	CPU utilization for five seconds: 90%; one minute: 86%; five minutes: 85%
e	CPU utilization for five seconds: 90%; one minute: 84%; five minutes: 83%

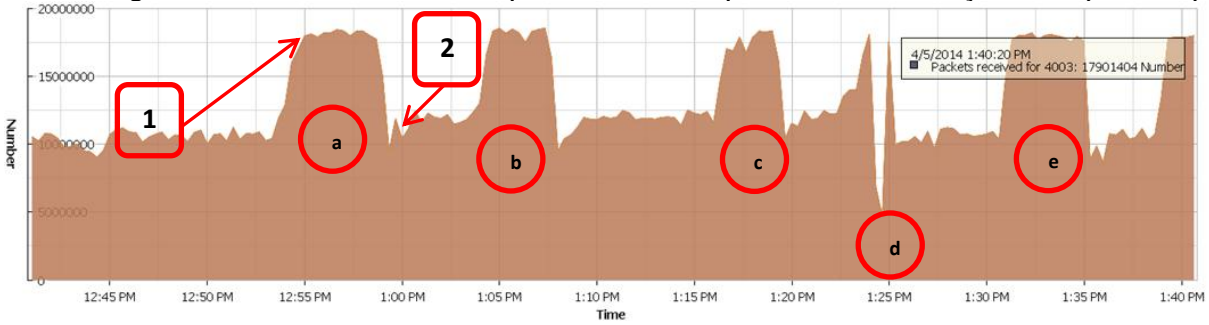
Após as 10 simulações, do teste 1, a rota de ataque determinada foi a mostrada na Figura 16.

Figura 16 – Rota de ataque mais frequente (teste 1).



Nos testes 2, 4, 6 e 8, o comportamento foi similar, pois o tráfego legítimo que foi injetado no roteador de borda Adamantos sempre estava concorrendo com o tráfego malicioso. O tráfego foi medido em cada equipamento para verificar a preservação do tráfego legítimo, em caso de bloqueios intermitentes. Após 10 repetições dos testes, o tráfego médio resultante apresentou o comportamento mostrado na Figura 17. Com relação ao volume do teste 8, o pico do tráfego no roteador Adamantos5 foi de 308 mil pacotes por segundo (pps) (ponto 1 da Figura 17). Após o bloqueio, o tráfego voltou ao fluxo normal com 167 mil pps (ponto 2 da Figura 17). A volumetria do tráfego foi menor em comparação à obtida nos testes 1, 3, 5 e 7, pois o limite da interface da máquina virtual onde estava sendo executado o sistema Arquitena foi alcançado rapidamente, que passou a não conseguir mais encaminhar os pacotes entrantes.

Figura 17 – Gráfico de picos de ataques na simulação 8 (volume).



A Figura 18 mostra os mesmos picos de ataque da Figura 17, mas com o viés do fluxo de pacotes no roteador Adamantos5. O tráfego nos roteadores Adamantos3, 4 e 6, foi afetado devido ao tráfego malicioso também entrar nesses roteadores. Os valores obtidos para as métricas discutidas na seção 5.4, nos testes 2, 4, 6 e 8 (Tabela 10), foram mostrados na Tabela 13.

Figura 18 – Gráfico de picos de ataques na simulação 8 (fluxo).

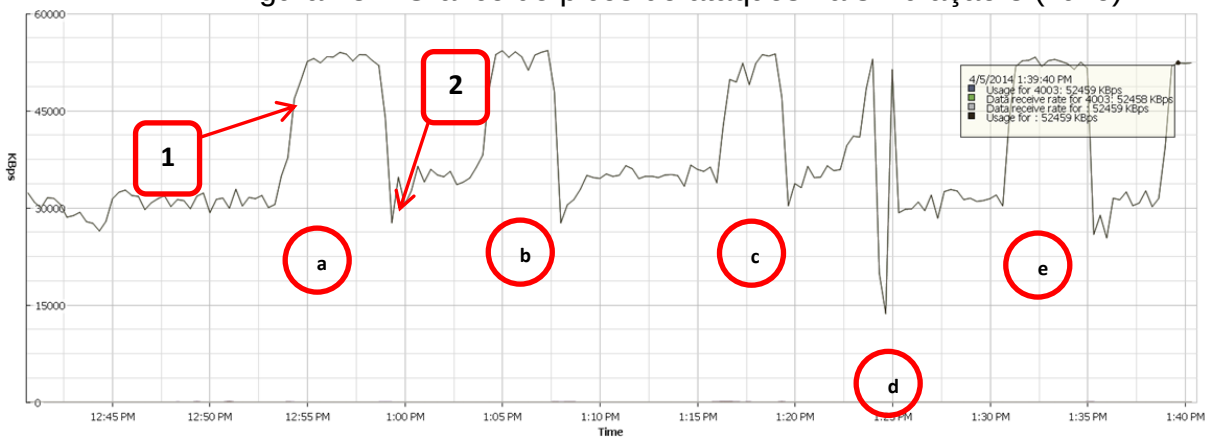


Tabela 13 – Resultados da simulação 8.

