# USER PROFILING BASEADO EM REDES NEURAIS

*Paulo Henrique Pisani[1], Silvio do Lago Pereira[2]*
*[1,2]Faculdade de Tecnologia de São Paulo – FATEC-SP*
*paulohpisani@yahoo.com.br, slago@ime.usp.br*

## *1. Introdução*

Na era atual, as pessoas possuem um ativo chamado *identidade* [1]. Identidades são usadas para autenticação em diversos contextos envolvendo sistemas informatizados, desde acesso a sistemas de suporte ao trabalho até a contas bancárias e comunidades de relacionamento.

A autenticação de usuários pode ser feita por meio de diversos métodos, alguns mais simples, baseados em senha, outros mais sofisticados, baseados em biometria. Contudo, de modo geral, após a autenticação inicial, um intruso está livre para acessar e utilizar o sistema assim como o usuário legítimo o faria [2]. Conforme pesquisas realizadas [3], a cada ano, mais de 10 milhões de pessoas são afetadas pela fraude de identidade.

## *2. User Profiling*

A fraude de identidade pode ser evitada por meio da aplicação de mecanismos que permitem uma constante autenticação do usuário, usando para isso características mais difíceis de serem roubadas (ex. ritmo de digitação do usuário). A partir do monitoramento do usuário, e da observação de tais características, padrões de comportamento do usuário são definidos. O processo de obtenção destes padrões para geração do *perfil* (*profile*) é conhecido como *perfilamento de usuário* (*user profiling*) [4].

A partir do perfil representando o comportamento normal do usuário, gerado no *período de treinamento* [5], um sistema de detecção de intrusão (*Intrusion Detection System* – IDS) pode detectar anomalias de comportamento. Para tanto, o IDS efetua uma comparação entre o perfil e os eventos observados, indicando aquilo que desvia do perfil como uma possível fraude.

A maior vantagem dos IDSs baseados em detecção de anomalias é fato de eles serem bastante efetivos na detecção de ameaças ainda não conhecidas [5].

## *3. Redes Neurais Artificiais*

Uma ferramenta bastante poderosa para a tarefa de reconhecimento de padrões são as *redes neurais artificiais* (RNAs). Redes neurais artificiais são comuns em aplicações de Inteligência Artificial (IA), cujo objetivo é obter sistemas computacionais que simulem comportamento inteligente, tais como aprendizado, percepção, raciocínio e adaptação [6].

RNAs, que têm como motivação as redes neurais biológicas, possuem a capacidade de operar com dados incompletos e com ruídos, oferecendo assim grande robustez no reconhecimento de padrões [7].

## *4. Aplicação*

Utilizando redes neurais artificiais, desenvolvemos um sistema simples para detecção de intrusões (IDS) baseado em *user profiling* [8]. As RNAs foram usadas para identificar desvios no comportamento dos usuários.

Para a definição do comportamento do usuário, bem como para a detecção de desvios, escolhemos características relativas ao seu ritmo de digitação. A análise da dinâmica de digitação, nesse contexto, é considerada um tipo de *biometria comportamental* [9].

Uma das aplicações da dinâmica de digitação é na verificação contínua de palavras-chave específicas [9]. Com base neste conceito, monitoramos uma lista de palavras consideradas sensíveis. Os testes iniciais foram realizados com 12 pessoas e obtivemos taxas de falsa aceitação de 22,5% e de falsa rejeição de 23,5% na média, sendo que para comandos com maior quantidade de caracteres as taxas foram de 10% e 15%.

## *5. Conclusões*

Com aumento significativo do uso de identidades, bem como das fraudes relacionadas ao mau uso destas, mecanismos de segurança adicionais ao uso de senhas passaram a ser fundamentais. Um destes mecanismos é o *user profiling* [5]. De fato, conforme observamos com os testes realizados, *user profiling* baseado em redes neurais artificiais é um mecanismo bastante efetivo e robusto para o reconhecimento de fraudes de identidade.

Porém, uma aplicação de *user profiling* mais completa necessita de um conjunto maior de características para atingir resultados mais precisos. Algumas dessas características poderiam ser a velocidade de movimentação do cursor, os tempos de clique do *mouse*, a quantidade de cliques, programas usados e consumo de tempo de CPU e memória. Tais características adicionais são objetos de estudo para trabalhos futuros.

## *6. Referências*

[1] *Taking control of your digital ID*, BBC News, 1 nov. 2006.

[2] D. J. Marchette, *Computer Intrusion Detection and Network Monitoring*, Springer, 2001.

[3] *Identity Theft Focus of Nat. Consumer Protection*, Week 2005. Federal Trade Commission, Fev. 2005.

[4] T. Goldring, *User Profiling for Intrusion Detection in Windows NT*, In Proc. of the 35th Symposium on the Interface, 2003.

[5] K. Scarfone & P. Mell, *Guide to Intrusion Detection and Prevention Systems (IDPS)*, National Institute of Standards and Technology (NIST), 2007.

[6] S. Russell & P. Norvig, *Artificial Intelligence – A Modern Approach*, 2nd ed., Prentice-Hall, 2004.

[7] A. P. Braga et. al. *Redes Neurais Artificiais*. LTC, 2007.

[8] P. H. Pisani, *User Profiling com Redes Neurais*, http://br.geocities.com/php_rna/default.htm, Acesso em: 11 ago. 2008.

[9] P. Elftmann, *Secure Alternatives to Password-based Authentication Mechanisms*, Lab. for Dependable Distributed Systems, RWTH Aachen Univ., 2006.