Detecção de Intrusão em Redes de Alta Velocidade

Benicio P. de Carvalho Filho
Inpe - LAC
benicio@lac.inpe.br

Antonio Montes Filho
Cenpra
antonio.montes@cenpra.gov.br

Resumo

Uso do "netflow" como alternativa para a detecção de intrusão em redes de alta velocidade, com o propósito de melhoria da eficiência e eliminação do problema de perda de pacotes.
Palavras-chave: *detecção de intrusão, netflow*

1. Introdução

Com a contínua expansão da Internet e a evolução tecnológica que tem possibilitado o aumento da velocidade das redes, a detecção de intrusão, apesar de relativamente recente, vem sendo forçada a buscar novos paradigmas. A análise do tráfego com o exame de cada pacote trafegado torna-se tarefa muitas vezes inviável, em função da grande capacidade de processamento necessária para que não ocorra perda de informação. A partir do uso do Netflow, investigamos a detecção de intrusão a partir de padrões normais de tráfego para a rede tratada e a busca de anomalias a partir desse padrão.

Ao longo dos últimos anos, o crescimento do número de incidentes tem refletido o próprio crescimento da Internet. A figura a seguir [Allen2000] ilustra essa situação, apresentando dados do CERT/CC para o número de incidentes registrados por ano:

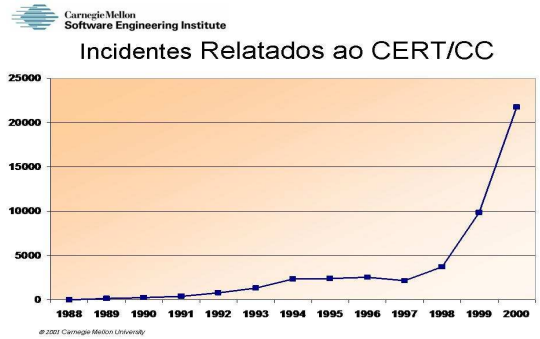

Figura 1. Incidentes reportados ao CERT

O crescimento do comércio eletrônico tende a exacerbar essa tendência de crescimento do número de incidentes. Enquanto, nos primórdios, os ataques externos eram lançados por aqueles interessados em explorar a Internet para benefício pessoal e como teste para suas habilidades, existe uma tendência crescente, hoje em dia, para intrusões motivadas por interesses financeiros, políticos e militares.

Outra tendência que se observa é a disseminação de técnicas e ferramentas de intrusão, facilitando a ação de atacantes com pouco conhecimento técnico a ampliando o universo de agentes hostis. Nos anos 80, os atacantes eram os especialistas em sistemas; eles possuíam um alto nível de conhecimento, além de métodos pessoais para se infiltrar nos sistemas. O uso de ferramentas automáticas e scripts era exceção e não regra. Hoje, qualquer um pode atacar uma rede, em razão do grande número de ferramentas de intrusão e scripts automáticos disponíveis através da própria Internet.

A figura a seguir [Allen2000] apresenta uma interessante relação observada entre a evolução dos tipos de ataques e o nível de conhecimento dos atacantes.

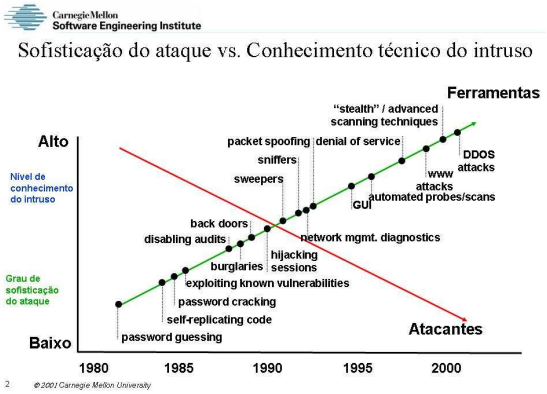

Figura 2. Ataques x atacantes

2. Detecção de intrusão

Detecção de intrusão é o processo de monitoramento e análise de eventos ocorridos em sistemas computacionais ou redes com o propósito de identificar sinais de problemas de segurança.

Os sistemas de detecção de intrusão são, para as redes e sistemas computacionais, como os sistemas de vigilância do mundo físico e, assim como estes, variam quanto às suas características e seu custo. Eles monitoram a rede em busca de comportamentos suspeitos, desempenhando um importante papel nas arquiteturas de segurança, sem eliminar, contudo, a necessidade de outras medidas.