Métrica	Calculado	Medido (TCP/UDP)	Medido (TCP/UDP) com fluxo de ativação	Agente/Sistema
FPmed	250 mil	167 mil	167 mil	Adamantos5
PDET	xx	95%	95%	Arquitena
TDAD	xx	2 seg	2 seg	Arquitena
TABD	xx	a – 4 min b – 3 min c – 3 min d – 2 min e – 3.5 min	a – 4 min b – 3 min c – 3 min d – 2 min e – 3.5 min	Arquitena
PFPO	xx	4%	4%	Arquitena
PFNE	xx	19%	19%	Arquitena
Tráfego legítimo	xx	65%	65%	Adamantos3

No teste 8, os seguintes agentes sofreram as ações a seguir:

- **Adamantos1 - 6** – tráfego de ataque distribuído (bloqueios intermitentes)
- **Éacos1 e 2** – ficaram ativos por acionamento do Minos1
- **Éacos3 e 5** - ficaram ativos por acionamento do Adamantos2
- **Éacos4 e 6** - ficaram ativos por acionamento do Adamantos1
- **Minos1** – tráfego excedido (acionou vizinhos imediatos)

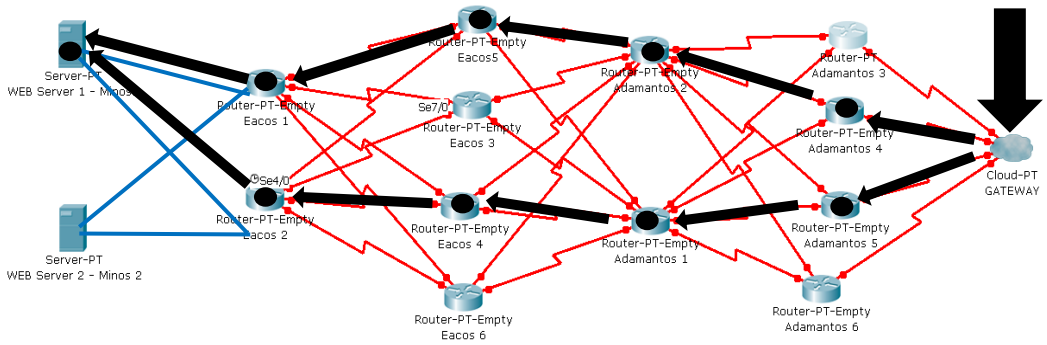
Nos testes 2, 4, 6 e 8 foi realizado medições da utilização de CPU depois do início do ataque. Antes da implantação do sistema Arquitena a porcentagem de CPU tinha a média de aproximadamente 10%, sem tráfego, e após a implantação do Sistema Arquitena a taxa de processamento subiu apenas 2% na fase de inicialização, quando o número de mensagens é maior. A tabela 14 mostra a taxa de utilização de CPU nas fases de ataque e bloqueio, no agente Adamantos5, de acordo com as figuras 17 e 18.

Tabela 14 – Utilização de CPU (testes pares).

Posição no tempo	**Comando no roteador: #*show processes***
1	CPU utilization for five seconds: 91%; one minute: 90%; five minutes: 90%
2	CPU utilization for five seconds: 35%; one minute: 34%; five minutes: 33%
a	CPU utilization for five seconds: 90%; one minute: 89%; five minutes: 89%
b	CPU utilization for five seconds: 90%; one minute: 89%; five minutes: 88%
c	CPU utilization for five seconds: 92%; one minute: 87%; five minutes: 85%
d	CPU utilization for five seconds: 91%; one minute: 82%; five minutes: 70%
e	CPU utilization for five seconds: 90%; one minute: 85%; five minutes: 84%

Após as 10 simulações do teste 8, as rotas de ataques mais utilizadas foram aquelas mostradas na Figura 19.

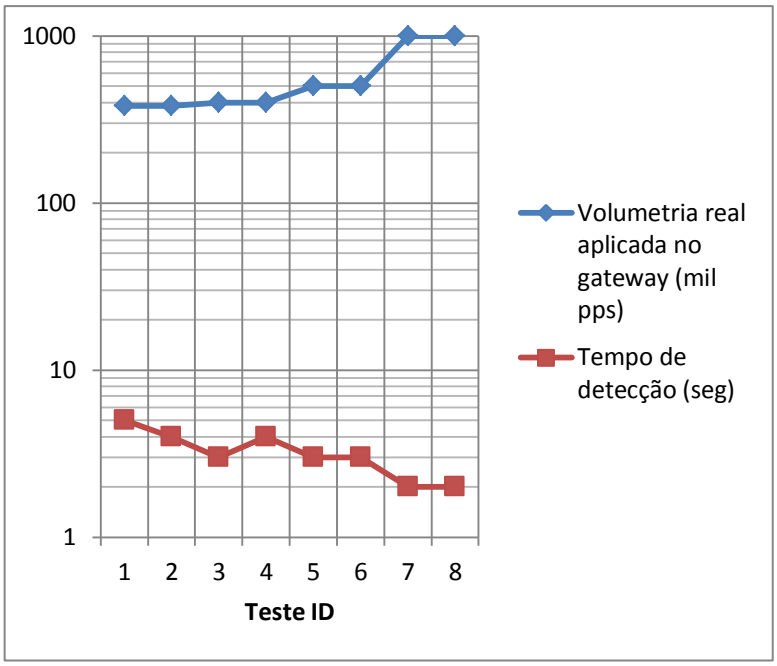
Figura 19 – Rotas de ataque mais frequentes (teste 8).



