1
Introdução

1.1.
Segurança em Redes de Computadores

A crescente dependência das empresas e organizações modernas a sistemas computacionais interligados em redes e a Internet tornou a proteção adequada destes sistemas uma prioridade. O conjunto de protocolos TCP/IP, utilizado na Internet, se tornou o padrão de fato para praticamente todas as redes de computadores e sistemas interconectados. Redes e sistemas móveis como os Sistemas de Celulares estão em processo de integração às redes TCP/IP e à Internet através de tecnologias de 2,5 e 3ª Geração. Redes sem fio (Wi-Fi) baseadas em padrões (IEEE 802.11) do IEEE (Instituto dos Engenheiros Eletricistas e Eletrônicos) também estão se popularizando e permitindo o acesso dos mais variados dispositivos a redes TCP/IP. Empresas operadoras de serviços de telecomunicações estão substituindo sistemas de voz e telefonia convencionais por versões que suportem o transporte de voz sobre o protocolo IP (VoIP ou Voz sobre IP e Telefonia IP). Paralelamente ao crescimento destas redes e da Internet, observou-se também um enorme crescimento em técnicas e ferramentas de ataques e intrusões a redes baseadas neste conjunto de protocolos. Ataques contra a confidencialidade, integridade e disponibilidade das redes de computadores e dos serviços oferecidos são extremamente comuns e cada vez mais poderosos e sofisticados. O objetivo deste trabalho é investigar a aplicação de técnicas de inteligência computacional – em especial duas classes de redes neurais artificiais – ao problema da detecção de padrões de ataques e intrusões em redes de pacotes baseadas no conjunto de protocolos TCP/IP.

Investimentos realizados em soluções e serviços em segurança da informação têm crescido anualmente. Relatório [1] intitulado “Enterprise Security Product Markets” estima que a indústria de produtos e serviços em segurança da inf*ormação crescerá de U$ 7,1 bilhões em 2002 para U$ 13,5 bilhões em 2006. Tradicionalmente, os gastos com segurança da informação se concentram em soluções e serviços preventivos como “firewalls”, esquemas de

criptografia e autenticação forte, filtros de conteúdo, dentre outros. É necessário, entretanto, ser capaz de detectar ataques e intrusões. Estes ataques podem ocorrer por falhas não previstas nas medidas preventivas adotadas ou, até mesmo, através de sistemas ou subsistemas não protegidos originalmente ou com vulnerabilidades originalmente desconhecidas. Estatísticas compiladas [2] pelo CERT/CC (*Computer Emergency Response Team/Command Center*) órgão sediado na universidade de Carnegie Mellon nos Estados Unidos, e apresentadas no gráfico da figura 1.1, demonstram o enorme crescimento no número de vulnerabilidade de segurança descobertas em softwares comerciais.