A detecção de intrusão é uma tecnologia relativamente recente, tendo a pesquisa na área adquirido maior relevância a partir de 1980. Essa pesquisa, entretanto, tem produzido um amplo espectro de estratégias na busca de soluções para consecução de seus objetivos.

Em razão do pouco tempo de existência e, consequentemente, de sua imaturidade, têm surgido diversas novas dificuldades a superar. Existe ainda uma distância muito grande entre os aspectos teórico e prático na detecção de intrusão, o que, muitas vezes, pode levar a pseudo-soluções inconsistentes ou conflitantes com outros elementos do sistema de segurança, ou ainda ao emprego equivocado de ferramentas e técnicas, com um inerente aumento do potencial de riscos.

A detecção de intrusão, dependendo da oportunidade em que se dá, objetiva uma possível medida emergencial que interrompa a atividade hostil, além de informações que possibilitem:

- Determinar a real ocorrência do ataque: evidências que comprovem, sem sombra de dúvida, a ocorrência do incidente.
- Localizar o ataque: o que ocorreu e onde ocorreu.
- Identificar o atacante: para as medidas cabíveis.
- Avaliação da estratégia do ataque, seu grau de periculosidade e até suas vulnerabilidades. Para determinar a resposta adequada.

O processo de detecção pode ser visto de diferentes formas. Pode resultar da observação de um ataque em progresso ou do reconhecimento dos resultados de um ataque após a sua ocorrência.

Sistemas de detecção de intrusão atuam através da coleta e análise do tráfego da rede. Nessa análise procura-se detectar algum padrão ou alguma anomalia que caracterizem atividades entendidas como ilícitas. Podem ser classificados:

Quanto à origem

- Estação ("Host Based"): Dados provenientes de log, accounting e aplicações de detecção de intrusão executadas numa estação são usados como entrada para o sistema.
- Rede(Network Based): Dados sobre tráfego na rede, bem como dados provenientes de múltiplas estações são usados para detectar intrusões.
- Múltiplo (Multi-Network/Infrastructure Based): Dados provenientes de múltiplas fontes são usados para detectar intrusões. Essas fontes são entidades constituintes de um domínio administrativo e os dados podem ser oriundos de de aplicações, estações, redes ou outros sistemas de detecção "multi-network".

Quanto ao tipo de análise

- Detecção por anomalia: O sistema de deteção de intrusão procura identificar desvios em relação a características de tráfego consideradas normais.
- Detecção por assinatura: O sistema de deteção de intrusão procura identificar atividades que correspondam a padrões conhecidos (assinaturas) ou à exploração de vulnerabilidades conhecidas dos sistemas. A maior parte dos sistemas de detecção de intrusão usados atualmente situa-se nesta categoria.

3. Sistemas de detecção de intrusão

Apresentamos a seguir alguns exemplos de sistemas de detecção de intrusão usuais, discutindo suas características e principais limitações:

Shadow [Shadow]

O Shadow, "Secondary Heuristic Analysis for Defensive Online Warfare", software de domínio público, usa estações como sensores e analisadores. Os sensores são usualmente posicionados em pontos importantes da rede, como a porta externa de um firewall, onde podem ter acesso a todo o tráfego, enquanto a estação de análise é localizada internamente ao firewall. Os sensores extraem os cabeçalhos dos pacotes e os salvam em um arquivo. Este é lido, numa periodicidade pré-determinada, pela estação de análise que, por sua vez, efetua uma operação de filtragem e gera um segundo arquivo. O Shadow não não provê alertas ao identificar eventos, pela possibilidade de ocorrência de grande número de falsos positivos, observada em outros sistemas.

O sensor usa o libpcap, desenvolvido pelo Network Research Group, do Lawrence Berkeley Lab., para prover a capacidade de captura dos pacotes, enquanto a capacidade de filtragem de pacotes da estação de análise é provida pelo tcpdump e por um script Perl, fornecido como parte do kit.

Este tipo de IDS necessita alguma forma de interceptação ou espelhamento do tráfego, além de exigir um tratamento dos pacotes que, dependendo da largura da banda, pode levar à perda substancial de informação pela incapacidade de tratar todos os pacotes à medida que chegam.

Snort [Snort]

O Snort é um sistema de detecção de intrusão para pequenas redes baseado na identificação de padrões. Suas principais características são a facilidade para criação de regras, a análise do tráfego em tempo-real, análise de protocolo, e capacidade de inspeção e busca de padrões no conteúdo dos pacotes. Pode ainda emitir alertas de diversos tipos para diferentes interfaces.

