Projeto Agiredes: agente inteligente para a análise de registros de conexão de redes

VERÔNICA CONCEIÇÃO OLIVEIRA DA SILVA[1]
ANALÚCIA SCHIAFFINO MORALES DE FRANCESCHI[2]
KAREN SELBACH BORGES[3]

RESUMO

O crescimento dos sistemas integrados e das redes aumenta a necessidade de gerência da infra-estrutura nas organizações. O projeto Agiredes tem desenvolvido agentes para gerência de redes que utilizam técnicas de IA para auxiliar no processo de tomada de decisão dos administradores de redes. O presente trabalho propõe a utilização de técnicas de inteligência artificial no desenvolvimento de um agente para identificação de possíveis ataques às redes privadas, com base na análise de pacotes, freqüência e volume de tentativas, enfocando o registro gerado entre as conexões da rede privada com a Internet. Desta forma, o agente consegue identificar a origem de um possível problema e isolá-lo antes de ocasionar a interrupção do serviço da rede.

***Palavras-chave**: segurança, Internet, inteligência artificial, gerência de redes.*

ABSTRACT

The growth of the system integration and networking increase the necessity of infrasctructure management in the organizations. The Agiredes project has developed agents for network management based on AI

[1] Acadêmico do Curso de Matemática Aplicada à Informática/ ULBRA – Bolsista PROICT/ULBRA

[2] Professor – Orientador do Curso de Matemática Aplicada à Informática/ULBRA (analucia@ulbra.tche.br).

[3] Professor do Curso de Informática/ULBRA

techniques which help them in the decision-making process. The present work proposes the employment of the artificial intelligence techniques for development of an agent able to identify a possible attack in the private network, based on the packages analysis, frequency and volume of attempts, it has been considered the generated register between the private connections and the Internet. The agent will reach the problem identification from origin and thus it can isolate it before causing the degradation of the network service.

Key words: *security, Internet, artificial intelligence, computer network mangement.*

INTRODUÇÃO

O crescimento de sistemas integrados através de redes de computadores e o surgimento de novas tecnologias tem aumentado o investimento em tecnologia da informação (TI) em grande parte das organizações. Os sistemas podem ser interligados através de investimentos realizados pela própria empresa, mantendo um serviço de rede dedicado (isto acontece em organizações de grande porte que possuem investimentos elevados em infra-estrutura de TI). Outra possibilidade é utilizar recursos e serviços terceirizados oferecidos por concessionárias de redes e telecomunicações. Na maioria das organizações de pequeno e médio porte estão presentes estes dois casos, porque é necessário manter uma rede local dentro da organização e utilizar as concessionárias para os serviços de interligação de longa distância.

Entretanto, independentemente do tipo de infra-estrutura, são necessárias atividades de controle e administração dos recursos disponíveis. As operações consistem em intensa monitoração e controle dos ativos de redes (software e hardware), e são conhecidas como gerência de redes. A gerência de redes pode ser realizada a nível de elementos de redes, que são os dispositivos físicos, ou a nível de serviços. Utiliza-se o modelo funcional definido pela OSI para reduzir a complexidade das atividades de gerência, dividindo-a em gerência de falhas, de configuração, de contabilização, de desempenho e de segurança (KUROSE, 2003).

No entanto, as ferramentas de gerência de redes disponíveis no mercado, que atendem a todas as funcionalidades citadas acima, possuem um custo muito elevado. Ferramentas mais baratas, que constituem uma alternativa, são direcionadas apenas para algumas destas funcionalidades. Normalmente, auxiliam no processo de monitoração, coleta dos dados, fornecem relatórios dos mais variados tipos e emitem alarmes após alguma falha predeterminada ter sido detectada. São sistemas passivos que não executam nenhum tipo de tomada de decisões aguardando que o administrador da rede solucione o problema. Estão disponíveis na Internet programas gratuitos para estas atividades, como é o caso do NAGIOS (2006), DO CACTI (2006), LOOK AT LAN (2006), etc. Observa-se, muitas vezes, que mesmo com software disponível as empresas não realizam o gerenciamento da rede de forma satisfatória. Grande parte das organizações de pequeno e médio porte resolvem os problemas à medida que eles surgem e não possuem o costume de realizar as atividades próativas para manter o funcionamento da rede. Atribui-se este fato à falta de conhecimento de certos administradores de redes sobre o assunto. Muitas vezes, os responsáveis pela rede conhecem estes utilitários, mas desconhecem como

