2
Sistemas de Detecção de Intrusão

2.1.
O que é Detecção de Intrusão ?

Intrusão foi definida por [8] como “qualquer ação ou conjunto de ações que objetivem comprometer a integridade, confidencialidade e disponibilidade de um sistema”. Podemos ainda definir uma intrusão, de forma mais abrangente, como qualquer violação a política de segurança de um sistema. De acordo com [9], detecção de intrusão é “o processo de monitoração de eventos que ocorrem em um sistema computacional ou rede de sistemas computacionais, analisando-os em busca de problemas de segurança”. Sistemas especialistas e automatizados são empregados na tarefa de detecção de intrusão. Estes Sistemas de Detecção de Intrusão – SDI - irão, normalmente, notificar um especialista humano sempre que detectarem alguma atividade considerada suspeita ou fora dos padrões normais. Detecção de intrusão em sistemas computacionais é uma tecnologia relativamente nova. Pesquisas nesta área se iniciaram a partir de 1980. Sistemas de Detecção de Intrusão mais recentes podem, opcionalmente, atuar automaticamente ao detectar uma anomalia. O desenvolvimento de sistemas de detecção de intrusão em tempo real é motivado, segundo Denning [9] , por quatro fatores primordiais:

- A maioria dos sistemas computacionais utilizados possui alguma falha de segurança que pode ser explorada por usuários mal intencionados. Pesquisar e corrigir todas estas potenciais falhas de segurança não é viável técnica e comercialmente.
- Inviabilidade de substituição de sistemas com falhas de segurança por versões mais seguras.
- Desenvolver sistemas completamente livres de falhas de segurança, mesmo que isto seja uma meta crítica do projeto, é considerado tarefa virtualmente impossível.
- Mesmo sistemas altamente seguros podem ser comprometidos por usuários internos que abusem de seus privilégios.

2.2. Tipos de Sistemas de Detecção de Intrusão

Sistemas de detecção de intrusão são tipicamente divididos em duas categorias – baseado em conhecimento ou baseado em comportamento - em relação a estratégia de detecção adotada e em outras duas categorias – baseado em redes ou baseado em *host* - em relação ao escopo e fonte de informações analisadas. Uma classificação mais completa, entretanto, está apresentada na figura 2.1.