Bro [Paxson1999]

O Bro é uma ferramenta de pesquisa em desenvolvimento pelo Lawrence Livermore National Laboratory. Ele vem sendo construído em partes, para explorar assuntos relacionados à robustez de sistems de detecção de intrusão, isto é, levantando quais características fazem esses sistemas capazes de resistir a ataques a eles dirigidos. Os objetivos de projeto incluem:

- *alta capacidade de monitoramento sem perda de pacotes.*
- *notificação em tempo-real, para assegurar resposta rápida a ameaças.*
- *mecanismo independente da política, para facilitar desenvolvimento, implementação e manutenção.*

- *extensibilidade, para mais facilmente se ajustar ao surgimento de novos ataques.*
- *capacidade de repelir ataques. Atacantes sofisticados irão muitas vezes tentar encontrar vulnerabilidades nos próprios sistemas de detecção de intrusão.*

O Bro tem uma hierarquia de funções de tres níveis: no nível mais baixo, usa a libpcap para capturar pacotes da rede. Isto desacopla a funcionalidade principal de detecção dos detalhes da rede em si. Também permite a rejeição em baixo nível de uma fração significativa dos pacotes entrando na rede. Assim, a libpcap irá capturar todos os pacotes associados com os protocolos de aplicação (finger, ftp, telnet, etc.) de que o Bro está tratando.

A próxima camada, de eventos, efetua verificações de integridade nos cabeçalhos dos pacotes. Se o pacote é mal-formado, será gerado um evento identificando o problema e o cabeçalho será descartado. É feita então uma verificação para decidir pela gravação do conteúdo completo do pacote, ou apenas do cabeçalho, ou nenhum tipo de registro.

Eventos gerados nesse processo são enfileirados para investigação por um "*script*" de interpretação de políticas, que reside no terceiro nível. O interpretador associa valores de eventos aos códigos do tratador de eventos que, por sua vez, poderá gerar novos eventos, notificações em tempo-real, ou gravação de dados.

Atualmente o Bro monitora quatro aplicações: finger, ftp, portmapper e telnet. É suportado em diversas variantes de UNIX e usado como parte do sistema de segurança do laboratório. Não tem experimentado perda de pacotes para um tráfego de 25Mbps com carga de análise de aproximadamente 200 pacotes/segundo.

Observemos que, em vista das velocidades comuns hoje em dia, este volume de tráfego está longe de ser alto.

4. Pontos fortes e fracos dos atuais Sistemas de Detecção de Intrusão:

Pontos Fortes

- Monitoramento, classificação e análise de eventos e de utilização.
- Identificação de anomalias em atividades dos usuários.
- Definição de limiares de segurança e acompanhamento de desvios.
- Políticas de segurança da informação "default".
- Reconhecimento de padrões de ataques conhecidos.
- Mecanismos de registro para auditoria.
- Possibilidade de uso por profissionais pouco experientes em segurança no acompanhamento de importantes atividades relativas à segurança.

Pontos Fracos ou Limitações

- Baixo nível de escalabilidade.
- Limitações relativas a recursos: Para detectar assinaturas de ataques os sistemas devem capturar, armazenar e analisar grandes volumes de dados, praticamente em tempo-real. O volume de dados pode ser imenso, devendo potencialmente os IDS's manter informações sobre conexões de milhares de máquinas.
- Técnicas de evasão como fragmentação, em que o atacante envia pacotes fragmentados e consegue enganar os IDS's quebrando a carga útil em pedaços menores. Isto pode, possivelmente, reduzir a efetividade do sistema de detecção de intrusão.
- Os sistemas baseados em padrões ou assinaturas estarão sempre um passo atrás dos ataques mais modernos, porque o padrão só será configurado após a ocorrência dos primeiros eventos do tipo.
- Não são efetivos em redes chaveadas.
- Encriptação. Os IDS's não são capazes de analisar dados criptografados, pela falta da chave. Isto permite a ocorrência, sem detecção, de ataques escondidos em conexões encriptadas.
- Não são capazes de compensar mecanismos de proteção falhos ou mal-configurados ou não existentes na infraestrutura de proteção.
- Não são efetivos contra técnicas de evasão ou ataques sofisticados.
- Não possuem a capacidade de investigar ataques automaticamente, sem intervenção humana.