utilizar seus recursos para gerenciar os dispositivos de redes.

Em particular, se for considerada a conexão de redes privadas através da rede pública (Internet) é necessário além das ferramentas de gerência, o uso específico de ferramentas de segurança (*firewalls*, *proxy*, DMZ (Demilitarized Zone), IDS (*Intrusion Detect System*)) para garantir o funcionamento e proteção das redes. Atualmente, todas as aplicações geram registros de funcionamento, porém cabe ao administrador analisar esses registros, conhecidos como *log*, e identificar possíveis problemas que afetariam a rede e seus dispositivos.

Os administradores têm por principal objetivo, manter a rede operacional para que a troca de informações seja realizada com agilidade e segurança. Nos últimos anos, muitas ferramentas foram desenvolvidas no intuito de auxiliar o administrador em suas tarefas diárias; tais ferramentas geram um grande volume de relatórios. Os administradores agregam mais atividades a cada dia, reduzindo o seu tempo para análise destes dados.

Especificamente na área de segurança de redes, as ferramentas de IDS como *snort*, *shadow*, *asgaard*, *tripwire* dentre outros, analisam o funcionamento da rede gerando uma *baseline* (perfil de comportamento da rede). Durante o processo de execução o agente verifica o comportamento da rede; se esta apresentar um desvio, serão gerado alarmes de intrusão que podem ou não ser evidências de um ataque. A geração de *baseline* não garante que o comportamento normal da rede seja o que foi capturado no instante na análise.

Este trabalho não está baseado em um comportamento previamente definido, e sim em uma análise dos dados trafegados em tempo real, levando em consideração a captura dos dados do pacote como porta destino e endereço de origem. A partir de cálculos estatísticos é possível considerar um ataque ou congestionamento de circuito e indicar ao administrador a origem para que este proceda conforme necessário, evitando que tentativas falhas de acesso à rede possam causar congestionamento no enlace ou possível estouro de *log,* o que pode ocasionar a interrupção do serviço de acesso à Internet. O objetivo é desenvolver uma ferramenta que auxilie o administrador de rede na identificação de possíveis ataques à rede a partir da análise dos seguintes dados: o conteúdo dos pacotes IP (*Internet Protocol*), a freqüência e o número total de tentativas de conexões. Desta forma, propõe-se uma solução alternativa para aquelas existentes no mercado.

MATERIAL E MÉTODOS

O projeto Agiredes tem como objetivo o desenvolvimento de agentes voltados para a área de gerência de redes de computadores. É possível considerar agentes como um paradigma para o desenvolvimento de sistemas, sejam simples ou complexos (HAYES, 1999). No entanto, existem diversos tipos de agentes, de forma que não é possível considerar com o mesmo nível de classificação a função de um robô (que é uma agente autônomo de hardware) e de um programa coletor de dados de gerência, por exemplo. Apresenta-se a seguir, uma taxonomia para agentes inteligentes no sentido de esclarecer alguns tópicos sobre o desenvolvimento e aplicação de técnicas de inteligência artificial (IA) (DE FRANCESCHI, 2003). Agentes de

