Anais do IX Congresso Brasileiro de Redes Neurais / Inteligência Computacional (IX CBRN)
Ouro Preto 25-28 de Outubro de 2009

UM SISTEMA IMUNOINSPIRADO BASEADO NO ALGORITMO DAS CÉLULAS DENDRÍTICAS PARA A DETECÇÃO DE INTRUSÃO EM REDES DE COMPUTADORES

Guilherme Costa Silva*, Luciano de Errico†, Walmir M. Caminhas*

*LABICOMP - Laboratório de Inteligência Computacional, Departamento de Engenharia Elétrica
Universidade Federal de Minas Gerais, 31270-901
Belo Horizonte, Minas Gerais, Brasil

†LabCOM - Laboratório de Redes de Comunicação, Departamento de Engenharia Eletrônica
Universidade Federal de Minas Gerais, 31270-901
Belo Horizonte, Minas Gerais, Brasil

Emails: guicosta@ufmg.br, l.errico@cpdee.ufmg.br, caminhas@cpdee.ufmg.br

Abstract— This paper discusses the use of the Dendritic Cell algorithm (DCA) for anomaly detection, using scenarios where ping scan activity indicates the start of an intrusion. The text details how DCA's performance can be analyzed, which metrics to use, how the anomaly coefficient behaves during the process, and the conditions that favor better performance and reduce false alarms.

Keywords— Dendritic Cell Algorithm, Artificial Immune Systems, Anomaly Detection, ping scan.

Resumo— Este artigo discute o uso do Algoritmo das Células Dendríticas (DCA) para a detecção de anomalias, usando cenários onde atividades de ping scan indicam o início de uma intrusão. O texto detalha como o desempenho do DCA pode ser analisado, as métricas usadas, como o coeficiente de anomalia se comporta durante o processo e as condições que favorecem melhor desempenho e reduzem os alarmes falsos.

Palavras-chave— Algoritmo das Células Dendríticas, Sistemas Imunes Artificiais, Detecção de Anomalias, ping scan.

1 Introdução

Os sistemas de detecção de intrusos (IDS - Intrusion Detection Systems) são aplicativos que monitoram sistemas e redes de computadores para detectar comportamentos maliciosos, tais como acesso não autorizado, exploração de vulnerabilidades ou obtenção de privilégios. A detecção pode seguir duas abordagens: assinaturas, buscando padrões de intrusão existentes; ou anomalias, buscando eventos fora dos perfis de comportamento normal. Mais informações sobre IDS podem ser encontradas em Biermann et al. (2001).

Os IDS baseados em detecção de anomalias tendem a produzir grandes quantidades de alarmes falsos, pois não consideram as condições do ambiente de aplicação ou a possível falta de contexto. Muitos estudos têm sido desenvolvidos no sentido de resolver o problema e heurísticas de inteligência computacional, como as consideradas em Engelbrecht (2002), têm sido usadas para melhorar precisão e eficácia na classificação de comportamentos.

Os Sistemas Imunoinspirados são um conjunto de heurísticas baseadas nos mecanismos do sistema imune humano, oferecendo robustez, tolerância e resposta contra anomalias, características essenciais para um IDS. De acordo com Timmis et al. (2008), as abordagens imunoinspiradas podem servir como investigações interdisciplinares, que incorporam modelos matemáticos e computacionais de teorias sobre o sistema imune, produzindo resultados válidos tanto para aplicações em engenharia quanto estudos sobre imunologia.

Uma teoria correlata é o Modelo do Perigo (Matzinger, 1994), que adota o perigo como causa das respostas imunes, servindo de inspiração para uma nova geração de sistemas imunoinspirados (Aickelin and Cayzer, 2002) e trazendo abordagens e representações mais adequadas para os problemas, segundo Aickelin et al. (2003). Uma dessas novas abordagens é o Algoritmo das Células Dendríticas (DCA - *Dendritic Cell Algorithm*) (Greensmith et al., 2005). Na biologia, as células dendríticas são agentes naturais que detectam anomalias no corpo humano, constantemente buscando por potenciais fontes de danos ao organismo. Um estudo detalhado do algoritmo é feito em Greensmith (2007), onde se verifica a sua sensibilidade e o seu comportamento quanto à correlação de sinais.