Antes mesmo do início das simulações e implantação do sistema Arquitena, foi medido em quanto tempo o servidor WEB parava de oferecer o serviço após o início do ataque DDoS. Em média, o serviço ficou indisponível após 2 segundos do início do ataque, quando realizado com a mesma volumetria dos testes 5 a 8. Quando se efetuou as mesmas verificações com a volumetria dos testes 1 a 4, o serviço ficou indisponível após 8 segundos. Após a implantação do sistema Arquitena, o servidor WEB (vítima) ficou indisponível apenas nas simulações dos testes 6 e 8, pois a volumetria alcançada foi muito alta em pouco tempo. O serviço ficou indisponível após 18 segundos do início do ataque, e em média o bloqueio do tráfego foi efetuado após 167 segundos, tornando o serviço disponível novamente.

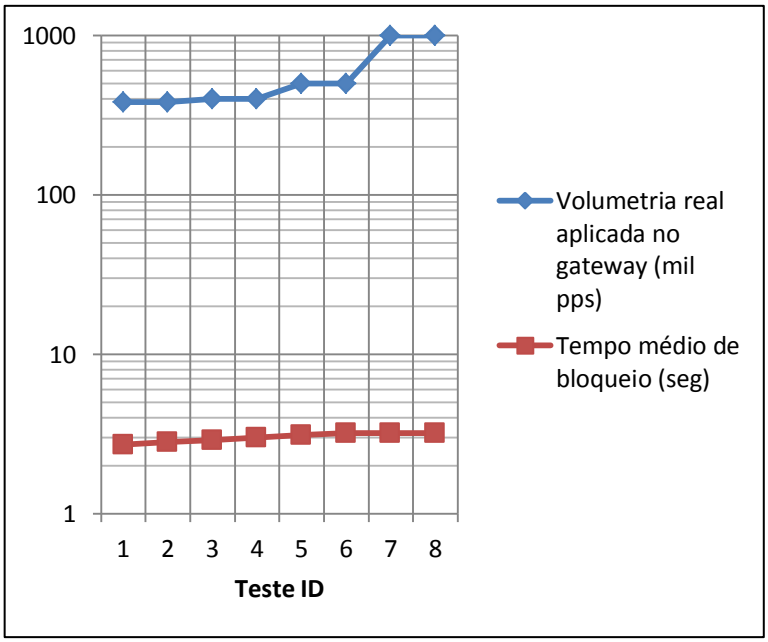
O Gráfico 1 mostra a tendência do tempo de detecção em relação à volumetria aplicada na rede, conforme os testes 1 a 8. O tempo de detecção foi diminuindo conforme a volumetria foi aumentada na entrada da rede, em cada um dos 8 testes realizados.

Gráfico 1 – Tempo de detecção x volumetria aplicada.



O Gráfico 2 mostra a tendência do tempo de bloqueio em relação à volumetria aplicada na rede, conforme os testes 1 a 8. O tempo médio de bloqueio foi aumentando conforme a volumetria foi aumentada na entrada da rede, em cada um dos oito testes realizados.

Gráfico 2 – Tempo médio de bloqueio x volumetria aplicada.



5.6.1. Comparações com trabalhos relacionados sem SMA

Em relação às técnicas sem SMA estudadas no Capítulo 3, a maioria dos trabalhos apresentavam os valores dos resultados obtidos nos testes. Com isso, foi possível comparar o desempenho baseado em algumas métricas definidas em ambas as técnicas. Neste ponto, é importante mencionar que os computadores utilizados nos trabalhos relacionados e os detalhes dos testes realizados são diferentes, o que pode afetar os valores exatos obtidos. Entretanto, dada à indisponibilidade de código fonte ou ambiente de testes de tais trabalhos, apenas o valor escalar apresentado foi utilizado para essa comparação. Cabe notar também que a abordagem *Path Attestation* não apresentou em (BHATTACHARJEE; RAGHAVAN; SANAND, 2011) resultados numéricos de desempenho que pudessem ser comparados com o sistema Arquitena. Por esse motivo, ela não foi incluída nesta comparação.

A Tabela 15 resume as semelhanças e vantagens em potencial do sistema Arquitena quando comparado às técnicas de defesa contra ataques DDoS sem utilização de agentes ou SMA mencionadas.

Tabela 15 – Comparação entre técnicas sem SMA e Arquitena.

MÉTRICA	Arquitena	Base estatística	HCI-MPR	*Path Attestation Scheme*	DefCOM
Classificação do tráfego	Não	Sim	Sim	Sim	Sim
Descarte de tráfego malicioso	Aleatório (baseado no tráfego real)	Discernimento estatístico	Discernimento estatístico	Verifica o caminho dos pacotes	Através de verificação no tráfego
Efetua marcação no pacote	Não: processo não intrusivo e rápido	Não	Não	Sim	Sim
Efetividade contra ataques DDoS (forjado ou não)	Contempla a maioria	Sim	Sim	Ataques com IPs de origem forjada	Sim

Tabela 15(cont.) – Comparação entre técnicas sem SMA e Arquitena.