5. Redes de alta velocidade e detecção de intrusão

Grande parte dos atuais sistemas de detecção de intrusão não é capaz nem mesmo de lidar confiavelmente, em tempo-real, com links Fast Ethernet saturados. À medida que as velocidades aproximam-se de Gigabit Ethernet, o desafio é cada vez maior para as implementações atuais de Sistemas de Detecção de Intrusão.

Isso pela necessidade de se tratar todo o tráfego sem perdas. Mesmo nos sistemas que não operam em tempo-real, é preciso tratar cada um dos pacotes para extrair as informações pertinentes e armazená-las. Para isso, cada pacote precisa ser processado para extração do cabeçalho e, em alguns casos da própria carga útil ou parte dela. Em interfaces de alta velocidade que possuam carga significativa de tráfego, não é possível aos sistemas atuais, em função de limitações de hardware, processar a tempo todos os pacotes; as perdas que ocorrem prejudicam a obtenção de resultados confiáveis.

Pelas razões apresentadas acima, existem algumas tentativas no sentido de encontrar alguma forma de lidar com grandes volumes de tráfego em redes de alta velocidade. Vamos entender aqui, por simplicidade, redes de alta velocidade como sendo Gigabit Ethernet ou superior.

Trabalhos significativos na área: alguns esforços importantes na área são:

Caso 1: Stateful intrusion detection for high-speed networks **-** *Reliable Software Group – University of California, Santa Barbara* [Kru2002]

Caso 2: low-cost network intrusion detection

Esta proposta visa também enfrentar os desafios apresentados pelas redes de alta velocidade e grande volume de tráfego. É uma solução de detecção estatística de anomalias. [Taylor et al.]

6. Netflow

A proposta aqui apresentada busca lidar com a deficiência da captura de pacotes em redes de alta velocidade fazendo uso de uma facilidade presente nos atuais equipamentos de rede da maioria dos fabricantes, o netflow.

O que é um fluxo:

Um fluxo é identificado como uma sequência unidirecional de pacotes entre um dado par origem-destino, ambos definidos pelo endereço IP na camada de rede e pelos números de porta na camada de transporte. Específicamente, um fluxo é uma combinação dos seguintes campos:

- Endereço IP de origem
- Endereço IP de destino
- Número de porta de origem
- Número de porta de destino
- Tipo de protocolo – camada 3
- ToS byte
- Interface lógica de entrada (ifIndex)

Estes sete campos definem univocamente um fluxo. Se um fluxo tem um campo diferente de outro fluxo, então é considerado um novo fluxo. Um fluxo contém outros campos que dependem do formato de regstro da versão considerada.

Netflow

O NetFlow [Cisco02] é uma ferramenta de monitoramento de tráfego desenvolvida pela CISCO NETWORKS. A tecnologia Netflow é parte integral do IOS, sistema operacional dos equipamentos CISCO, e coleta e mede dados à medida em que chegam a interfaces específicas de roteadores e switches.

Arquitetura

O Netflow inclui três componentes chave que executam as seguintes atividades:

- Cache de Fluxo: coleta fluxos IP entrantes em interfaces de roteadores ou switches e prepara os dados para exportação.
- Coletor de Fluxo: captura dados exportados de múltiplas fontes, agrega e armazena.
- Analisador: Ferramenta de análise de tráfego de rede para recuperar, apresentar e analisar dados Netflow coletados.

A figura abaixo [Cisco01] ilustra o formato do registro do Netflow:

byte 0	*byte 1*	*byte 2*	*byte 3*
Flow Entry			
source IP address			
destination IP address			
next hop IP address			
input intf index		output intf index	
packets			
bytes			
start time of flow			
end time of flow			
source port		destination port	
pad	TCP flags	IP protocol	TOS
source AS		destination AS	
src netmask	dst netmask	padding	

Figura 6. Registro netflow

Exportando NetFlow

Todos os dados circulando através do roteador (apenas alguns dos campos dos cabeçalhos dos pacotes) ou switch são armazenados temporariamente no roteador e, após a expiração, são agrupados por fluxos em datagramas UDP para serem exportados para um coletor. Apesar da falta de garantis de entrega, o UDP é usado por ser mais rápido e simples do que o TCP.

Versões

Muitos fabricantes de equipamentos de rede têm implementado suas próprias versões de Netflow. Atualmente, diversas versões são usadas: A versão 1 não é mais suportada por roteadores, a versão 5 é a mais completa, as versões 6 e 7 são usadas em switches, a versão 8 é uma versão de agregação em router e a versão 9 é a mais recente, sendo versátil e extensível. A versão pública atual de monitor de Netflow suporta o Netflow não agregado (versões 1, 5, 6 e 7) e é capaz de analisar dados de equipamentos de diversos fabricantes.