hardware normalmente são autônomos e desenvolvidos com o auxílio de técnicas de IA. Segundo Roisenberg (ROISENBERG, 1998) agentes autônomos são definidos como sistemas computacionais que operam em ambientes dinâmicos e imprevisíveis. Eles interpretam dados obtidos pelos sensores que refletem eventos ocorridos no ambiente e executam comandos em motores que produzem efeitos no ambiente (FRANKLIN, 1999). O grau de autonomia de um agente está relacionado à capacidade de decidir por si só como relacionar os dados dos sensores com os comandos dos motores em seus esforços para atingir objetivos, satisfazer motivações, etc. As principais aplicações são encontradas na área da robótica, nas quais destacam-se: limpeza de material nuclear, escavações de ambientes perigosos e exploração de outros planetas (LESSER, 1999). Os agentes de software podem ser encontrados em um número muito maior de aplicações. Eles podem ou não ser desenvolvidos com o auxílio de técnicas de IA. Agentes passivos foram classificados como agentes que não possuem autonomia e não utilizam técnicas de IA. Assemelham-se a simples programas e são construídos através de instruções. É o caso de agentes e gerentes de gerência de redes. Os agentes normalmente coletam informações dos objetos gerenciados e repassam aos gerentes. Os gerentes repassam as informações para os administradores de redes que realizam a tomada de decisões.

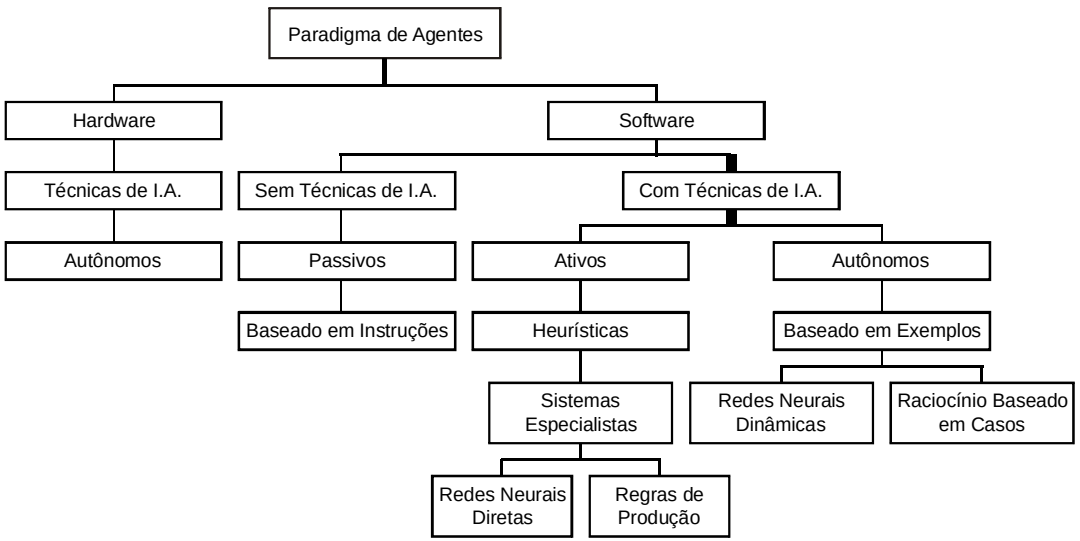

Figura 1 - Paradigma de Agentes

Os agentes de software desenvolvidos com o auxílio de técnicas de IA podem ser ativos ou autônomos (Figura 1). Os agentes ativos são desenvolvidos com o auxílio de heurísticas (ABOELELA, 1999). Na área de gerência de redes, estas heurísticas são fornecidas pelo administrador da rede. Servem para definir regras de produção ou redes neurais diretas para auxi-

liar na solução dos problemas. Assemelham-se aos sistemas especialistas e não possuem autonomia. Agentes autônomos possuem características dinâmicas (FAN, 1997). Devem ser desenvolvidos com o auxílio de técnicas que satisfaçam esta característica. Podem ser desenvolvidos com o auxílio de redes neurais artificiais (RNAs) dinâmicas ou através de técnicas de raciocínio baseado em casos (RBC) (DE FRANCESCHI 2003).