MÉTRICA	Arquitena	Base estatística	HCI-MPR	*Path Attestation Scheme*	DefCOM
Efetua detecção, mitigação e bloqueio do tráfego malicioso	Sim	Não (Apenas detecção)	Não (Apenas detecção)	Não (Apenas detecção)	Não (Apenas detecção)
Tráfego legítimo tende a ser preservado	Sim	Sim	Sim	Sim	Sim
Tráfego legítimo pode ser priorizado	Não	Sim	Sim	Não	Sim
Depende da estabilidade do roteamento na rede	Não	Não	Sim	Sim	Não
Efetua renovação do valor do tráfego	Sim (constante)	Não (somente na fase inicial de testes)	Não (somente na fase inicial de testes)	Não (somente na fase inicial de testes)	Sim (constante)
Mitigação usa o próprio equipamento de rede	Não (usa agentes em rede virtual)	Sim	Sim	Sim	Sim
Tráfego legítimo compete com tráfego malicioso	Sim	Sim	Sim	Sim	Sim

Em resumo, a principal vantagem do Arquitena sobre as técnicas, sem utilização de agentes ou SMA mencionadas é de ter as funções de mitigação e bloqueio após a detecção. Outra vantagem sobre todas as outras técnicas de defesa é o fato da solução proposta não utilizar o equipamento de rede para efetuar o processo de mitigação: ao invés disso, ele utiliza uma rede virtual análoga a rede real, evitando onerar o processamento do equipamento de rede caso este esteja sobre um ataque de DDoS. Finalmente, uma vantagem do Arquitena em relação às técnicas apresentadas em (BHATTACHARJEE; RAGHAVAN; SANAND, 2011) e (MANN; KUMAR, 2011) é que ele não depende da estabilidade do processo de roteamento de rede para prosseguir com o processo de mitigação de ataques DDoS.

Outro ponto a se destacar é que nem todas as técnicas estudadas apresentam as mesmas métricas utilizadas pelo Arquitena. Por exemplo, o DefCOM não apresentava o valor da taxa de detecção e nem os valores de falso-positivo e falso-negativo. A tabela 16 mostra a comparação entre essas técnicas sem utilização de agentes ou SMA.

Tabela 16 – Comparação de métricas entre técnicas sem SMA (numérica).

DefCOM	**Arquitena (média testes 1 - 8)**
Tempo médio para bloqueio: **450 segundos**	Tempo médio para bloqueio: **240 segundos**
Tempo de detecção em um nó: **1.58 – 2.45 segundos**	Tempo de detecção médio em um agente: **3 segundos**
Base estatística	**Arquitena (média testes 1 - 8)**
Taxa de detecção: **50% (10 atacantes)** **98,4% (100 atacantes)**	Taxa de detecção: **70% (10 atacantes)** **95% (90 atacantes)**
Falso-positivo: **1,8% (volumetria baixa de ataque)** **2,9% (volumetria alta de ataque)**	Falso-positivo: **4% (volumetria baixa de ataque)** **1,5% (volumetria alta de ataque)**
HCI-MPR	**Arquitena (média testes 1 - 8)**
Taxa de detecção: **95,2% (10 mil pps)** **99,8% (1 mil pps)**	Taxa de detecção: **70% (312 mil pps)** **95% (167 mil pps)**

Quando comparado com a técnica DefCOM (MIRKOVIC; ROBINSON; REIHER, 2003), o sistema Arquitena teve tempo médio de detecção maior, mas o tempo médio para bloqueio foi menor. Essa diferença é devido ao esforço em se

comunicar com o equipamento de rede e coletar a estatística, ao contrário do DefCOM em que cada nó da rede executa as tarefas pertinentes de alerta e detecção. Através desta estratégia de utilização dos próprios equipamentos para classificar o tráfego, o processamento do equipamento pode ser onerado. Já a técnica baseada em estatística (GUPTA; MISRA; JOSHI, 2008) teve valores piores de (1) taxa de detecção quando a quantidade de atacantes era menor, e (2) taxa de falso-positivo quando a volumetria era maior. Contudo, o sistema Arquitena teve uma pequena desvantagem quando a quantidade de atacantes era maior e, a volumetria de ataque era menor. O sistema Arquitena independe da variação da quantidade de agentes, porém, mantém uma taxa alta de detecção. Assim, a taxa de detecção é diretamente proporcional à volumetria do ataque e não à quantidade de agentes.

Em relação à taxa de detecção, o sistema HCI-MPR (MANN; KUMAR, 2011) teve vantagem quando o tráfego era menor, bem como uma taxa de detecção melhor quando a volumetria de ataque era maior. Esses valores são discutíveis, entretanto, a quantidade de pacotes aplicada na rede de simulações do sistema Arquitena é muito superior ao sistema HCI-MPR. A abordagem de defesa adaptativa (KOTENKO; ULANOV, 2007) não apresentou resultados numéricos de desempenho que pudessem ser comparados com o sistema Arquitena e, por esse motivo, não foi incluída nesta comparação.

5.6.2. Comparações com trabalhos relacionados com SMA

Em relação às técnicas com SMA estudadas no Capítulo 3, a maioria dos trabalhos apresentavam os valores dos resultados obtidos nos testes. Com isso, foi possível comparar o desempenho baseado em algumas métricas definidas em ambas as técnicas. Neste ponto, é importante mencionar que os computadores

utilizados nos trabalhos relacionados e os detalhes dos testes realizados são diferentes, o que pode afetar os valores exatos obtidos. Entretanto, dada à indisponibilidade de código fonte ou ambiente de testes de tais trabalhos, apenas o valor escalar apresentado foi utilizado para essa comparação. Cabe notar que a abordagem de Arcabouço de defesa Adaptativa não apresentou em (KOTENKO; ULANOV, 2007) resultados numéricos de desempenho que pudessem ser comparados com o sistema Arquitena. Por esse motivo, ela não foi incluída nesta comparação.