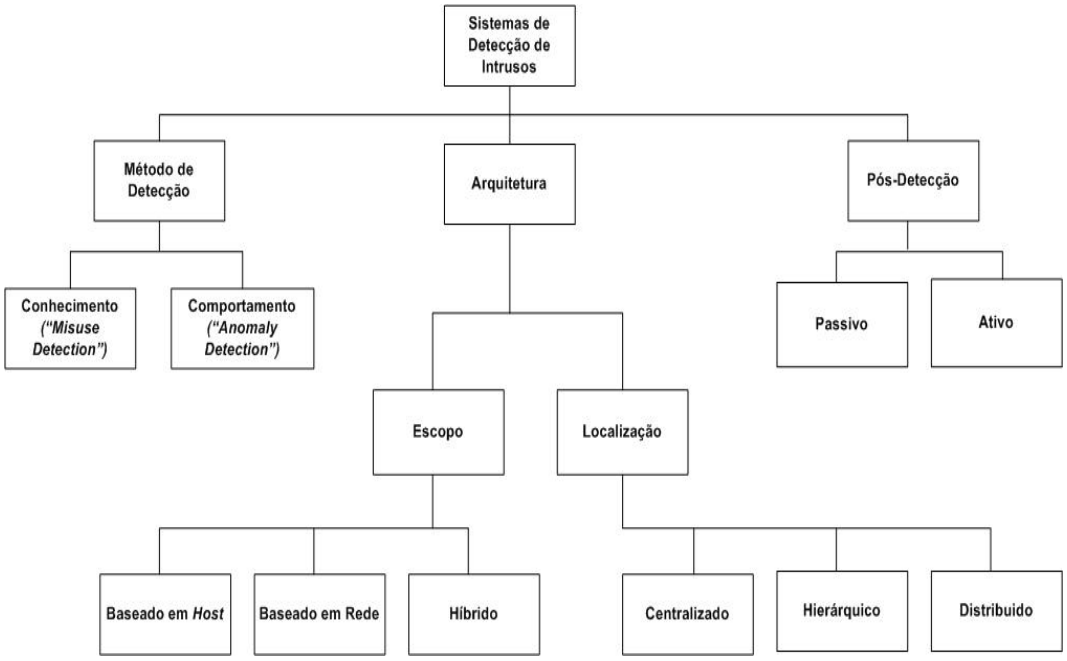

Figura 2-1 : Classificação SDI

2.2.1. Método de detecção

Quanto ao método de detecção empregado um sistema de detecção de intrusão pode ser classificado como baseado em uso indevido (ou conhecimento) ou baseado em anomalia (ou comportamento). Existem ainda sistemas híbridos que combinam as duas técnicas.

Detecção por uso indevido (*"Misuse Detection"*)

Sistemas de detecção de intrusos por uso indevido ("misuse detection") são sistemas que procuram por padrões de ataques e intrusões previamente conhecidos. São também conhecidos como sistemas de detecção baseado em padrões de assinaturas de ataques ou em sistemas de detecção de intrusão baseado em conhecimento. A premissa básica dos sistemas que se enquadram nesta classe é que existem ataques com características precisas, bem definidas e que podem ser facilmente codificadas em um sistema especialista. As falhas nos softwares *fingerd* e *sendmail* exploradas pelo *Worm Morris* [3] [4] [5] em 1989 e as falhas no servidor Web "*Internet Information Services*" da Microsoft exploradas mais recentemente pelo *Worm CodeRed* [6] são exemplos de ataques cujos padrões podem ser codificados como conhecimento em um sistema de detecção baseado em uso indevido. Sistemas de detecção desta classe normalmente possuem uma base de assinaturas de ataques conhecidos que deve ser constantemente atualizada a medida que novos ataques sejam descobertos.

Um exemplo de regra codificada em um SDI baseado em uso indevido, para detecção de um ataque de negação de serviço denominado "*Ping of Death*" (ping da morte – envio de um pacote ICMP Echo Request muito grande que causa indisponibilidade dos serviços de rede em vários sistemas operacionais devido a *bug* ou erro de programação presente nos softwares vulneráveis) consistiria em considerar como ataque todo pacote "ICMP Echo Request" maior do que 64000 bytes. Snort [14] é um SID baseado em uso indevido que possui uma linguagem própria para criação de regras representando intrusões ou atividades suspeitas. Uma regra Snort capaz de detectar atividades suspeitas como a gerada pelo já mencionado *Worm* CodeRedv2 seria :

ALERT TCP ANY ANY -> ANY 80 (CONTENT: ".IDA?"; DSIZE: > 289; MSG: "ATAQUE CODEREDV2 GENÉRICO"; FLAGS: A+; NOCASE;)

Basicamente, a regra monitora por pacotes TCP de qualquer endereço e porta de origem (ANY ANY), para qualquer endereço de destino desde que a porta de destino seja 80 (ANY 80 - Porta 80 no protocolo TCP é normalmente associada ao serviço HTTP) que possua tamanho (DSIZE corresponde a "*Datagram Size"*) maior que 289 bytes e apresentem em seu conteúdo a seqüência de caracteres ".IDA". Encontrando um pacote que satisfaça todas

estas condições o Snort gerará um alerta com a mensagem “Ataque CodeRedv2 Genérico”. Um sistema de detecção como Snort é normalmente configurado para verificar milhares de regras semelhantes a usada neste exemplo, e deve ter esta base de regras ou de conhecimento constantemente atualizada para refletir novos padrões de ataques.