Este trabalho utiliza o DCA na detecção do ping scan, método usado por hackers para obter informações sobre a rede, antes de realizar um ataque. O artigo analisa o progresso do coeficiente MCAV (mature context antigen value), que indica o grau de anomalia de um evento, e sua relação com a presença real de anomalia no sistema. Outras contribuições do artigo são a utilização de novas bases de dados para a detecção de perigo e a introdução de melhorias na analise de detecção do

ping scan, para reduzir os alarmes falsos.

2 Uso do DCA na detecção do ping scan

O ping scan é uma ferramenta usada por administradores de redes para buscar endereços IP ativos, mas que também é utilizada por invasores, para levantar endereços e portas disponíveis. Neste trabalho o DCA é utilizado para a detecção do ping scan, com base na metodologia descrita em Greensmith (2007) e nas seguintes ferramentas:

- `nmap`: utilitário mais comum de ping scan (processo a ser detectado pelo DCA);
- `scp`: utilitário para transferência segura de arquivos (processo normal do usuário, a ser classificado como não intrusão).

Para aplicar o DCA nestas condições, são utilizados o programa *strace* e a pasta */proc/net* do sistema operacional Linux[1] e timestamps[2] geradas pelo sistema.

O `strace` registra as chamadas de sistemas executadas através do `ssh`, protocolo de conexão remota e segura entre dois computadores em rede, usado neste contexto para simular uma intrusão. O registro do `strace` é usado pelo DCA como o vetor de antígenos A.

Os sinais são coletados através de scripts que utilizam dados de arquivos da pasta `/proc/net`, formatados como um arquivo de registro para ser usado pelo DCA como o vetor de sinais de entrada S e são mapeados a seguir, com seus respectivos significados biológicos e computacionais:

1. PAMP - S_{i0}: padrão molecular associado ao patógeno, indicador inerente de anomalia que corresponde nesta abordagem aos pacotes ICMP "Destino inacessível"recebidos por segundo, gerados para informar endereços não encontrados na rede;
2. Sinais Necróticos - S_{i1}: evidenciam uma possível situação de perigo e que, aqui, corresponde à taxa de pacotes enviados por segundo;
3. Sinais Apoptóticos - S_{i2}: indicam uma situação normal ou segura e corresponde ao inverso da variação da taxa de pacotes enviados;
4. Inflamação - S_{i3}: tipo de sinal que amplifica os efeitos dos demais sinais, não utilizado nesta abordagem, pois nào é significante, conforme Greensmith (2007).

onde i é o índice do conjunto de sinais dentro da categoria.

Estes sinais de entrada, após processados, formam os seguintes sinais de saída:

1. moléculas coestimulatórias (O_0): representam a migração da célula (do estado inicial para o de semi-madura ou madura), ocorrida quando a variável atinge um limiar pré-determinado;
2. citocinas semi-maduras (O_1): sinal que representa a supressão da resposta imune (possui valor de contexto 0);
3. citocinas maduras (O_2): sinal que representa a ativação da resposta imune (possui valor de contexto 1).

Estas saídas são calculadas pela Equação 1, onde j corresponde à categoria do sinal de entrada, p corresponde à categoria do sinal de saída e W corresponde à matriz de pesos, que representa as relações entre as entradas e as saídas, resumidas a seguir:

- o sinal PAMP possui maior influência na migração e produção de citocinas maduras em relação aos sinais necróticos;
- a produção de citocinas semi-maduras está totalmente relacionada com os sinais apoptóticos;
- os sinais apoptóticos podem também suprimir a produção das citocinas maduras, influenciando negativamente em seu sinal de saída correspondente.