Fabricantes

Os seguintes fabricantes possuem equipmentos capazes de exportar Netflow:

- Cisco Networks
- Enterasys
- Extreme Networks
- Foundry Networks
- Juniper Networks.
- Riverstone Networks.

Atributos

O Netflow não agregado cria um registro de fluxo contendo informações detalhadas descrevendo cada fluxo IP. Existe, entretanto, alguma variação na implementação conforme o equipamento e o fabricante, podendo-se observar a falta de parte ou da totalidade de algum campo. A tabela a seguir lista os atributos de um registro da versão 5 do Netflow e identifica algumas questões relativas a cada campo:

- Intervalo de amostragem.
- Sete tipos básicos – necessária descrição não agregada do fluxo IP. Algumas implementações permitem mascaramento de endereços IP.

- Próximo Hop.
- Índice SNMP de I/O ifIndex – Algumas implementções não apresentam informação de interface. Isto significa que os fluxos não podem ser atribuídos às interfaces e a análise de uso da interface não é possível.
- Pacotes e octetos.
- Tempos inicial e final – Tipicamente os tempos são expressos em centésimos de segundo, embora algumas implementações usem segundos.
- AS de origem e destino – Importante no monitoramento de roteadores BGP.

7. Ambiente para pesquisa

Foi montada uma arquitetura composta por um roteador Enterasys com netflow v.5, exportando os dados para um coletor-analisador instalado em um HP tc2120 com 768MB de memória e processador P4 de 2.4 GHz. O sistema operacional utilizado no coletor-analisador é o Linux SUSE versão 9.1.

Para tratamento dos dados foi instalado o pacote de ferramentas flow-tools, desenvolvido na Ohio State University [Fullmer2000].

8. Considerações sobre serviços e anomalias

Um sistema de detecção de intrusão é um componente importante, mas não único, da arquitetura de segurança de uma rede. Como tal, sua função deve ser complementar à dos outros elementos. Como ferramenta complementar, sua função será direcionada para tipos específicos de ameaças.

Considerando os bloqueios implementados nas entradas das redes pelos sistemas de Firewall e a relação de serviços e hosts autorizados, procuram-se estabelecer perfis de tráfego considerados normais. Os serviços considerados no trabalho serão, em princípio, aqueles comumente autorizados cujo mal-uso se pretende detectar, tenham eles tanto origem externa como interna.

Uma decisão importante, na implementação de um sistema de detecção de intrusão, é a de quais tipos de intrusão queremos detectar. Neste trabalho, vamos focalizar, principalmente, a detecção de "backdoors" e o uso de serviços lícitos para atividades ilícitas.

Alguém poderia, por exemplo, configurar um serviço SSH na Internet, escutando na porta 443, e configurar um cliente SSH na Intranet para acessar aquele serviço. Tal arranjo torna virtualmente impossível, para qualquer administrador, a detecção da real natureza do tráfego. Se o administrador, entretanto, tem instrumentos para identificar a ocorrência de desvios de determinadas características associadas ao tráfego normal da rede, ele pode tomar ciência desses eventos e proceder a uma investigação mais detalhada já direcionada para a identificação da causa da perturbação.

Obviamente, uma irregularidade natural a buscar será também a existência de serviços não autorizados. Assim, todo o tráfego associado a portas que não correspondem a serviços conhecidos ou autorizados será considerado para fins de emissão de alertas.

A seguir, apresentamos algumas considerações relativas aos serviços normalmente autorizados:

FTP: Como a função do FTP é a transferência de arquivos, o comportamento esperado de conexões FTP é apresentar ocorrência maior de pacotes com grandes quantidades de bytes transferidos, e uma baixa ocorrência de pacotes com "payload" pequeno. Assim, para as conexões FTP estaremos observando desvios em relação a esse comportamento, que podem indicar um uso indevido como, por exemplo, um "backdoor".

SSH: O SSH tem duas finalidades específicas: o SSH propriamente dito e o SFTP, com características de tráfego distintas. O SFTP será analisado, em princípio, como o FTP, enquanto o SSH, que deve, normalmente, apresentar pequena quantidade de bytes transferidos por pacote, será analisado quanto a desvios em relação a esse padrão.

