A Aplicação de uma Arquitetura de Máquinas de Comitê na Autenticação de Usuários através da Dinâmica de Digitação

Mauro Roisenberg, Sérgio Roberto de Lima e Silva Filho

[1]Departamento de Informática e Estatística – Universidade Federal de Santa Catarina (UFSC)
Caixa Postal 476 – 88.040-90 – Florianópolis – SC – Brasil

{mauro,sergio}@inf.ufsc.br

Abstract. *This paper describes the application of neural network committee machines to the authentication problem, particularly for the authentication of computer users through the classification of keystroke dynamics patterns. Keystroke dynamics is a biometric characteristic that encompass the features of users when typing the computer keyboard. The proposed methodology also provides continuous authentication. The results and the associated discussion describe a successful application of neural networks committee machines to a difficult, but important, real world task taken from the security area. In addition the techniques described are widely applicable to other complex pattern classification problems.*

Resumo. *Este artigo descreve a aplicação de máquinas de comitê de redes neurais aplicadas ao problema de autenticação, especificamente a autenticação de usuários de computadores através da classificação de padrões da dinâmica de digitação. Dinâmica de digitação é uma característica biométrica que engloba os atributos dos usuários quando estão utilizando o teclado do computador. A metodologia proposta provê também autenticação contínua. Os resultados obtidos e a discussão associada descreve uma aplicação promissora de máquinas de comitê de redes neurais a um problema real e importante oriundo da área de segurança. Além disso, as técnica aqui descritas podem ser largamente aplicadas a outros problemas complexos de classificação de padrões.*

1. Introdução

Um grande problema nos computadores e sistemas computacionais é a necessidade de um mecanismo de autenticação do usuário, que possibilite que apenas usuários válidos tenham acesso aos recursos computacionais críticos, sejam estes os próprios acessos ao computador, a um determinado sistema ou ainda a uma determinada informação. Para isto, devem existir mecanismos eficientes de autenticação para que pessoas não autorizadas não consigam se passar por pessoas válidas e acessar estes recursos restritos [5], [7], [12] e [15].

Hoje em dia as técnicas de autenticação se baseiam na utilização de um ou mais dos mecanismos de segurança: através de algo que somente o usuário sabe; através de algo que somente o usuário possui; e através de algo que o usuário é.

Fazendo parte desta última técnica de autenticação está a forma de autenticação conhecida como prova por biometria. Esta forma é baseada em uma característica fisiológica ou comportamental do indivíduo. Como exemplos de características fisiológicas

têm-se a impressão digital, leitura da íris, padrão das linhas da mão, etc., já as características comportamentais podem ser a dinâmica da digitação, reconhecimento de voz, entre outros. Esta forma de autenticação é uma das mais seguras já que não se pode roubar o que o indivíduo é.

Os principais mecanismos de autenticação são efetuados apenas na entrada do sistema e, portanto, o usuário não necessita se autenticar durante todo o tempo em que está utilizando um determinado recurso computacional, mas apenas no acesso inicial ao mesmo. Entretanto, muitos usuários não se preocupam em fechar a sessão ou bloquear o acesso ao computador quando se ausentam de seu local de trabalho, de modo que ele fica vulnerável a ataques de pessoas não autorizadas [7].

O objetivo deste artigo é descrever um método de autenticação seguro, barato e contínuo do usuário através da característica biométrica de dinâmica da digitação. O mecanismo utilizado para a autenticação envolve o uso de Redes Neurais Artificiais (RNAs) através da sua capacidade de classificação de padrões [13]. Procura-se mostrar que esta é uma tarefa bastante complexa para ser executada por apenas uma RNA necessitando de uma estratégia mais elaborada para classificação de padrões que envolve o uso de um conjunto de RNAs formando um comitê de especialistas, chamado Máquinas de Comitê.

Entre os fatores que levam à utilização da dinâmica de digitação como o método de autenticação pode-se destacar: o baixo custo de implementação, pois é necessário apenas um teclado e um software para coleta, treinamento e verificação do padrão do usuário; e o fato de ser "não intrusivo", ou seja, o usuário não se sente constrangido de ter de provar sua identidade, sendo que o processo de autenticação pode ser realizado de forma contínua e transparente, enquanto o usuário utiliza o teclado para acesso usual as suas tarefas no sistema.