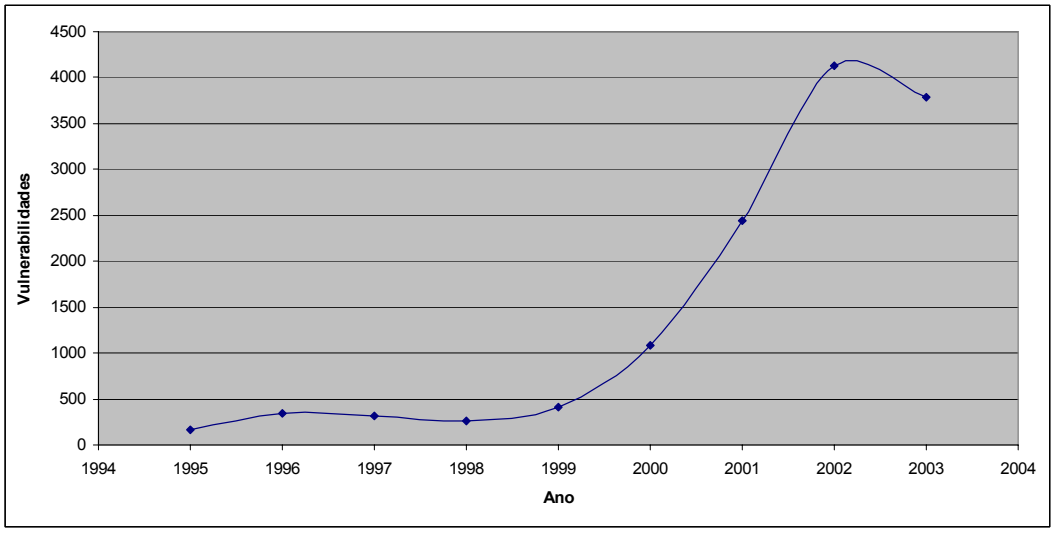

Figura 1-1 : Vulnerabilidades CERT/CC

O gráfico mostrado na figura 1.2, também referente a dados coletados pelo CERT/CC, demonstra a evolução do número de incidentes de segurança em redes reportados a esta organização por empresas e entidades governamentais nos Estados Unidos.

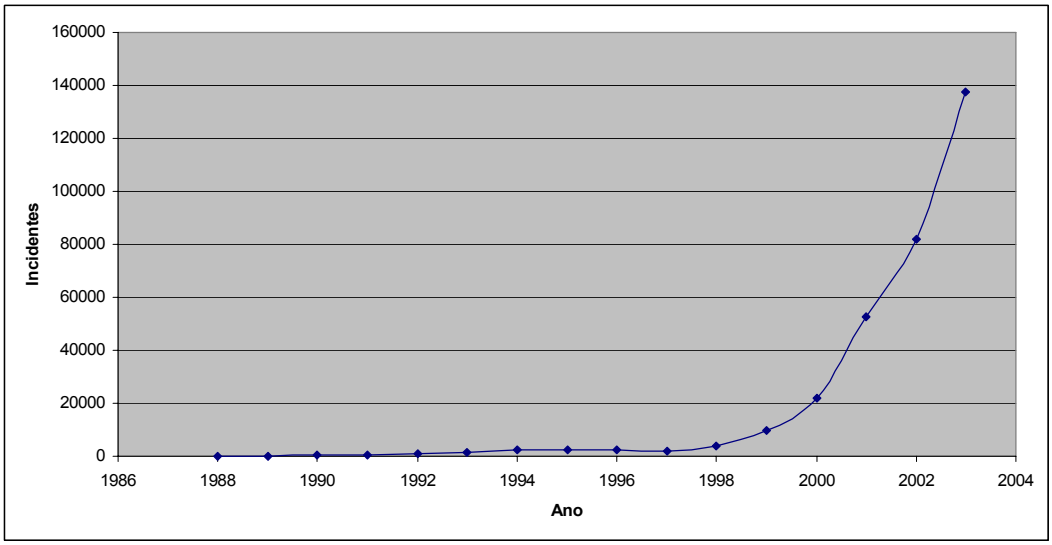

Figura 1-2 : Incidentes CERT/CC

Parcela considerável destes incidentes reportados ocorreu devido a existência de vulnerabilidades, ou falhas, nos softwares utilizados. Alguns incidentes tiveram e tem repercussão em mídia nacional e, muitas vezes, mundial. Como em 1989, quando a Internet ainda era essencialmente uma rede com fins militares e de pesquisa, e registrou-se o primeiro grande incidente de segurança. Um *Worm*[1] [3] [4] [5] que explorava falhas nos softwares *sendmail* e *fingerd*.em sistemas operacionais Unix infectou e tornou inoperante 10% de todos os sistemas conectados a Internet. Ou em julho de 2001, quando aproximadamente 359.000 computadores conectados a Internet foram contaminados pelo *Worm* CodeRedv2 [6] em um intervalo inferior a 14 horas. O CodeRedv2 explorava vulnerabilidades existentes no software para servidor Web denominado "*Internet Information Services"* da empresa Microsoft para infectar suas vítimas. O prejuízo mundial com sistemas inoperantes, recursos empregados para impedir a propagação do *Worm* e removê-lo de computadores já infectados foram estimados em U$ 2,6 bilhões de dólares [6]. As mesmas vulnerabilidades exploradas em massa pelo CodeRedv2 podiam já estar sendo exploradas por criminosos em busca de vantagens em escala menor e em ataques direcionados. A grande maioria destes incidentes, entretanto,