Depois de esclarecidas as diferenças entre os vários tipos de agentes, descreve-se a metodologia que está sendo utilizada para o desenvolvimento de agentes inteligentes no projeto Agiredes. A metodologia utilizada é a seguinte: define-se o problema, que pode ser formalizado através de um objeto matemático do tipo P = {D, R, q}, consistindo de dois conjuntos não vazios, D os dados e R os resultados possíveis e de uma relação binária q Ì D x R, a condição que caracteriza uma solução satisfatória, associando a cada elemento a solução única desejada. Desta forma o problema é representado como uma função.

Segundo Barreto (BARRETO, 2001) existem diversas formas de definir uma função, dentre elas destacam-se enumeração exaustiva, declarativamente, por um programa ou através de exemplos. Quando um problema não é completamente definido para todo valor de seus dados, em que se conhece apenas a definição do problema para um subconjunto dos dados possíveis. Neste caso, a solução não é única: todas as funções que forem iguais dentro da região em que o problema é definido são válidas. Neste caso, é melhor ter uma solução aproximada do que os dados para definir a função. Este trabalho defende que resolver o problema será então encontrar um modo de implementar a função ou aproximá-la com as ferramentas disponíveis. Ou seja, através de exemplos treinar a rede e obter-se valores estimados da solução para os outros valores, utilizando a propriedade de generalização das RNAs.

Com base nesta metodologia, está sendo desenvolvido um agente adaptativo através de exemplos, que utiliza redes neurais recorrentes para encontrar possíveis ataques à rede. O agente concentra-se na análise de dados trafegados, considerando dados contidos no cabeçalho do pacote e outras informações. O conjunto a seguir, está baseado na metodologia descrita anteriormente:

P = {D, R, q}

Onde:

P = probabilidade de ataques no link da rede pública;

D = (elemento de falha, endereço IP de origem, endereço IP de destino, tipo de protocolo, política de segurança);

R = (atividade normal, atividade irregular, atividade suspeita e alerta de invasão);

q é a relação entre os dados e os resultados que deverá ser obtida através do treinamento da rede neural artificial.

A rede neural que será utilizada foi escolhida com base na metodologia e na taxonomia dos agentes apresentada. Após análise do problema, sugere-se o uso de uma rede com ciclos, ou recorrente. Isto porque não existe um conjunto bem definido de entradas e saídas desejadas na definição do problema.

Após esta definição é necessário coletar informações para elaboração do agente. Elas serão obtidas através da análise de diversos registros de *logs* e situações relatadas pelo administrador da rede do cenário de testes. O cenário que está sendo utilizado é de uma rede de uma empresa multinacional do ramo de varejo, que atualmente possui uma única conexão com a rede pública e enfrenta os problemas de segurança, como àqueles citados inicialmente. A infra-estrutura de rede que está sendo utilizada como cenário pode ser visualizado na Figura 2.