Para análise do desempenho do sistema de autenticação foi utilizada uma matriz de confusão representando as Taxas de Falsa Aceitação (FAR - False Acceptance Rate), ou seja, as vezes que um usuário não-autorizado é aceito pelo sistema como sendo um usuário legítimo, e de Falsa Rejeição (FRR - False Rejection Rate), ou seja, as vezes que a um usuário legítimo não é permitido acessar o sistema [15].

O restante deste artigo está organizado da seguinte maneira: a segunda seção apresenta uma revisão de alguns trabalhos já desenvolvidos na área. A seção 3 apresenta a fundamentação teórica de Máquinas de Comitê de Redes Neurais. A metodologia proposta é apresentada na seção 4. Os experimentos realizados e os resultados obtidos são discutidos na seção 5. Finalmente, a seção 6 apresenta as conclusões.

2. Revisão de Trabalhos Correlatos

De acordo com [11] *"os mesmos fatores neurofisiológicos que tornam a assinatura manuscrita única também são exibidos pelo padrão de digitação dos usuários. Quando uma pessoa digita em um teclado, esta deixa uma assinatura digital na forma de latências entre as teclas digitadas"*. O método de identificação através da dinâmica da digitação está baseado na hipótese de que cada indivíduo quando está digitando textos em um teclado segue padrões de digitação diferentes, ou seja, quando um indivíduo esta digitando textos o ritmo da digitação varia de pessoa para pessoa sendo que este é único. Conseqüentemente, se o ritmo de cada indivíduo é único, é possível identificá-lo através desta característica.

No estudo da dinâmica da digitação de uma pessoa existem diversas características que podem ser adquiridas e mensuradas à medida que o usuário digita textos no teclado, como por exemplo [1], [10] e [16]:

- Latência entre digitações consecutivas - a latência pode ser medida de várias maneiras, como, por exemplo, através da latência entre duas teclas pressionadas, latência entre duas teclas soltas, latência entre o pressionar de uma tecla e soltar de outra, entre outros;
- Duração de tempo que a tecla é mantida pressionada;
- Velocidade total de digitação;
- Freqüência de erros;
- Correlação entre as teclas pressionadas, principalmente quando digita-se letras maiúsculas e acentuação;
- A pressão exercida sobre as teclas, entre outras.

As técnicas de identificação através da dinâmica da digitação encontradas na literatura podem ser classificadas quanto ao momento da autenticação. Na abordagem estática a autenticação do usuário é realizada geralmente no acesso do usuário ao sistema. Já na abordagem dinâmica o usuário é autenticado continuamente pelo sistema.

No século 19, observações dos operadores de telégrafo mostraram que cada operador tinha seu próprio padrão de digitar as mensagens e os operadores conseguiam reconhecer outros operadores apenas pela forma que eles digitavam as mensagens [10]. Recentemente, muitos outros estudos foram realizados sobre o tema da dinâmica da digitação.

É importante ressaltar que ao digitar letras de um texto através de um teclado, a latência de tempo de digitação entre teclas não é um número fixo, mas um conjunto de valores que seguem alguma distribuição de freqüência. Uma noção superficial do tipo de problema que se necessita resolver pode ser visto na Figura 1. É possível observar claramente que se está diante de um problema de classificação de padrões onde existe sobreposição de padrões e o estabelecimento de uma fronteira de decisão é extremamente complexo. Neste caso, a latência de tempo de digitação observada entre apenas duas teclas fixas pode não ser suficiente para determinar claramente a que classe, isto é, a que usuário pertence determinada latência. Além disso, as distribuições de freqüência para cada duas teclas diferentes seguem distribuições diferentes para cada usuário. Assim, é como se o problema de classificação de padrões se estendesse para uma dimensão n, onde n é o número de diferentes combinações de pares de teclas de um teclado.

Deste modo, as abordagens normalmente utilizadas para autenticação de usuários através da dinâmica da digitação, em sua maioria são técnicas estatísticas de classificação de padrões.

Em 1980 Gaines et. al. [8] publicou um trabalho utilizando a dinâmica da digitação para autenticação num grupo de sete secretárias utilizando métodos estatísticos (t-tests) para classificação e verificação do acesso. Gaines utilizou três textos, sendo o primeiro um texto em inglês, o segundo um conjunto de palavras randômicas e o terceiro um conjunto de frases randômicas. Todos os textos tinham tamanhos variando de 300 a 400 palavras e cada um dos textos foi coletado duas vezes sendo que a segunda coleta foi realizada 4 meses depois da primeira. A latência entre dígrafos foi utilizada como característica medida, mas apenas para os dígrafos que ocorreram mais que 10 vezes. O método de Gaines gerou uma FAR de 0% e FRR de 4%.