Os dois exemplos apresentados ilustram regras de um SDI baseado em uso indevido e baseado em redes. Estes sistemas também podem ter foco em eventos relacionados a um computador ou sistema específico (baseado em *host*). Neste caso, suas regras lidariam com elementos associados apenas a um computador ou sistema. Um exemplo de regra de um SDI baseado em *host* poderia gerar um alerta se o percentual de ocupação do processador superasse determinado valor por um intervalo de tempo específico, uma determinada seqüência de chamadas a funções do sistema operacional ocorresse e/ou determinados arquivos críticos fossem alterados.

Outra estratégia de representação do conhecimento usado para detecção usualmente empregada para esta classe de SDIs consiste em codificar o conhecimento sobre o que é uma intrusão como uma seqüência de troca de estados do sistema. Pares de atributos-valores representam os possíveis estados em que o sistema em analise pode se encontrar. Ações acarretam transições entre estes estados. Determinadas transições ou seqüências de transições podem ser consideradas como decorrentes de técnicas de intrusão (ver figura 2.2).

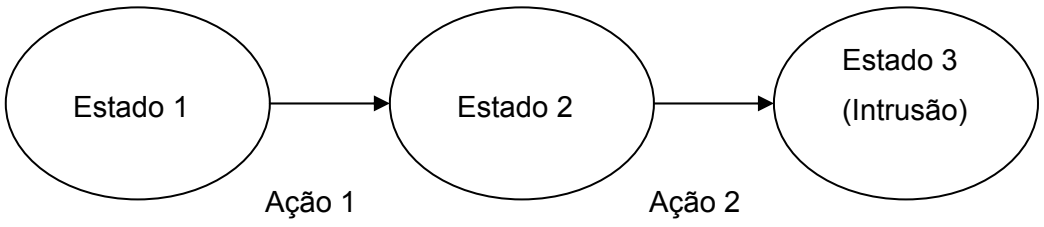

Figura 2-2 : Transição de estados

USTAT [15] [16] – "*Unix State Transition Tool*" – é um exemplo clássico de ferramenta de detecção de intrusão que modela o conhecimento sobre o que é uma intrusão através de transições entre estados do sistema monitorado.

É importante ressaltar que, independentemente do foco (rede ou *host*) os sistemas baseados em uso indevido necessitam de conhecimento especialista humano codificado (*hard-coded*) para operar.

A principal desvantagem dos sistemas de detecção de intrusos por uso indevido é a sua incapacidade de detectar novos ataques, ou mesmo, ataques que ainda não façam parte da sua base de assinaturas e padrões. Este é o principal fator motivador para pesquisa e desenvolvimento de sistemas de detecção de intrusos baseado em anomalias.

Detecção por anomalia ("*Anomaly Detection*")

Sistemas de detecção de intrusão por anomalia, também chamados sistemas de detecção de intrusão baseados em comportamento, procuram determinar ou criar modelos que representem o comportamento normal ou esperado do sistema computacional ou rede em análise e alertam sempre que desvios neste comportamento esperado forem encontrados. A premissa básica destes sistemas é que atividades de intrusão ou ataque fazem parte do subconjunto composto por atividades anômalas. Idealmente, o conjunto de atividades maliciosas será igual ao conjunto de atividades anômalas. Nesta situação o sistema não gerará nenhum falso-positivo e nenhum falso-negativo. Na prática, as seguintes situações podem ocorrer com probabilidade diferente de zero:

- Atividade intrusiva porem normal: esta situação é extremamente grave pois gera a situação de falso-negativo, ou seja, o sistema não detecta um ataque ou intrusão.
- Atividade não intrusiva e normal: : situação denominada positivo-negativo. Atividade normal e não intrusiva, portanto não será detectada pelo sistema.
- Atividade não intrusiva e anômala: situação denominada falso-positivo. O sistema alerta indicando uma intrusão que na verdade não ocorreu.
- Atividade intrusiva e anômala: situação ideal em que o sistema detecta corretamente o ataque ou intrusão.