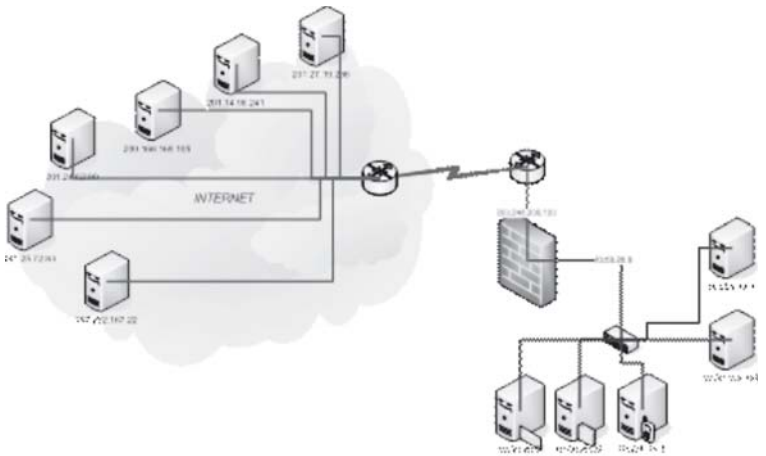

Figura 2 - Ambiente de testes

Com base no foco do problema, serão utilizados ainda os seguintes recursos disponíveis no sistema operacional: *tcpdump* para a captura dos pacotes de dados, *nmap* será utilizado para a identificação de portas de comunicação do *firewall*. Pesquisa-se ainda, a possibilidade de utilizar os registros de *log* da rede. Na Figura 3 apresenta-se a arquitetura proposta para o funcionamento do agente. A rede neural será utilizada para identificar a possibilidade de ataques ou não pelos dados analisados. Caso um ataque seja identificado, as ações para impedir o ataque deverá ser ativada.

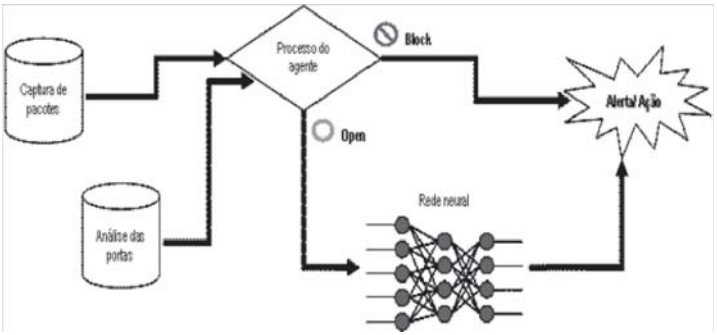

Figura 3 - Arquitetura do agente

RESULTADOS

Entre os resultados obtidos destaca-se a modelagem da rede neural que será utilizada pelo agente (Figura 4). A partir da definição do problema e da técnica de IA que deverá ser empregada, é necessário definir as informações de entrada e as informações de saída da rede neural. Quando as entradas e as saídas não são bem definidas pelo problema, deve-se utilizar uma rede recorrente, que é capaz de adaptar-se, ou seja, estabilizar os pesos sinápticos da rede de forma a produzir uma saída adequada conforme a sua topologia.

A arquitetura do agente apresentada anteriormente, concentra-se na análise de portas de comunicação e conteúdo do pacote referente às informações as quais necessita, tais como endereços IP, portas de origem e destino, freqüência de tentativas, erros, dentre outros; em sua MIB *(Management Information Base)* processa as informações coletadas gerando uma saída de alerta ou ação.