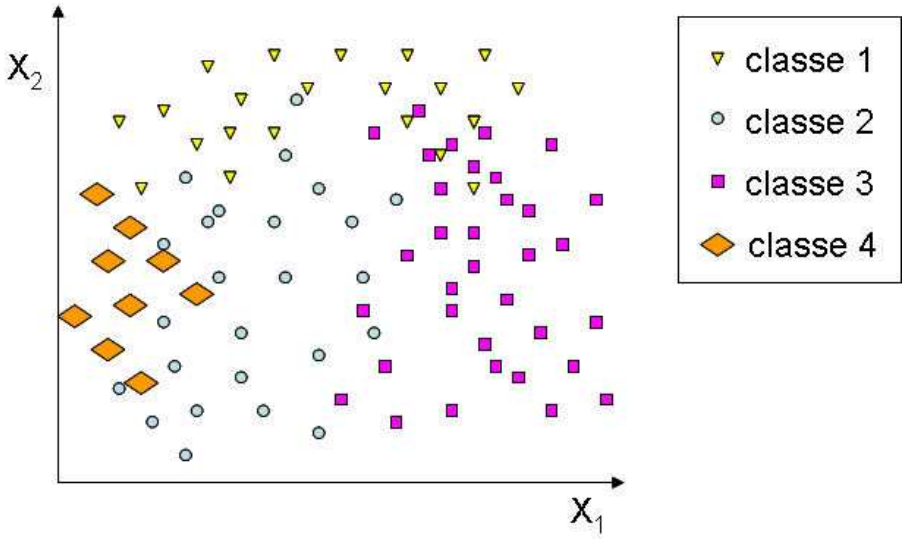

Figura 1. Um exemplo esquemático de vetores bi-dimensionais (X_1, X_2) pertencentes a 4 classes, a fim de ilustrar a dificuldade de se estabelecer uma superfície de separação entre as classes

Em 1990 Joyce e Gupta [11] desenvolveram um trabalho tratando da abordagem estática de autenticação através dos dados de login, senha, primeiro e último nome de 33 usuários, sendo coletados 8 vezes esta informação para treinamento e 5 para testar o modelo. A característica estudada foi o tempo entre teclas sendo o classificador estatístico. Este modelo resultou em uma FAR de 0,25% e FRR de 16,67%.

Existem também alguns trabalhos que abordam a autenticação de usuários através da dinâmica de digitação utilizando-se de Redes Neurais Artificiais [2], [3], [4], [5], [6], [11], [12], [14] e [15], entretanto, na grande maioria destes trabalhos, a abordagem é estática e os usuários são autenticados apenas no início da sessão.

3. Redes Neurais em Máquinas de Comitê

As RNAs podem ser vistas como um grande número de simples processadores interconectados formando um sistema computacional paralelo. Este paradigma da Inteligência Artificial (IA) procura usar alguns princípios organizacionais como aprendizado, generalização, adaptação, tolerância a falhas e representação distribuída numa rede de pesos onde os nodos são neurônios artificiais. Uma das principais características deste paradigma é a sua habilidade em aprender relações não lineares complexas de entrada e saída através de exemplos da casos conhecidos da relação.

Em tarefas de classificação de padrões, as RNAs são freqüentemente utilizadas para modelar fronteiras de decisão com bom desempenho de generalização através de um modelo com um grau de complexidade intermediário. Entretanto, no caso específico da dinâmica da digitação, o grau de liberdade das variáveis do problema é bastante elevado e a superposição de padrões pode ocorrer com freqüência quando se analisa poucos pares de teclas. Deste modo, a utilização de uma única RNA apresenta um desempenho insatisfatório dada a complexidade do problema da autenticação de usuários através da dinâmica de digitação.

Quando uma tarefa é complexa o melhor a fazer é subdividi-la em pequenas tarefas

simples e combinar as soluções destas tarefas para resolver o todo. Segundo [9] este é o princípio de "dividir e conquistar" muito utilizado na engenharia.