A Tabela 17 resume as semelhanças e vantagens em potencial do sistema Arquitena quando comparado às técnicas de defesa contra ataques DDoS baseadas em agentes ou SMA estudadas.

Tabela 17 – Comparação entre técnicas com SMA e Arquitena.

MÉTRICA	Arquitena	Aprendizado por Reforço	Defesa Adaptativa	Reconhecimento de padrões	Agentes móveis
Descarte de tráfego malicioso	Aleatoriamente, mas monitora tráfego	Discernimento estatístico	Baseia-se em base de endereços reais	Discernimento estatístico e observação do tráfego	Verifica origem do pacote em cada servidor
Facilidade para diferenciar tráfego	Sim (renovação do valor do tráfego constante, assim como a definição de limites)	Sim (caso tráfego não seja muito próximo do limiar do tráfego normal)	Sim (renovação do valor do tráfego constante, assim como a definição de limites)	Sim (devido ao monitoramento)	Agentes móveis verificam em cada servidor
Efetua detecção, mitigação e bloqueio do tráfego malicioso	Sim	Somente detecção	Sim	Somente detecção	Somente detecção

Tabela 17(cont.) – Comparação entre técnicas com SMA e Arquitena.

MÉTRICA	Arquitena	Aprendizado por Reforço	Defesa Adaptativa	Reconhecimento de padrões	Agentes móveis
Abordagem é adaptável às variações de tráfego	Sim	Não	Sim	Não	Não
Custo de defesa	Utiliza apenas a coleta de dados no equipamento de rede, não onera processamento local	Utiliza o próprio equipamento de rede para classificar o tráfego (onera processamento local)	Usa adaptação para diminuir o custo de defesa, não onera processamento local	Utiliza o próprio equipamento de rede para classificar o tráfego (onera processamento local)	Utiliza o próprio equipamento de rede para classificar o tráfego (onera processa-mento local)
Arquitetura descentrali-zada na comunica-ção dos agentes	Sim (elimina ponto único de falha)	Sim (elimina ponto único de falha)	Sim (elimina ponto único de falha)	Sim (elimina ponto único de falha)	Sim (elimina ponto único de falha)
Agentes específicos para detecção	Não (todos os agentes atuam como detectores e mitigadores)	Sim	Sim	Sim	Sim
Comunica-ção entre os agentes é constante	Não	Sim (onera o sistema de análise)	Sim (onera o sistema de análise)	Sim (onera o sistema de análise)	Sim (onera o sistema de análise)
Nível de confiança para diferenciar o tráfego malicioso	Aumenta (tempo de detecção tende a diminuir devido à renovação do valor do tráfego)	Aumenta (tempo de detecção tende a diminuir devido à renovação do valor do tráfego)	Aumenta (tempo de detecção tende a diminuir devido à renovação do valor do tráfego)	Aumenta com o acréscimo de agentes na rede	Aumenta, mas o tempo de detecção também

Em resumo, as principais vantagens do Arquitena em relação às técnicas apresentadas baseadas em agentes ou SMA é o fato do sistema proposto ter as

funções de mitigação e bloqueio após a detecção. As técnicas apresentadas em (AKYAZI; UYAR, 2008) e (BAIG; SALAH, 2009), por exemplo, apresentam somente a fase de detecção de ataques DDoS. Além disso, o sistema Arquitena automaticamente pode desativar alguns agentes caso eles não sejam necessários para o processo de mitigação. Essa característica diferencia-o das demais técnicas, incluindo a técnica de Defesa Adaptativa (KOTENKO; ULANOV, 2007) e Aprendizado por Reforço Cooperativo (MALIALIS; KUDENKO, 2013).

A tabela 18 mostra a comparação entre essas técnicas com utilização de agentes ou SMA e o sistema Arquitena.

Tabela 18 – Comparação de métricas entre técnicas com SMA (numérica).

Aprendizado por Reforço	**Arquitena**
Porcentagem de preservação de tráfego legítimo: **52,86% (tráfego de ataque constante) - Baseline** **57,08% (tráfego de ataque constante) – MARL** **85,19% (tráfego de ataque constante) – CTL**	Porcentagem de preservação de tráfego legítimo: **65% (tráfego de ataque constante)**
Reconhecimento de padrões	**Arquitena**
Taxa de detecção: **88% (tráfego intenso)** **95% (tráfego não muito intenso)**	Taxa de detecção: **70% (tráfego intenso)** **95% (tráfego não muito intenso)**
Falso-positivo: **5,28% (128 agentes)** **1,55% (1024 agentes)**	Falso-positivo (12 agentes): **1,5% (volumetria alta de ataque)** **4% (volumetria baixa de ataque)**

Tabela 18(cont.) – Comparação de métricas entre técnicas com SMA.