2.2.2. Arquitetura

A classificação de sistemas de detecção de intrusão quanto a sua arquitetura leva em consideração a fonte de informações de monitoração – rede, "*host*" (um computador específico) ou híbrido – e a forma de distribuição das tarefas e do processamento dos módulos que compõem o SDI.

A figura 2.3 ilustra um exemplo clássico e bastante comum de aspectos de localização de uma solução de detecção de intrusão na prática. Demonstra uma rede local representativa de redes de corporações,universidades,organizações modernas interconectada a Internet por um "firewall" – dispositivo de segurança que realiza a separação entre redes distintas – com uma rede distinta para oferta de serviços na Internet – denominada DMZ ("*Demilitarized Zone*"). Cada segmento de rede presente – local, DMZ e Internet – possui um Sistema de Detecção de Intrusos – SDI - baseado em redes monitorando aquele segmento específico. Complementarmente, servidores considerados críticos possuem Sistema de Detecção de Intrusos baseado em "*hosts*" instalado e monitorando aspectos específicos destes servidores. Finalmente, uma ferramenta de gerência dos SDIs está presente e é capaz de coletar, centralizar informações dos diversos módulos.

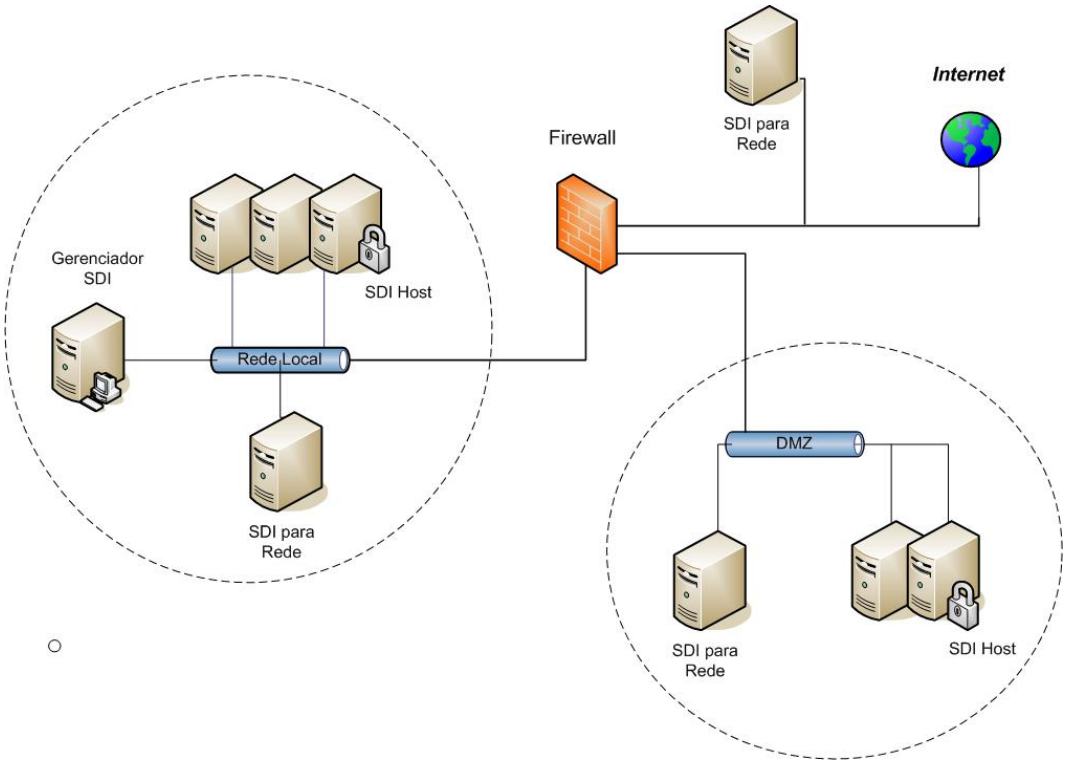

Figura 2-3: Exemplo clássico de solução de detecção de intrusos

Escopo de atuação

Em relação a fonte de informações analisadas e monitoradas para se detectar uma intrusão, um SDI pode ser classificado como SDI para *Host,* SDI para rede ou SDI híbrido.