*HTTP:*O HTTP pode apresentar perfis de tráfego distintos em função do conteúdo disponibilizado pelos web-servers. Por essa razão, é importante um conhecimento dos web-servers autorizados na rede e a elaboração do perfil.

SMTP: As conexões SMTP originadas de fora da rede são, numa rede com firewall configurado convenientemente, autorizadas apenas para alguns hosts. Aquelas originadas internamente para servidores SMTP externos são permitidas e pode-se estabelecer um padrão normal de tráfego contra o qual serão buscadas as anomalias. Um aumento grande do número de conexões pode significar, por exemplo, uma infestação por algum worm em equipamentos de usuários.

9. Um resultado prático

Caso: tráfego anormal entre redes do campus do Inpe: A partir de suspeita levantada pelo administrador de uma rede departamental, usamos os dados de netflow para avaliar um volume anormal de tráfego em sua rede, que durou diversos dias. Não vamos entrar em detalhes quanto ao uso das ferramentas, mas filtrando os dados por data, hora e rede de origem e ordenando por quantidade de dados transferidos temos o seguinte:

```
host: # flow-cat /var/netflow/ft/ft-v05.2004-09-15.20* | flow-nfilter -f flow.acl -F depxx_out | flow-stat -f10 -S3

# --- ---- ---- Report Information --- --- ---
#
# Fields:   Total
# Symbols:  Disabled
# Sorting:  Descending Field 3
# Name:     Source/Destination IP
#
# Args:     flow-stat -f10 -S3
#
#
# src Ipaddr      dst Ipaddr      flows          octets          packets
#
150.163.xx.yy   150.163.27.8    25             311827069       208377
150.163.xx.zz   150.163.105.9   826            1630183         19244
```

```
150.163.xx.yy   150.163.105.9   108          248737          3054
150.163.xx.yy   146.164.48.5    57           124901          1578
150.163.xx.yy   129.6.15.28     56           113037          1374
150.163.xx.yy   200.19.119.69   50           111820          1360
```

Agora, filtrando por porta de destino:

```
host:~ # flow-cat /var/netflow/ft/ft-v05.2004-09-15.21* |flow-nfilter -f flow.acl -F
depout|flow-stat -f5 -S3
# --- ---- ---- Report Information --- --- ---
#
# Fields:   Total
# Symbols:  Disabled
# Sorting:  Descending Field 3
# Name:     UDP/TCP destination port
#
# Args:     flow-stat -f5 -S3
#
#
# port     flows            octets           packets
#
25        56              328565768        219412
53        642              458270           6094
123       464              227188           2783
```

Com esses relatórios identificamos origem, destino, porta de destino e ficou fácil determinar a causa do tráfego anormal, que foi o envio inadvertido de uma certa quantidade de mensagens de correio eletrônico de tamanho exagerado.

10. Conclusão

Como a pesquisa está usando dados reais de tráfego numa rede operacional, tem sido possível testar o uso do netflow para avaliação de situações suspeitas em conjunto com outras ferramentas já em uso no ambiente, como o Shadow e o Snort. Resultados preliminares apontam as seguintes vantagens:

a. Eficiência: Baixo volume de dados em comparação com a captura de pacotes (Shadow);
b. Maior facilidade de manuseio, pelo menor volume de dados e maior flexibilidade para análises, com resultados mais rápidos e objetivos;

11. Bibliografia

[Kru2002] Kruegel, C. et alii, *Stateful Intrusion Detection for High-Speed Networks,* Proceedings of the IEEE Symposium on Security and Privacy, IEEE, 2002.

[Allen2000] Allen, J. et alii, State of the Practice of Intrusion Detection Technologies, Technical Report CMU/SEI-99-TR-028, Carnegie Mellon Univ., Software Engineering Inst., Pitsburgh, 2000.

[Shadow] http://www.nswc.navy.mil/ISSEC/index.html.

[Paxson1999] Paxson, V., Bro: A System for Detecting Network Intruders in Real-Time, Computer Networks, 31(23-24), pp. 2435-2463, 14 Dec. 1999.

[Snort] http://www.snort.org/docs/snort_manual/

[Cisco01] Cisco IOS NetFlow Technology Data Sheet, Cisco Systems.

[Cisco02] Netflow Services and Applications, Cisco Systems.

[Fullmer2000] Fullmer, M. & Romig, S., The OSU Flow-tools Package and Cisco NetFlow Logs, Proceedings of the LISA 2000.