Reconhecimento de padrões	**Arquitena**
Falso-negativo: **38,72% (128 agentes)** **11,37% (1024 agentes)**	Falso-negativo (12 agentes): **11% (volumetria alta de ataque)** **19% (volumetria baixa de ataque)**
SMA com agentes móveis	**Arquitena**
Tempo médio de detecção: **3,02 seg (2 verificações)** **13,7 seg (todas as verificações)**	Tempo médio de detecção: **5 seg (volumetria alta de ataque)** **2 seg (volumetria baixa de ataque)**

O sistema Arquitena apresentou algumas vantagens em relação a certas técnicas com utilização de sistemas multiagentes. Através do sistema Arquitena, o tráfego legítimo foi preservado em 65% mesmo sem qualquer classificação ou priorização do tráfego legítimo, enquanto que técnicas com Aprendizado por Reforço (MALIALIS; KUDENKO, 2013) apresentaram taxas de preservação de apenas 52,08%. Na técnica com utilização de agentes móveis (AKYAZI; UYAR, 2008), o tempo de detecção varia conforme as verificações que os agentes realizam nos servidores, a fim de saber a origem do tráfego. Quanto mais verificações, maior é o tempo de detecção. O sistema Arquitena apresentou variação no tempo de detecção somente em relação à volumetria, independendo da quantidade de verificações nos equipamentos da rede. Quando comparado com o SMA que utiliza reconhecimento de padrões (BAIG; SALAH, 2009), o Arquitena tem ganho expressivo no percentual de falso-negativos, mesmo com a quantidade de agentes mantida em um valor fixo.

6 CONCLUSÃO

Os ataques de DDoS contra empresas de infraestrutura, inclusive ISP e HSP (*Hosting Service Provider* - Provedor de Serviço de Hospedagem), continuam a ser uma ameaça operacional, sendo que no ano de 2014 esta ainda é a principal preocupação de empresas desse tipo (ARBOR, 2014). Espera-se, portanto, um aumento na demanda por serviços de detecção e mitigação de ataques DDoS nos próximos anos, em resposta ao aumento de ataques em termos de frequência e volumetria.

Prover tais serviços de forma eficiente e eficaz é, entretanto, algo desafiador. Devido à natureza distribuída de tais ataques, ações isoladas de mitigação na rede dificilmente são capazes de reduzir o tráfego pertinente a ataque DDoS a níveis pouco perceptíveis. Outro motivo que torna a mitigação dos ataques DDoS complicada é o correto discernimento do tráfego legítimo durante um ataque. O comportamento da rede sob um ataque DDoS pode exaurir os recursos dos equipamentos de rede (e.g., roteadores).

Neste contexto, o presente trabalho apresenta uma solução baseada em abordagem com Sistema Multiagentes para monitorar, detectar e bloquear ataques DDoS baseados em inundação. O sistema é composto de múltiplos agentes que tendem a conservar parte do tráfego legítimo automaticamente, sem a necessidade de intervenção manual. Todos os agentes, de forma deliberativa, tornam mais preciso o padrão limite de pacotes em cada equipamento, nos intervalos de tempo configurados. Portanto, o sistema tenta ser suficientemente inteligente para distinguir horários de pico no tráfego da rede, podendo tratar várias origens de ataques simultaneamente, sejam elas internas ou externas ao ISP. Além disso, a solução

proposta utiliza uma rede virtual análoga à real, de forma que cada agente monitora um equipamento da rede. Dessa forma, o agente não fica instalado no equipamento de rede, evitando onerar o processamento do equipamento e deixando-o livre para executar suas funções originais.

O sistema Arquitena tem a vantagem de estar presente em toda a rede do ISP, sendo que alguns agentes não ficam ativos todo o tempo, evitando coleta de dados desnecessária pelo equipamento de rede enquanto não são detectados ataques. Essa característica de cooperação entre os agentes tende a melhorar a eficiência do sistema, sem afetar criticamente sua robustez. Outro ponto a ser destacado é que os equipamentos tradicionais de segurança de perímetro não são desenhados para resolverem problemas recorrentes aos ataques DDoS. Pelo contrário, esses equipamentos inúmeras vezes são alvos dos ataques DDoS. Assim, o Arquitena pode se tornar uma ferramenta complementar a esses mecanismos de defesa mais comuns, implantados na rede do ISP. Além disso, este sistema não está atrelado a qualquer protocolo de transporte ou de rede para funcionar corretamente. Este sistema não foi projetado para mitigar diretamente ataques de vulnerabilidade. Assim, a atuação do Arquitena é complementar à soluções que visam prevenir a exploração de tais vulnerabilidades, como *firewalls*.

Em conclusão, a implantação do sistema Arquitena torna a rede do ISP ou HSP mais resiliente, e essa vantagem pode ser estendida para os clientes que utilizam os recursos do ISP ou do HSP. Dessa forma, os ataques de longa duração e volumosos, ataques no estilo rajadas, mas também com grande volume de tráfego, podem ser mitigados. Sendo assim, a capacidade do enlace, capacidade de sessão,

capacidade de serviço da aplicação e acesso contínuo a base de dados podem ser preservados.

Através de simulações propostas, pode-se observar que o sistema Arquitena consegue obter baixas taxas de falso-positivos e falso-negativos, além de ter alta taxa de detecção. Essas métricas são primárias para qualificar um sistema de defesa contra ataques DDoS.