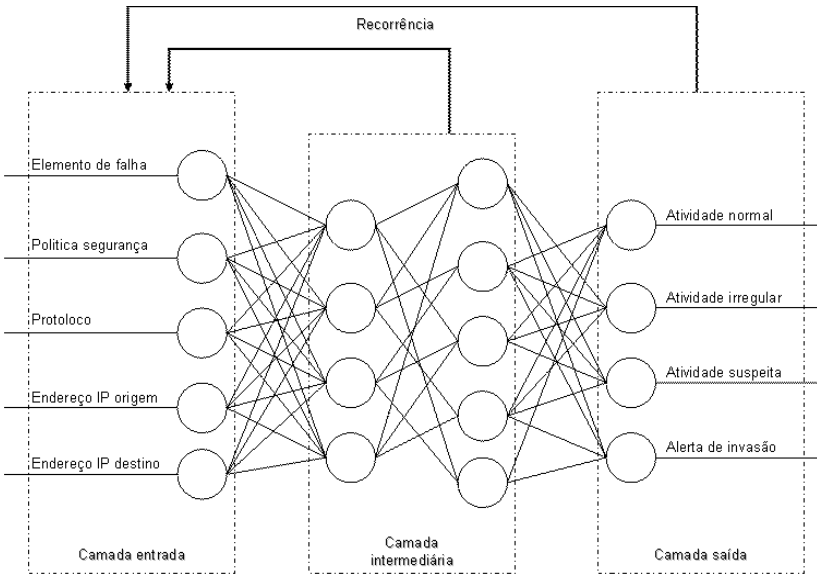

Figura 4 - Modelagem da rede neural recorrente

A rede neural será utilizada para definir uma classificação a partir dos dados de entrada, identificando as ações que serão sugeridas como saída do agente. Conforme os dados fornecidos na entrada o sistema deverá apontar uma saída. A rede será treinada com exemplos da rede e deverá convergir conforme as saídas: atividade normal, atividade irregular, atividade suspeita e alerta de invasão.

A tarefa do administrador é verificar o registro de log para estabelecer critérios de segurança, que foram utilizados para a identificação dos níveis de classificação das atividades. Estes dados alimentarão a rede neural recorrente no primeiro instante, afim de, se obter a saída coerente com a análise do administrador.

Observando os registros é possível perceber que há diversos acessos efetivados e outros tantos negados devido a políticas de segurança previamente estabelecidas, no primeiro instante busca-se alguma anomalia como freqüentes acessos em um determinado espaço de tempo, em seguida é analisado a efetividade ou não do acesso, assim como o endereço IP de origem (Figura 5).

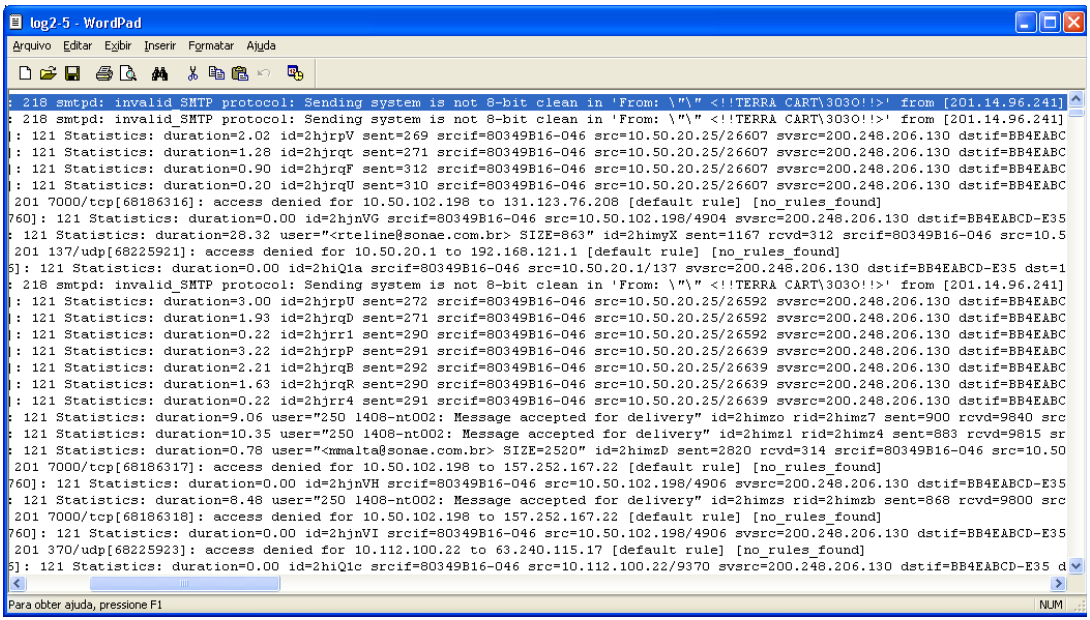

Figura 5 - Registro de acessos

Acessos constantes de um mesmo endereço de origem em um curto espaço de tempo sugerem um alerta de invasão. Endereços IPs de origem que pertençam a rede analisada em questão, sugerem uma falha de configuração, o que é classificado como atividade irregular; e constantemente repetida, uma atividade suspeita.