[1] *Worm* é o nome tipicamente usado para identificar um tipo, entre diversos, de software malicioso. Um *Worm* é um software malicioso independente (não precisa de software hospedeiro como o Vírus) e que é capaz de se propagar, usualmente através da rede, para outros computadores/sistemas.

dificilmente são detectados, e na grande maioria das vezes não são reportados a organizações como o CERT/CC.

No Brasil, o NBSO (*NIC BR Security Office*) [7] é o grupo de resposta a incidentes de segurança criado pelo Comitê Gestor da Internet no Brasil. Sua função é receber, analisar e responder a incidentes de segurança em computadores, envolvendo redes conectadas à Internet brasileira. O NBSO, assim como o CERT/CC, também mantém estatísticas referentes aos incidentes e ataques a ele reportados. Os gráficos apresentados nas figuras 1.3 e 1.4 a seguir apresentam números divulgados pelo NBSO para o período de julho a setembro de 2004.

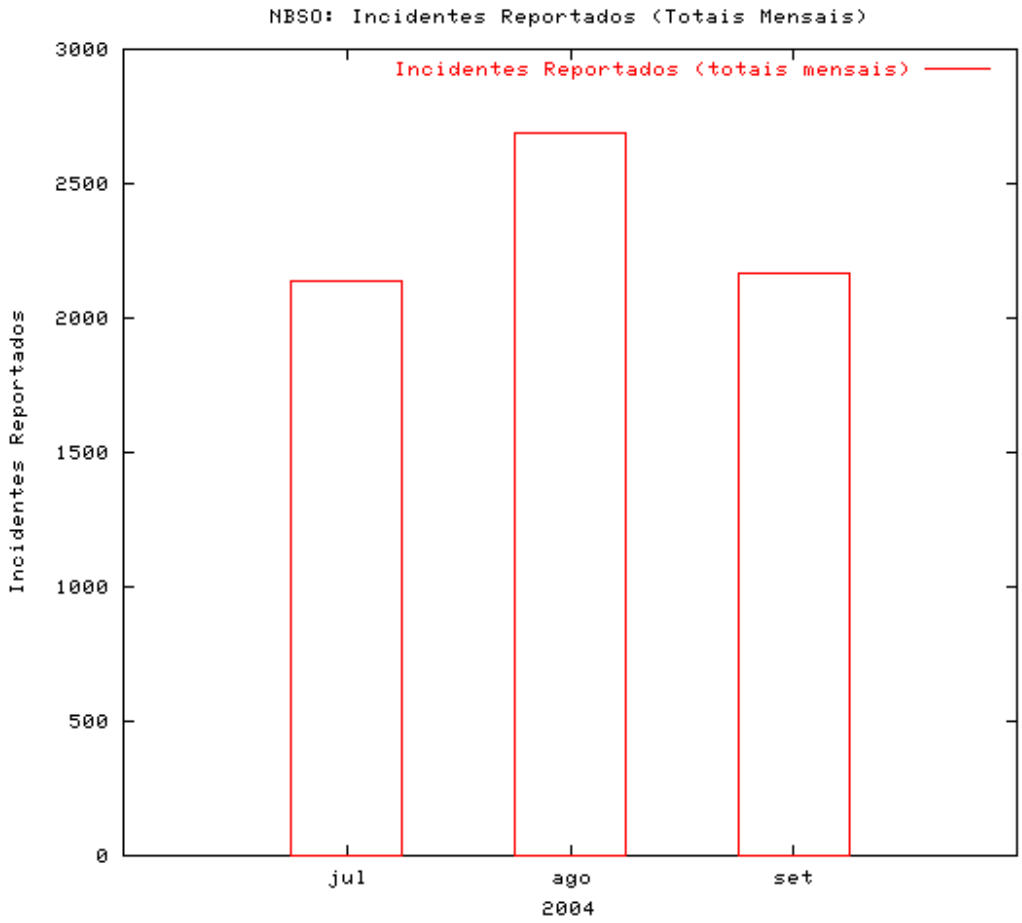

Figura 1-3 : Incidentes NBSO (mês)