6.1. Trabalhos futuros

O sistema de defesa proposto se baseia somente na volumetria do ataque para discernir no processo de mitigação, que pode incluir o bloqueio total do tráfego. Assim, uma análise mais profunda poderia considerar a utilização de um sistema híbrido de defesa, isto é, considerando o tráfego anômalo de divergências na assinatura do tráfego para classificar o tráfego malicioso.

Outro assunto interessante de pesquisa envolve a utilização de um SMA baseado em arquitetura distribuída (Jade). Tal análise permitiria verificar se essa arquitetura pode contribuir para melhorar os resultados alcançados nesta pesquisa em termos de taxa de falso-positivos, falso-negativos e tempo de detecção.

Algumas pequenas adições ou modificações na arquitetura do Arquitena também podem ser de interesse. Por exemplo, a fim de não necessitar de um agente externo, o sistema Arquitena poderia ser desenvolvido para assimilar o tráfego dos equipamentos, através de requisição direta, e assim tomar a decisão de mitigação e bloqueio mais rapidamente.

Adicionalmente, o sistema Arquitena poderia ter um processo de descobrimento da rede através de probes. Dessa forma, o sistema de forma

autônoma poderia desenhar toda a topologia da rede sem qualquer intervenção manual.

6.2. Publicações

A seguinte publicação foi resultado da pesquisa realizada durante esta dissertação:

- Conferência: em (PEREIRA; SIMPLICIO; BRANDAO, 2013) foi apresentado em uma sessão técnica os principais objetivos do sistema Arquitena, utilizando um Sistema Multiagentes para mitigar ataques DDoS em redes de ISPs.

REFERÊNCIAS

AKYAZI, U.; UYAR, A. **Distributed Intrusion Detection using Mobile Agents against DDoS attacks**. Computer and Information Sciences – ISCIS 23rd International Symposium.2008. p. 1-6.

ARBOR. **Worldwide Infrastructure Security Report**. 2014. Disponível em: <http://pages.arbornetworks.com/rs/arbor/images/WISR2014.pdf>. Acesso em: 10 mar. 2014.

AUSTIN J. **How to Do Things With Words**. Oxford University Press, Oxford. 1962.

BHATTACHARJEE, R.; SANAND, S.; RAGHAVAN, S. **Path Attestation Scheme to Avert DDoS Flood Attacks**. International Federation for Information Processing. 2010.

BAIG, Z.; SALAH, K. **Multi-agent patter recognition mechanism for detecting distributed denial of service attacks**. IET Information Security. 2009.

BAQER, M.; KHAN, A.; BAIG, Z. **Implementing a graph neuron array for pattern recognition within unstructured wireless sensor networks**. *Proc. Workshops Embedded Ubiquitous Comput*. 2005. p. 208.

BORDINI, R.; HÜBNER, J.; WOOLDRIDGE, M. **Programming Multi-Agent Systems in AgentSpeak using Jason**. Wiley. 2007.

CERT. **Cartilha sobre ataques**. Disponível em: <http://cartilha.cert.br/ataques/>. Acesso em: 09 jan. 2014.

DOYLE, J.; CARROLL, J. **CCCIE Professional Development - Routing TCP/IP** – Volume I**,** CiscoPress.com, 2005. Disponível em: <http://www.net130.com/tutorial/cisco-pdf/routingtcpipv1.pdf>. Acesso em: 18 ago. 2013.

DYNAGEN. **Software Dynagen**. Disponível em: <http://dynagen.org/>. Acesso em: 06 fev. 2014.

DYNAMIPS. **Software Dynamips**. Disponível em: <http://www.gns3.net/dynamips/>. Acesso em: 06 fev. 2014.

FERNANDÉZ, V. et al. **Evaluating Jason for Distributed Crowd Simulations**. Disponível em: <http://www.uv.es/grimo/publications/icaart2010.pdf>. Acesso em: 30 set. 2012.

FININ, T.; LABROU, Y.; MAYFIELD, J. **KQML as agent communication language**. University of Maryland Baltimore County. 1995.

FRANKLIN, S.; GRAESSER, A. **Is it an agent or just a program? A taxonomy for autonomous agents**. Proceedings of the Agent Theories, Architectures, and Languages Workshop. Berlin: Springer-Verlag. 1996.

GNS3. **Software GNS3**. Disponível em: <http://www.gns3.net/**>**. Acesso em: 19 mar. 2014.

GUPTA, B.; MISRA, M.; JOSHI, R. **An ISP Level Solution to Combat DDoS Attacks using Combined Statistical Based Approach**. Journal of Information Assurance and Security 2. 2008. p.102-110.

HAMMOUD, D.; MAAMRI, R.; SAHNOUN, Z. **Machine Learning in an Agent**: A Generic Model and an Intelligent Agent based on Inductive Decision Learning. Journal of Artificial Intelligence, 4. 2011. p. 29-44.

IPFIX. **Protocolo IPFIX**. Disponível em: <http://datatracker.ietf.org/wg/ipfix/charter/>. Acesso em: 15 jul. 2013.

JADE. **Software Jade**. Disponível em: <http://jade.tilab.com/>. Acesso em: 16 mar. 2014.

JFLOW. **Software Jflow**. Disponível em: <http://www.juniper.net/techpubs/software/erx/junose82/swconfig-ip-services/html/ip-jflow-stats-config2.html>. Acesso em: 01 nov. 2013.