Detecção para *Hosts (HIDS)*

Sistemas que monitoram e procuram por intrusões em um único computador ("*host*"). Consistem, tipicamente, de sistemas especialistas que monitoram chamadas de funções do sistema operacional, acesso a arquivos de considerados críticos, uso de recursos como processador, disco e memória procurando por padrões que determinem um ataque ou intrusão ou por desvios significativos em relação ao perfil de uso considerado normal e regular. A característica principal e que distingue esta classe é que a visão do SDI está restrita apenas ao "*host*" ou computador no qual ele esta instalado e operando.

Detecção para Redes (*NIDS*)

Sistemas que monitoram segmento ou segmentos de rede procurando por conexões e/ou pacotes maliciosos. São compostos, invariavelmente, por um módulo de captura de pacotes que trafegam na rede – normalmente denominados *sniffers*. Este módulo opera configurando a placa de rede utilizada para monitoração em modo de funcionamento promíscuo. Uma placa de rede em modo de funcionamento normal recebe todos os pacotes até detectar o endereço de destino do mesmo. Se o pacote não estiver endereçado ao próprio computador, ou não for um pacote *broadcast* (endereçado a todos os nós da rede) ou *multicast* (endereçado a um grupo de computadores) endereçado a um grupo ao qual aquele computador pertença, ele será descartado imediatamente. Já em modo promíscuo a placa de rede capturará todos os pacotes observados independentemente do endereço de destino. Para o protocolo TCP/IP as figuras 2.3 e 2.4 ilustram os campos que fazem parte do cabeçalho e serão capturados pelo módulo *sniffer.*

VERS	HLEN	TIPO DE SERVIÇO	COMPRIMENTO TOTAL	
IDENTIFICAÇÃO			FLAGS	DESLOCAMENTO DO FRAGMENTO
TEMPO DE VIDA	PROTOCOLO	VERIFICAÇÃO DA SOMA DO CABEÇALHO		
ENDEREÇO IP DE ORIGEM				
ENDEREÇO IP DE DESTINO				
OPÇÕES IP (SE HOUVER)				PADDING
DADOS				
...				

Figura 2-4 – Cabeçalho IP

PORTA DE ORIGEM			PORTA DE DESTINO
NÚMERO DE SEQÜÊNCIA			
NÚMERO DO RECONHECIMENTO			
HLEN	RESERVADO	BITS DE CÓDIGO	JANELA
SOMA DE VERIFICAÇÃO			PONTEIRO URGENTE
OPÇÕES (SE HOUVER)			ENCHIMENTO
DADOS			
...			

Figura 2-5 – Cabeçalho TCP

O protocolo UDP – *User Datagram Protocol* – não possui o mesmo conceito de conexão que o TCP. Seu objetivo é fornecer um mecanismo de entregas de datagramas sem nenhum tipo de verificação ou garantia de entrega. Várias aplicações se beneficiam deste tipo de serviço. Do ponto de vista de um sistema de detecção de intrusão, entretanto, é interessante criar, mesmo para o UDP, o conceito de uma "conexão virtual". Na prática, os sistemas de detecção de intrusos enxergarão uma "conexão virtual" sempre que uma seqüência de pacotes UDP forem trocados entre dois endereços IPs, dentro de um intervalo de tempo predefinido e mantendo portas de origem e destino fixas. A figura 2.5 apresenta os campos do cabeçalho UDP. Pode-se observar que é uma estrutura de dados bem mais simples que a do TCP.

PORTA DE ORIGEM UDP	PORTA DE DESTINO UDP
COMPRIMENTO DE MENSAGEM UDP	SOMA DE VERIFICAÇÃO UDP
DADOS	
...	