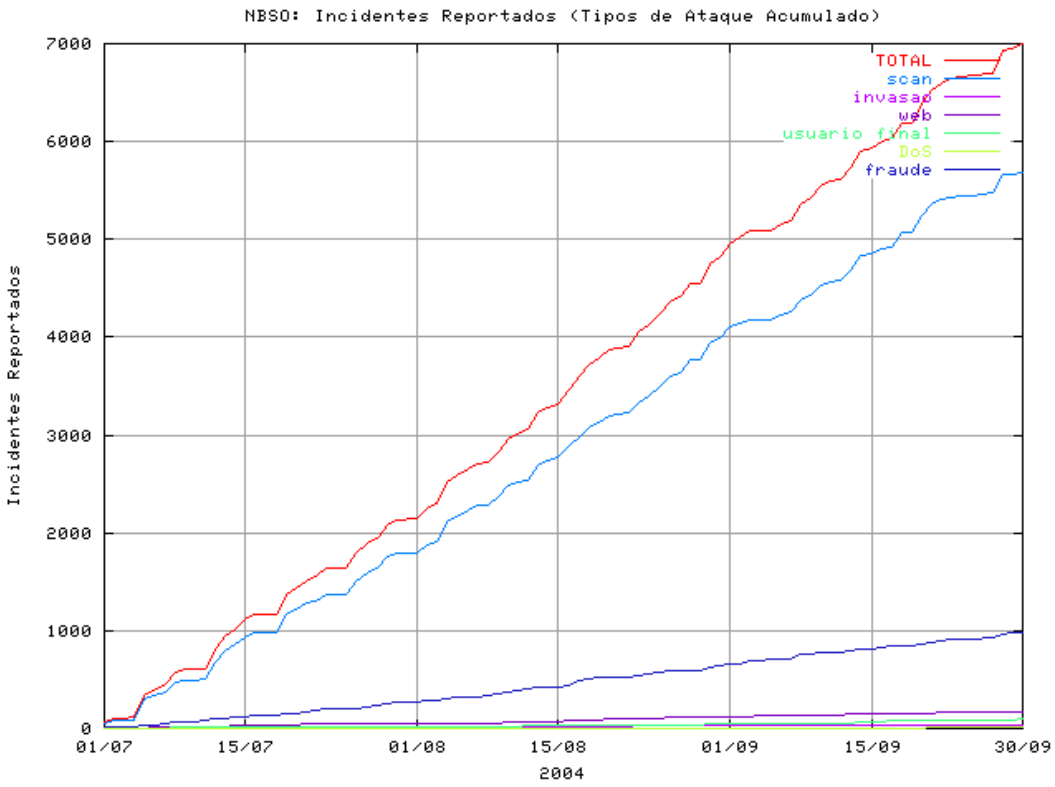

Figura 1-4 : Ataques NBSO (Acumulado)

Tanto os dados do CERT/CC norte-americano quanto os dados divulgados pelo NBSO são uma amostra dos problemas relativos à segurança da informação enfrentados atualmente. Problemas estruturais relativos ao processo de desenvolvimento de software precisam ser solucionados para reduzir o enorme número de vulnerabilidades existentes. Medidas preventivas devem continuar sendo adotadas por organizações conectadas a Internet. Entretanto, brechas de segurança sempre irão ocorrer. A capacidade de detectar rapidamente uma situação de uso indevido de recursos computacionais permite que o dano potencial seja eliminado ou minimizado, além de possibilitar uma reavaliação e alteração das medidas preventivas em vigor. Devido ao enorme volume de dados a serem analisados, sistemas computacionais especialistas denominados Sistemas de Detecção de Intrusão são empregados para auxiliar na detecção de ataques e uso inapropriado. Tradicionalmente, estes sistemas dependem de conhecimento humano especialista codificado ("*hard-coded*") internamente. Esta abordagem limita a atuação e eficácia destes sistemas quando novos padrões de ataques e intrusões surgem, obrigando a uma constante atualização deste conhecimento especialista para manter a utilidade do sistema.