JUNEJA, D.; CHAWLA, R.; SINGH, **A. An Agent-Based Framework to CounterAttack DDoS Attacks.** International Journal of Wireless Networks and Communications. Volume 1, Number 2. 2009. p. 193-200.

KOTENKO, I.; ULANOV, A. **Multi-agent Framework for Simulation of Adaptive Cooperative Defense Against Internet Attacks**. Autonomous Intelligent Systems: Multi-Agents and Data Mining Lecture Notes in Computer Science. Volume 4476. 2007. p. 212-228.

KWAK, Y.; ANBARI, F**. Benefits, obstacles, and future of six sigmas approach**. Elsevier Direct (Technovation 26). 2006. p. 708–715.

LABOVITZ, C.; MALAN, R.; JAHANIAN, F. **Origins of Internet Routing Instability**. IEEE Infocom. 1999

LITHGOW-SMITH, B.; TAMMA, V.; WOOLDRIDGE, M. **An Ontology for Coordination**. In Applied Artificial Intelligence, 25. 2011. p. 235—265.

LUCK, M.; MCBURNEY, P.; GONZALES-PALACIOS, J. **Agent-based computing and programming of agent systems**. 2006. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.117.3420<. Acesso em: 08 jun. 2012.

MALIALIS, K.; KUDENKO, D. **Large-Scale DDoS Response Using Cooperative Reinforcement Learning**. EUMAS-2013 11th European Workshop on Multi-Agent Systems. Toulouse, France. 2013

MANN, P.; KUMAR, D. **A Reactive Defense Mechanism based on an Analytical Approach to Mitigate DDoS Attacks and Improve Network Performance**. International Journal of Computer Applications. 2011.

MIRKOVIC, J.; ROBINSON, M.; REIHER, P. **Alliance Formation for DDoS Defense**. Proceedings of the 2003 workshop on New security paradigms. Ascona, Switzerland. 2003. p. 18-21.

MIRKOVIC, J. et al. **Internet Denial of Service**: Attack and Defense Mechanisms. Prentice Hall. 2004.

MIRKOVIC, J. et al. **Benchmarks for DDoS Defense Evaluation**. in MILCOM. 2006. Disponível em: <http://dl.acm.org/citation.cfm?id=1897120>. Acesso em: 07 jun. 2012.

NETFLOW. **Protocolo Netflow**. Disponível em: <http://www.cisco.com/en/US/products/ps6601/products_ios_protocol_group_home.html>. Acesso em: 20 fev. 2014.

MCKEOWN, N. et al. **OpenFlow: Enabling Innovation in Campus Networks**. Seattle, USA: SIGCOMM, 2008. Disponível em: < http://archive.openflow.org/documents/openflow-wp-latest.pdf>. Acesso em: 25 mar. 2014.

PADGHAM, L.; WINIKOFF, M. **Developing Intelligent Agent Systems:** A practical guide. Wiley. 2006.

PEAKFLOW. **Equipamento PeakFlow**. Disponível em : <http://www.arbornetworks.com/products/peakflow>. Acesso em: 05 jan. 2014.

PEREIRA, J.; SIMPLICIO, M.; BRANDAO, A. **A multiagent approach for detecting and mitigating DDoS attacks**. São Paulo: WESAAC, 2013. p. 61-66.

RAO, A. **AgentSpeak(L):** BDI Agents speak out in a logical computable language**.** In Van de Velde, W. and Perram, J. W., editors, Agents Breaking Away: Proceedings of the Seventh European Workshop on Modelling Autonomous Agents in a Multi-Agent World, (LNAI Volume 1038). Springer: Verlag. Berlim, Alemanha. 1996. p. 42-55.

RAO, A.; GEORGEFF, M. **An Abstract Architecture for Rational Agents**. In Proceedings of Principles of Knowledge Representation and Reasoning (KR). 1992 p. 438-449.

RFC 3013. **Documento RFC 3013 – Recommended Internet Service Provider Security Services and Procedures**. Disponível em: <http://tools.ietf.org/html/rfc3013>. Acesso em: 10 jan. 2014.

RIOREY. **Taxonomy of DDoS Attacks**. 2014. Disponível em: <http://static.squarespace.com/static/53319b01e4b0ec02b601ca49/t/537ab649e4b02004337c19aa/1400550985376/RioRey_Taxonomy_DDoS_Attacks_2.6_2014.pdf>. Acesso em: 11 fev. 2014.

T50. **Software T50.** Disponível em: <http://t50.sourceforge.net/>. Disponível em: 03 out. 2013.

WEERDT, M.; MORS, A.; WITTEVEEN, C. **Multi-agent Planning An Introduction to planning and coordination**. 2005. Disponível em: <http://www.st.ewi.tudelft.nl/~mathijs/publications/easss05.pdf>. Acesso em: 28 set. 2011.

WETHERALL, D.; TANEMBAUM, A. **Redes de Computadores**. Pearson, 5º ed. 2011.

WHITE, D.; BORGATTI, S. **Betweenness centrality measures for directed graphs.** Social Network 16 335-346. North Holland. 1994

WILLINGER, W.; ALDERSON, D.; DOYLE, J. **Mathematics and the Internet**: A source of Enormous Confusion and Great Potencial. AMS. 2009.

WOOLDRIDGE, M. **An Introduction Multi Agent Systems**. Wiley, 2º ed. 2009.