$$O_p = \frac{\sum_i \sum_{j \neq 3} (W_{ijp} * S_{ij})}{\sum_i \sum_{j \neq 3} W_{ijp}} * (1 + S_{i3}), p \in \{0, 1, 2\} \quad (1)$$

Com o processamento e cálculo dos sinais de saída e a posterior migração da célula, as concentrações de O_1 e O_2 são comparadas, determinando o contexto da célula. Uma vez atribuídos os valores de contexto, para cada antígeno é definido um coeficiente MCAV, calculado pela Equação 2.

$$MCAV = \frac{mDC(antigeno)}{smDC(antigeno) + mDC(antigeno)} \quad (2)$$

[1] Os experimentos foram realizados em um computador com processador Intel Core2Duo de 4GB de RAM rodando Linux Ubuntu 8.04 e Matlab R2008a. Foram usados o `Nmap` versão 4,53, para realizar o *Ping Scan*, e o OpenSSH version 4.7, para fazer login remotos e prover `scp`.

[2] Apesar da coleta dos dados ser feita em tempo real, foi necessário registrar também o tempo da aquisição dos dados, para a correlação adequada entre antígenos e sinais.

O seguinte pseudocódigo descreve o funcionamento do DCA:

```
inicia DCA(Sinais de entrada:S,Antígenos:A)
    Inicializa_parâmetros(I, J, K, M, N, P, Q)
    Criar População DC, com M indivíduos
    Enquanto(ainda há dados para analisar)
        Atualiza(A,S)
        Para cada indivíduo m de M
            Coleta Q antígenos de A até N
        Para cada saída p de O
        Para cada sinal i e categoria j de S
            Calcula saída O_p com S_ij e W_ijp(Eq. 1)
            atualiza O_p
        Se O_1 > Limiar Tm de DC
            Se O_2> O_3
                  contexto de DC_m ← 0
            Senão
                  contexto de DC_m ← 1
        migra DC_m e a substitui
    Calcula MCAV para cada tipo de antígeno(Eq. 2)
Fim
```

2.1 Alterações no algoritmo

Para possibilitar a análise da detecção, algumas alterações foram realizadas no DCA, tais como: cálculo do MCAV após a migração e sua atualização em cada ciclo (o que torna possível uma análise iterativa das ocorrências de anomalia no sistema, uma vez que o DCA original apenas classificava o processo); armazenamento temporal dos valores do MCAV; armazenamento do início da execução dos processos e de quando foi acusada a anomalia; implementação de valores lógicos que indicam existência da anomalia e do sinalizador de alterações no MCAV, cuja estagnação indica o encerramento do processo; e os cálculos do tempo percentual da anomalia durante sua execução (Equação 3) e da taxa de anomalia durante o teste (Equação 4).

$$perc.anomalia = \frac{t_{anomalia}}{t_{total}} \quad (3)$$

$$tx.anomalia = \frac{t_{processo_f} - t_{anomalia_i}}{t_{processo_f} - t_{processo_i}} \quad (4)$$

2.2 Detalhamento dos testes

A descrição dos testes é mostrada na Tabela 1; os parâmetros usados, na Tabela 2; os pesos, na Tabela 3; e a normalização realizada, importante para minimizar e otimizar a diversidade dos dados, na Tabela 4.

[3]usado apenas na discussão da subseção 3.2

Tabela 1: Quantidade de IPs rastreados pelo *nmap* e tamanho do arquivo transmitido pelo *scp*.

	IPs rastreados	Tamanho (MB)
Cenário de Perigo - só *nmap*		
Teste 1	17200	-
Teste 2	29600	-
Teste 3	15240	-
Cenário Normal - só *scp*		
Teste 1	-	2 X 4MB
Teste 2	-	30
Teste 3	-	300
Cenário Misto- *nmap* e *scp*		
Teste 1	3795	4
Teste 2	8855	50
Teste 3	16445	100
Teste 4 [3]	16445	300

Tabela 2: Parâmetros-padrão do DCA para os testes com ping scan.