Haykin [9] afirma que: *"na aprendizagem supervisionada, a simplicidade computacional é alcançada distribuindo-se a tarefa de aprendizagem entre um número de especialistas, que, por sua vez, divide o espaço de entrada em um conjunto de subespaços"*. Esta combinação de especialistas constitui uma máquina de comitê. A idéia é que uma máquina de comitê supostamente produz resultados melhores que quando utilizado qualquer especialista individualmente, pois ela utiliza o conhecimento de vários especialistas para chegar a uma decisão.

No caso das RNAs organizadas como máquinas de comitê, existem uma série de classificações encontradas na literatura. No caso mais simples, chamado de *média de ensemble*, podemos imaginar um conjunto de redes treinadas sobre os mesmos conjuntos de treinamento. Como cada rede é inicializada com pesos diferentes, a convergência do aprendizado conduzirá a modelos, ou fronteiras de decisão diferentes para o mesmo problema, como se fossem visões diferentes de diferentes especialista frente a um mesmo problema. As saídas produzidas por cada um destes especialistas deverão então ser combinadas de alguma forma para produzir uma saída global y [9].

A Figura 2 exemplifica uma máquina de comitê baseada na média de ensemble:

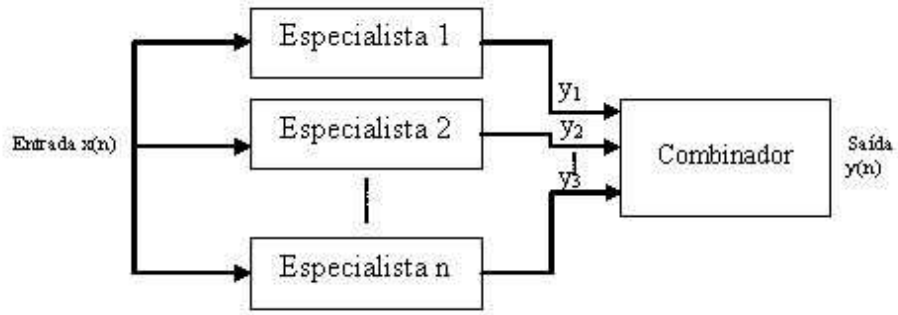

Figura 2. Máquina de Comitê baseada na Média de Ensemble

4. Arquitetura Proposta e Metodologia

4.1. Arquitetura Proposta

Na arquitetura proposta, um conjunto de redes neurais organizadas em máquina de comitê foi utilizada para autenticação de cada usuário, ou seja, para cada usuário cadastrado no sistema e que deseja se autenticar, existe um conjunto de redes neurais formando uma máquina de comitê que é capaz de reconhecer o padrão de digitação corresponde ao usuário em questão.

No nosso caso específico, cada usuário é representado por um conjunto de 13 RNAs, sendo então que a autenticação do usuário se dá pela combinação das classificações destas 13 RNAs que formam o comitê. A diferença em relação a uma arquitetura convencional de máquinas de comitê por média de ensemble é que o conjunto de treinamento das RNAs do comitê não é o mesmo, ou seja, cada RNA que compõe o comitê é treinada para classificar o usuário legítimo contra apenas um usuário não-legítimo, como pode ser visto na Figura 3.

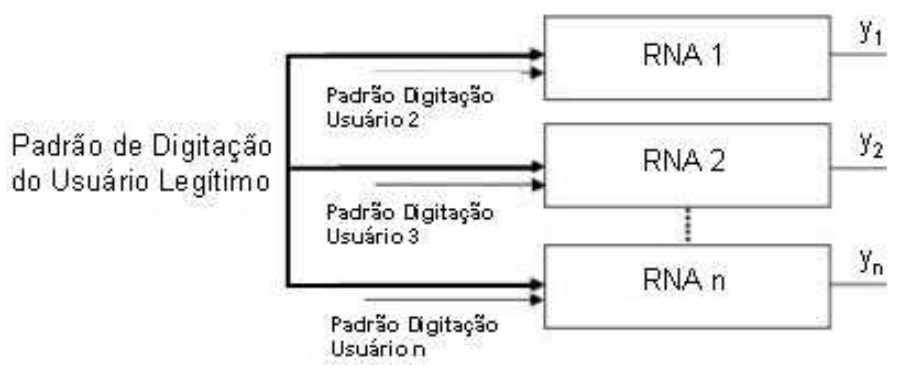

Figura 3. Máquina de Comitê proposta para autenticação de um usuário específico (Usuário 1)