1.2. Trabalhos anteriores

Pesquisa em técnicas e mecanismos para detecção de intrusão em sistemas computacionais vem sendo realizada há cerca de três décadas. [10] é uma excelente referência para pesquisas em detecção de intrusão em sistemas computacionais. Artigo técnico clássico de James P. Anderson [11], em que os tipos de atacantes contra sistemas computacionais são definidos, é citado como precursor das pesquisas nesta área. Denning [9] demonstrou que intrusões e padrões de ataque poderiam ser descobertos por sistemas que determinassem o comportamento normal de um sistema computacional e alertasse para quaisquer desvios encontrados. O foco destes primeiros trabalhos envolvia a detecção de intrusões em um "*host*" ou computador específico, através da monitoração constante de variáveis aleatórias associadas a este computador. [11] emprega redes neurais MLP para, através dos comandos de sistema operacional tipicamente digitados por cada usuário do sistema, detectar uma seqüência de comandos excêntrica e que represente um padrão de intrusão.

Com o surgimento da Internet e a popularização das redes de computadores, o foco das pesquisas se dividiu entre computadores específicos e as redes de pacotes que os interligam. [13] apresentou uma solução baseada em mapas auto-organizáveis (Kohonen) para detecção de intrusão em redes TCP/IP.

Empregar redes neurais artificiais que requerem algoritmo de treinamento supervisionado, para detectar intrusão em redes de pacotes TCP/IP, depende da existência de base de dados com padrões normais e padrões de ataques identificados. O emprego de técnica de inteligência computacional baseada em redes neurais artificiais ao problema de detectar atividades intrusivas em redes de pacotes TCP/IP é o objeto deste trabalho. Para isto será utilizada uma base de dados de treinamento e testes desenvolvida em projeto MIT/DARPA e disponibilizada na Internet. É objetivo explorar a capacidade de generalização de conhecimento previamente aprendido pelas redes neurais artificiais para a detecção de novos padrões de ataques e intrusões. Finalmente será empregada a técnica de decisão por comitê de redes neurais especialistas no intuito de melhorar a performance do sistema de detecção proposto.

1.3. Organização do texto

Os capítulos que compõem esta dissertação estão organizados da seguinte forma : o capítulo 2 fornece uma visão geral sobre sistemas de detecção de intrusos e os ataques/intrusões que estes sistemas devem detectar, bem como os problemas e dificuldades normalmente enfrentados por estes sistemas.

O capítulo 3 apresenta uma taxonomia para ataques computacionais e discute exemplos de ataques específicos que compõem as quatro classes macro de ataques selecionadas para este trabalho. O conjunto de ataques e intrusões utilizados neste trabalho foram catalogados em projeto realizado em 1998 e 1999 pelo Laboratório Lincoln do MIT – Massachussets Institute of Technology - e patrocinado pelo governo americano através da DARPA – Defense Advanced Research Projects Agency.

O capítulo 4 apresenta uma introdução as redes neurais artificiais, com ênfase nas redes Perceptron de múltiplas camadas (Multi-Layer Perceptron ou MLP destacando as principais características que motivaram sua aplicação a detecção de intrusão e ataques em redes TCP/IP. Será discutido ainda a aplicação do conceito e de técnicas de Comitê de Redes Neurais Artificiais Especialistas.

O capítulo 5 apresenta a arquitetura do sistema de detecção de intrusos baseado em redes neurais artificiais proposto, juntamente com os diversos cenários analisados, detalhando todos os procedimentos e técnicas empregados.

Finalmente, o capítulo 6 apresenta os resultados obtidos para cada cenário estudado e o capítulo 7 as conclusões finais sobre este trabalho e sugestões para trabalhos futuros.