Nome	Símbolo	Valor
No. de sinais por categoria	I	1
No. de categorias dos sinais	J	3
No. de antígenos processados	K	500
No. de células	M	100
Máximo de antígenos por célula	N	50
Máximo de sinais de saída	P	3
No. de receptores de antígenos	Q	1
Limiar mediano de migração	t_m	15

3 Resultados

3.1 Resultados obtidos

A Tabela 5 mostra os valores do índice MCAV nos processos e a Tabela 6 traz a análise temporal. Nos cenários de perigo e misto, o processo de interesse é o `nmap`, que caracteriza uma intrusão legítima; No cenário normal, o processo é o `scp`, que influencia significativamente no comportamento dos sinais. Para os testes realizados, foram realizadas 20 execuções de cada teste, com limiar de anomalia de 50% e o desvio padrão observado foi relativamente baixo, em média 0.2×10^{-3}.

Os dados das referidas tabelas trazem informações relevantes sobre o comportamento da detecção através do DCA: o MCAV médio indica qual foi a tendência do processo analisado, o MCAV máximo indica o quão anômalo o processo foi considerado e o MCAV final é a classificação final adotada em Greensmith et al. (2005).

Apesar da resposta satisfatória ao classificar o

[4]Este sinal é menor devido ao seu efeito altamente supressivo no cálculo do sinal de saída

[5]Neste teste, foram analisados dois processos `scp`.

Tabela 3: Pesos sugeridos, baseados em critérios predefinidos.

w_{ijp}	$j=0$	$j=1$	$j=2$
$p=0$	2	1	3
$p=1$	0	0	3
$p=2$	2	1	-3

Tabela 4: Normalização dos sinais.

S_{ij}	$j=0$	$j=1$	$j=2$
Faixa de valores	[0 10]	[10 100]	[10 1000]
Mínimo normalizado	0	0	0
Máximo normalizado	100	100	10^{4}
Cálculo	$s_{0,0} = s * 5$	-	-

nmap corretamente, o DCA também classificou o *scp* como anomalia em testes do cenário normal. Este fato será investigado na próxima seção, usando o teste 4 do cenário misto.

3.2 Redução dos alarmes falsos

Para descobrir a razão do alarme falso ocorrido em alguns testes do cenário normal, foi introduzido o teste 4 do cenário misto que, como os demais testes, apresentou a detecção normal do processo que realiza o *ping scan*. Porém, o upload do arquivo também foi acusado. Esse resultado é apresentado na Figura. 1.

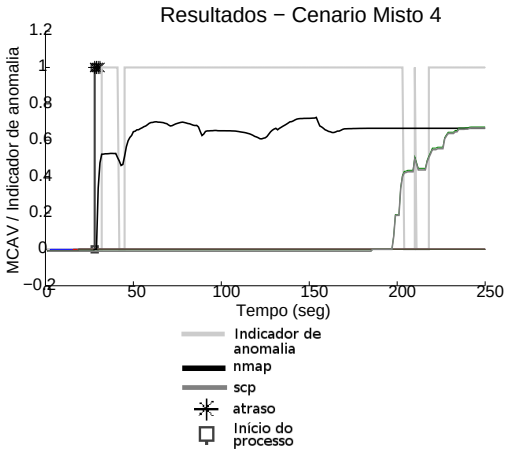

Figura 1: Teste 4 do cenário misto: apesar da detecção do `nmap`, o `scp` também foi acusado.

A taxa de pacotes enviados do `scp` é muito maior do que a taxa do `nmap`. Isso significa que S_{i1} está muito alto nas situações de segurança e nem mesmo S_{i2} consegue suprir a ação do upload.

Supondo uma grande diferença entre a taxa de uploads e ping scans, a normalização de S_{i1} foi alterada conforme indicado na Tabela 7 e o teste foi repetido. A avaliação do MCAV é apresentada na Figura 2.