Na etapa de autenticação do usuário, o padrão de digitação, ou seja, a latência entre dígrafos do texto que o usuário estiver digitando forma o conjunto de autenticação, que então é apresentado a máquina de comitê do usuário em questão. Neste caso, algumas redes do comitê conseguirão classificar corretamente o padrão de entrada enquanto que outras não serão capazes de fazer a classificação correta. A saída final da máquina de comitê, que é uma combinação das saídas de todas as RNA que compõem o comitê é que decidirá se o usuário é válido ou é um intruso. Esta combinação considera o número de redes que reconheceram o padrão de entrada como legítimo e quantas redes não foram capazes de reconhecê-lo. O valor de limiar do combinador, ou seja, a proporção de "votos" de reconhecimentos/não-reconhecimentos adequados são obtidos através de experimentos e se refletem nas taxas de falsa aceitação (FAR) e falsa rejeição (FRR), sendo uma das grandes contribuições deste trabalhos.

4.2. Metodologia dos Experimentos

Neste trabalho foram analisados 20 usuários sendo que estes não precisavam apresentar nenhuma característica especial, como por exemplo, terem curso de datilografia, mas deveriam apenas conhecer e saber utilizar um microcomputador e seu teclado. Destes usuários, podem-se identificar 2 grupos. No grupo 1 estão 14 usuários que tiveram seus dados coletados divididos em dois conjuntos: conjunto de treinamento e conjunto de testes. Já o grupo 2, constituído por 6 usuários, tiveram seus dados utilizados apenas no teste do sistema biométrico, sendo que estes usuários agiram apenas como "intrusos" e são essenciais para verificar a eficácia do sistema biométrico quanto à falsa aceitação de usuários.

Como informação alvo, foram utilizados uma frase fixa contendo 32 caracteres e um texto fixo com 347 caracteres correspondendo ao trecho de uma música popular, visando minimizar a interrupção da digitação pelo usuário para ler o conteúdo do texto, já que as interrupções geram coletas de medidas de tempo que geralmente não correspondem ao tempo padrão de digitação do usuário.

Em cada processo de coleta o usuário teve que digitar todas as informações alvo 5 vezes num mesmo dia. Este processo foi realizado uma única vez pelos usuários do grupo 2 e foi repetido 6 vezes pelos usuários do grupo 1, sendo que cada coleta foi realizada obrigatoriamente em dias distintos. Como resultado destas coletas tem-se que cada

usuário do grupo 1 forneceu 30 amostras de cada informação alvo, já os usuários do grupo 2 forneceram 5 amostras de cada informação alvo.

A característica da dinâmica da digitação extraída do processo de coleta neste trabalho resumiu-se ao tempo da latência entre o pressionamento de duas teclas consecutivas.

As amostras coletadas pelos usuários do grupo 1 foram divididas em conjunto de treinamento e conjunto de testes, sendo que estes conjuntos foram definidos da seguinte maneira:

- Conjunto de treinamento: composto pelas coletas ímpares de cada dia, ou seja, como cada usuário coletou 5 vezes cada informação alvo, a amostra de número 1, 3 e 5 de cada dia para a informação em questão formam o conjunto de treinamento do usuário para esta informação alvo.
- Conjunto de testes: as amostras pares compõem o conjunto de testes de cada usuário para a informação alvo

Portanto cada usuário do grupo 1 terá um conjunto de treinamento e um de testes para a frase fixa e o texto fixo. Já os usuários do grupo 2 terão apenas conjuntos de testes de cada informação, sendo que o conjunto deste grupo de usuários é formado por todas as coletas da informação, ou seja, cada conjunto de teste possui 5 amostras de coleta. Nos usuários do grupo 1, cada conjunto de treinamento é formado por 18 amostras de coleta de dados enquanto o conjunto de testes possui 12 amostras.

5. Experimentos e Resultados

Visando validar a arquitetura proposta e analisar seu desempenho, 2 experimentos foram propostos: no primeiro, o valor de limiar do combinador das saídas da redes que formam a máquina de comitê dos usuários é mantido fixo; no segundo experimento, os valores de limiar podem variar entre as diferentes máquinas de comitê.