Figura 2-6 – Cabeçalho UDP

Todos os pacotes capturados serão usualmente pré-processados e enviados para análise e inspeção pelo módulo de detecção. Este pré-processamento envolve a codificação dos bits recebidos em campos do cabeçalho dos protocolos correspondentes e pode incluir também o reconhecimento de campos de protocolos de aplicação como HTTP, SMTP, FTP

etc. Normalmente o SDI realizará também a reconstrução de seqüências de pacotes trocados entre dois computadores em registros de conexão. Esta é uma etapa importante para um SDI baseado em rede, pois muitas técnicas de ataques e intrusões não são detectadas analisando-se um único pacote ou analisando vários pacotes sem a manutenção de um contexto entre eles.

Sistemas Híbridos

Sistemas híbridos utilizam como fonte de informações de análise e avaliação tanto características relativas aquele computador específico, quanto pacotes transmitidos e recebidos em rede.

Localização

Implementações práticas de sistemas de detecção de intrusão apresentam uma combinação de agentes de monitoração ou sensores – de rede ou de *host* – com módulos de processamento dos dados capturados, módulos de detecção e ferramentas de gerência. A figura 2.7 ilustra os módulos normalmente encontrados em um SDI. De acordo com a arquitetura e localização destes módulos, estes sistemas podem ser classificados em centralizados, hierárquicos e distribuídos.