Tabela 5: Valores do índice MCAV para os diferentes cenários de testes realizados, sendo P = Cenário de perigo, N = Cenário normal e M = Cenário misto.

		nmap			scp		
Teste		max	med	final	max	med	final
P	1	0,80	0,71	0,79	-	-	-
	2	0,78	0,70	0,73	-	-	-
	3	0,77	0,66	0,74	-	-	-
N	1^{5}	-	-	-	0.01	0.003	0.01
					0,01	0,009	0,01
	2	-	-	-	0,59	0,37	0,59
	3	-	-	-	0,90	0,75	0,89
M	1	0,68	0,66	0,67	0	0	0
	2	0,74	0,63	0,66	0,01	0,005	0,001
	3	0,85	0,74	0,70	0,38	0,20	0,38

Tabela 6: Análise temporal da execução do DCA implementado, onde: T_N = duração do teste, T_{an} = instante onde a anomalia é sinalizada, T_{ex} = instante onde o processo foi iniciado, VP = verdadeiro positivo, VN = verdadeiro negativo e FP = falso positivo.

Teste		T_N	T_{an}	T_{ex}	Eq. 4	Eq. 3	Resposta
P	1	77s	16	13	95%	77%	VP
	2	171s	18	15	98%	87%	VP
	3	138s	32	12	84%	75%	VP
N	1	63s	-	-	-	-	VN
	2	53s	19	9	64%	33%	FP
	3	109s	30	20	88%	70%	FP
M	1	128s	18	16	97%	73%	VP
	2	178s	43	17	83%	68%	VP
	3	450s	41	39	96%	76%	VP

Tabela 7: Normalização alternativa.

	Sinal Necrótico (S_{i1})
faixa inferior	[10 100]
faixa superior	[1900 2000]
mínimo normalizado	0 (de 0 até 10 e de 2000 a $+\infty$)
máximo normalizado	100 (de 100 até 1900)

A Tabela 8 traz o comparativo dos testes feitos no cenário misto 4, antes e depois da normalização. Com esses resultados, os dois últimos testes do cenário normal foram refeitos, resultando na classificação correta do `scp`, conforme mostra a Tabela 9.

Para avaliar os resultados, foram realizados alguns testes usando limiares de anomalia entre 0% e 100%. Com os resultados obtidos, uma curva

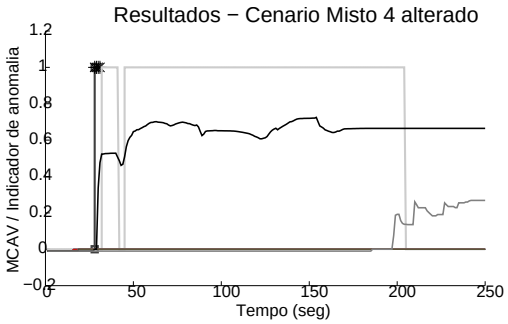

Figura 2: Teste 4 do cenário misto: alterada a normalização, o `scp` foi classificado corretamente como um processo normal.

Tabela 8: Comparativo dos testes feitos no cenário misto 4.

Normalização	Antes		Depois	
Processo	`Scp`	`nmap`	`Scp`	`nmap`
MCAV Máximo	0,67	0,72	0,27	0,72
MCAV Médio	0,44	0,64	0,18	0,64
MCAV Final	0,67	0,66	0,27	0,66
Anomalia começa	216	30	-	30
Equação 3	51%	68%	-	68%
Equação 4	83%		68%	

ROC (Fawcett, 2006) foi desenvolvida com os valores de sensibilidade (verdadeiros positivos) e especificidade (verdadeiros negativos) obtidos através dos cálculos de limiares, como mostra a Figura 3.

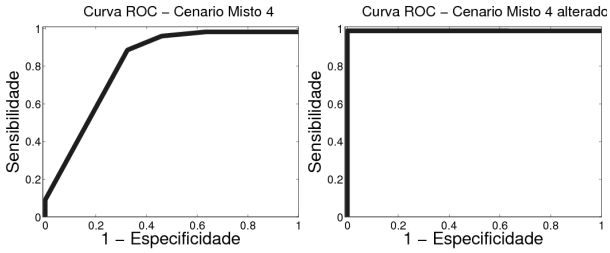

Figura 3: Curvas ROC para os testes feitos no cenário misto 4, mostrando que a normalização melhorou o desempenho do algoritmo.