A classificação das atividades deverá analisar as informações do registros de acesso como mensagens de erro, tipo de protocolo, bem como as portas utilizadas para tal tráfego, endereços IP de origem e a freqüência de tentativas.

CONSIDERAÇÕES FINAIS

Problemas ocasionados por falhas de segurança e gerência de redes deficiente, podem causar graves danos às organizações. O investimento em pesquisa na área de gerência de redes é necessário para garantir um serviço estável e eficiente, evitando a queda do sistema ou a degradação do serviço. A área de informática é uma necessidade crescente em todos os ramos profissionais; dispositivos de computação devem contribuir para o crescimento profissional e pessoal e não tornar-se um problema com sua utilização. Pesquisadores e profissionais da área de IT devem estar focados às necessidades de seus usuários e de suas próprias, oferecendo sistemas es-

táveis, ágeis e seguros, utilizando-se de recursos existentes e novas propostas de soluções.

O presente trabalho apresenta um impacto significativo, tanto a nível social quanto técnico-científico. O primeiro está relacionado à mudança de comportamento dos profissionais que trabalham na área de gerência de redes. Ao invés de investir tempo na busca de soluções para os problemas, poderão se organizar e melhor aproveitar seu tempo de serviço na realização de ações preventivas evitando o mau funcionamento da rede de computadores. Já o aperfeiçoamento dos agentes baseados em SNMP, trará grande contribuição científica na medida em que as tecnologias pesquisadas poderão oferecer respostas mais inteligentes para os profissionais que precisam administrar redes de computadores complexas.

REFERÊNCIAS BIBLIOGRÁFICAS

ABOELELA, E.; DOULIGERIS, C. Fuzzy temporal reasoning model for event correlation in network management. In: IEEE CONFERENCE ON LOCAL COMPUTER NETWORKS, 24., 1999, Lowell. **Proceedings…** Lowell, Mass.: IEEE Press, 1999. p.150-160.

BARRETO, J.M. **Inteligência artificial no limiar do século XXI:** abordagem híbrida simbólica, conexionista e evolucionária. 3. ed. Florianópolis, 2001.

CACTI – The complete rrdtool-based graphing solution. 2006. Disponível em: <http://www.cacti.net/>. Acesso em: 25 fev. 2006.

FAN, Z.; MARS, P. Access flow control scheme for ATM networks using neural-network-based traffic prediction. **IEEE Proceedings Communications**, v.144, n.5, p. 295-300, oct. 1997.

FRANCESCHI, A. S. M. de. **Aplicação de técnicas de inteligência artificial no desenvolvimento de agentes para gerência de redes**. 2003. 144f. Tese (Doutorado em Engenharia Elétrica) - Universidade Federal de Santa Catarina, Florianópolis, 2003.

FRANKLIN, S. Autonomous agents as embodied AI. **Cybernetics and Systems**, v.28, n.6, p.499-520, 1997.

HAYES, C. C. Agents in a Nutshell. **IEEE Transactions on Knowledge and Data Engineering**, v.11, n.1, jan./feb.1999.

KUROSE, James. **Redes de computadores e a Internet:** uma nova abordagem. 1.ed. São Paulo: Addison Wesley, 2003. 515p.

LESSER, V.R. "Cooperative Multiagent Systems: A personal view of the state of the art". **IEEE Transactions on Knowledge and Data Engineering**, vol.11, n.1, p.133-142, jan./feb.1999.

LOOK at Lan – Network Monitoring and Management Solutions. 2006. Disponível em: <http://www.lookatlan.com>. Acesso em: 25 fev. 2006.

NAGIOS. 2006. Disponível em: <http://www.nagios.org>. Acesso em: 25 fev. 2006.

ROISENBERG, M. **Emergência da inteligência em agentes autônomos através de modelos inspirados na natureza**. 1998. Tese (Doutorado em Engenharia Elétrica) - Universidade Federal de Santa Catarina, Florianópolis, 1998.