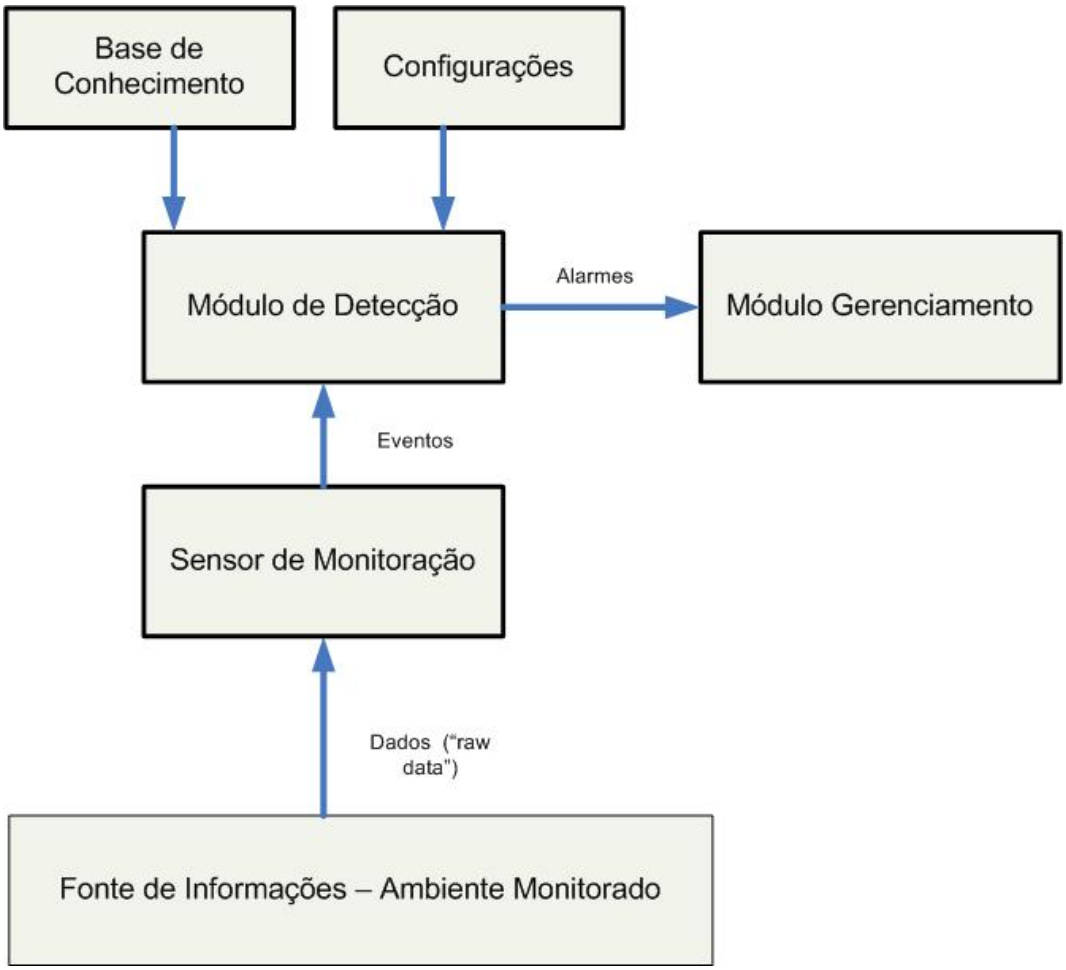

Figura 2-7 : Módulos de um SDI

Centralizado

SDIs centralizados possuem em sua arquitetura, um ou mais módulos de monitoração (também denominados agentes), responsáveis pela coleta dos dados a serem analisados e que transmitem todas estes dados coletados para um módulo central (também denominado gerente) responsável pela análise, detecção e resposta. Este modelo, conhecido como agente-gerente em sistemas de gerência, pode ser empregado independentemente da classificação quanto a estratégia e escopo dos módulos agentes.

Hierárquico

Sistemas hierárquicos, assim como os sistemas centralizados, apresentam o modelo agente-gerente, porém permitindo sua configuração em vários níveis distintos. No sistema centralizado, todos os agentes de monitoração estavam

subordinados (enviavam informações e se reportavam) a um único módulo central ou gerente. Sistemas hierárquicos permitem subdivisões neste modelo. Agentes que coletam informações referentes a rede (NIDS) podem estar subordinados ao módulo gerente de redes. Agentes que analisam computadores específicos (HIDS) podem estar subordinados ao módulo gerente de *hosts*.

Distribuído

Sistemas de detecção de intrusão distribuídos permitem que parte (ou todo) do processamento de análise e detecção seja realizado pelo módulo (agente) de coleta. Distribuem assim a inteligência do sistema entre os módulos gerentes e os diversos agentes que compõem o sistema.

2.2.3. Pós-detecção

Esta classificação dos sistemas de detecção de intrusão leva em consideração as ações tomadas quando um padrão de intrusão é detectado.

Passivo

Sistemas de detecção de intrusão que, ao detectar um padrão de ataque, apenas geram notificações para um operador ou administrador especialista através de console de gerenciamento, são classificados como SDIs passivos. Estas notificações ou alertas podem ser geradas e enviadas através de diversos padrões e protocolos (e-mails para o administrador responsável, alertas SNMP etc). O IETF [17] – *Internet Engineering Task Force* - possui um grupo de trabalho denominado *Intrusion Detection Exchange Format* [18] cuja missão é padronizar os mecanismos de comunicação, bem como a linguagem de comunicação, entre sistemas de detecção e entre agentes e módulos de gerência. Os seguintes documentos RFC – *Request for Comments* foram gerados por este grupo de trabalho com o objetivo de padronizar a troca de informações entre módulos de um SDI ou entre vários SDIs distintos :

- Intrusion Detection Message Exchange Requirements
- Intrusion Detection Message Exchange Format
- Intrusion Detection Exchange Protocol (IDXP)

Ativo

Sistemas de detecção de intrusão podem, ao detectar um padrão de ataque, tentar atuar ativamente no sistema ameaçado para protegê-lo de forma automática da ameaça. Esta classe de sistema tem sido alvo de pesquisas recentes e tem se destacado com denominação própria : sistemas de prevenção de intrusão. SDIs ativos e baseado em redes podem, por exemplo, enviar pacotes de término (pacotes TCP com flags FIN e RST) de conexão para a fonte de fluxos de pacotes que sejam considerados intrusão. Podem possuir integração com roteadores de borda – roteadores localizados no perímetro entre as redes locais de uma organização e redes externas com a Internet - ou Firewalls e realizar a criação de regras que filtrem determinado fluxo de pacotes ou determinado endereço de origem.