A mudança na normalização do DCA proporcionou melhorias significativas na precisão da detecção, alcançando taxas de erro quase nulas com o mapeamento adequado dos sinais.

4 Conclusões

Foi implementada uma métrica de detecção no DCA que permitiu a visualização de tendências dos antígenos em relação aos sinais e a análise temporal da ocorrência de intrusão. Foi possível perceber também que o *ping scan* possui uma tendência crescente em relação a seu MCAV que, porém, não é estável, e que seu valor final assume valores menores do que o valor máximo. Já a transferência de arquivo, que possui crescimento gradual, tendência a permanecer constante e valor final normalmente igual ao valor máximo, teve esse comportamento alterado após a introdução da normalização.

Tabela 9: Comparativo dos testes feitos nos cenários normais.

Cenário	Normal 2		Normal 3	
`scp`	Ant	Dep	Ant	Dep
MCAV Máximo	0,59	0	0,90	0,43
MCAV Médio	0,37	0	0,75	0,23
MCAV Final	0,59	0	0,89	0,37
Anomalia começa	36	-	34	-
Equação 3	64%	-	88%	-
Equação 4	33%	0%	70%	0%

Em termos gerais, o algoritmo apresentou resultados semelhantes aos da literatura na detecção do *ping scan*. Na transferência do arquivo houve grandes diferenças, por causa dos testes com transferências maiores e tempo mais longo, produzindo respostas indesejadas. Isto foi resolvido com a normalização, gerando resultados mais condizentes com o esperado. O DCA detectou os processos com um atraso relativamente pequeno.

Essas contribuições apresentam um novo ponto de vista do algoritmo em relação aos trabalhos na literatura e a confirmação das vantagens do DCA e do uso de sinais e antígenos em correlação para detecção de anomalias.

Agradecimentos

O projeto teve o apoio do CNPq e da FAPEMIG. Agradecemos ao Thiago Guzella, por discussões relacionadas ao DCA, e à Dra. Julie Greensmith e ao Dr. Uwe Aickelin (Universidade de Nottingham, UK), pelos dados gentilmente cedidos e cujos experimentos conduziram à validação das implementações.

Referências

Aickelin, U. and Cayzer, S. (2002). The danger theory and its application to artificial immune systems, In: *Proceedings of the 1st International Conference on Artificial Immune Systems (ICARIS-2002)*, pp. 141–148.

Aickelin, U., Bentley, P., Cayzer, S., Kim, J. and Mcleod, J. (2003). Danger theory: The link between ais and ids, *In Proc. of the Second Internation Conference on Artificial Immune Systems (ICARIS-2003)*, pp. 147–155.

Biermann, E., Cloete, E. and Venter, L. M. (2001). A comparison of intrusion detection systems, *Computers and Security* **20**(8): 676–683.

Engelbrecht, A. (2002). *Computational Intelligence: An Introduction*, Halsted Press, New York, NY, USA.

Fawcett, T. (2006). An introduction to roc analysis, *Pattern Recogn. Lett.* **27**(8): 861–874.

Greensmith, J. (2007). *The dendritic cell algorithm*, PhD thesis, University of Nottingham.

Greensmith, J., Aickelin, U. and Cayzer, S. (2005). Introducing dendritic cells as a novel immune-inspired algorithm for anomaly detection, *Artificial Immune Systems*, LNCS, Springer, pp. 153–167.

Matzinger, P. (1994). Tolerance, danger and the extended family, *Annual Reviews in Immunology* pp. 991–1045.

Timmis, J., Andrews, P., Owens, N. and Clark, E. (2008). An interdisciplinary perspective on artificial immune systems, *Evolutionary Intelligence* **1**(1): 5–26.