5.1. Experimento 1 - Todas as Máquinas de Comitê com mesmo Limiar

Este experimento utilizou a arquitetura de máquinas de comitê proposta para tentar reconhecer a dinâmica de digitação de usuários cadastrados no sistema. A fim de verificar como a quantidade de teclas influencia o desempenho da autenticação, este experimento foi dividido em 2 etapas. Na primeira a autenticação era realizada por máquinas de comitê treinadas com uma frase fixa, enquanto na segunda etapa, as máquinas de comitê foram treinadas com um texto fixo. Em ambas as etapas, o valor de limiar das máquinas de comitê foi estabelecido em um valor fixo.

5.1.1. Resultados do Experimento 1

Como proposto na metodologia, a dinâmica de digitação de cada usuário foi treinada contra a dinâmica de digitação de outros usuários 2 a 2, num total de 13 RNAs treinadas compondo a máquina de comitê. A saída de cada RNA é combinada de modo que os resultados autenticam ou não o usuário de acordo com o valor de limiar proposto para a arquitetura.

Após o treinamento, dados válidos do conjunto de teste do usuário em questão são apresentados à máquina de comitê como forma de avaliar as taxas de falsa rejeição (FRR). Um conjunto de 170 dados válidos são utilizados nesta verificação. Já para avaliar as taxas de falsa aceitação (FAR) um conjunto com 2210 dados de cada um dos outros 13 usuários do grupo 1 e 490 dados de usuários do grupo 2 são apresentados à máquina de comitê.

Este experimento obteve a seguinte matriz de confusão, com suas respectivas taxas de falsa aceitação e de falsa rejeição.

Tabela 1. Resultados do Experimento 1

FRASE FIXA		**TEXTO FIXO**	
FAR	FRR	FAR	FRR
3,77 %	4,11 %	1,67 %	0,58 %

Na etapa de análise da frase fixa, os melhores resultados foram obtidos quando se utilizou um valor de limiar para o combinador tal que 8 das 13 redes classificassem as entradas como válidas e quando elas possuíssem uma taxa de acerto superior a 65% dos dígrafos apresentados.

Na etapa de análise do texto fixo, quando os mesmos valores de limiar foram utilizados, obteve-se uma FRR de 2,35% e uma FAR de 2,09%. Entretanto, os melhores resultados foram obtidos quando se utilizou um valor de limiar para o combinador tal que 9 das 13 redes classificassem as entradas como válidas e quando elas possuíssem uma taxa de acerto superior a 55% dos dígrafos apresentados.

Como a variância da latência de digitação entre 2 dígrafos é muito alta, o sistema de classificação deve ser de certa forma tolerante a esta alta variabilidade. Assim, podemos observar que os melhores resultados de classificação ocorreram quando utilizou-se valores intermediários para o limiar de aceitação e de taxa de acerto na classificação dos dígrafos apresentados. Valores altos de limiar levavam a taxas muito elevadas de falsa rejeição enquanto que valores muito baixos conduziam a altas taxas de falsa aceitação.

5.2. Experimento 2 - Valores de Limiar Específicos para cada Usuário

O segundo experimento utilizou os mesmos dados do primeiro experimento, entretanto, como observou-se que usuários diferentes possuíam variâncias diferentes na latência entre dígrafos, isto é, enquanto para alguns usuários a variância na latência entre dígrafos era grande, outros usuários possuíam um "ritmo de digitação" mais constante e portanto com uma menor variância, propomos que o valor de limiar do combinador das máquinas de comitê fosse distinto para cada usuário.

Como no experimento anterior, analisou-se o desempenho de classificação tanto para redes treinadas com uma frase fixa, como com um texto fixo. Os resultados, por usuário, podem ser vistos na Tabela 2.

Analisando a Tabela 2 podemos observar que, na classificação utilizando a frase fixa, 6 usuários foram perfeitamente identificados utilizando-se a estratégia de valores de limiar distintos para cada usuário. Além disso, podemos observar apenas 1 usuário válido foi incorretamente classificado pelo sistema.

Tabela 2. Resultados do Experimento 2

	FRASE FIXA		**TEXTO FIXO**	
Usuário	FAR	FRR	FAR	FRR
usuário 1	0 %	0 %	0 %	0 %
usuário 2	1,55 %	0 %	0 %	0 %
usuário 3	0 %	0 %	0 %	0 %
usuário 4	0 %	0 %	0 %	0 %
usuário 5	0 %	0 %	0 %	0 %
usuário 6	2,07 %	0 %	0 %	0 %
usuário 7	0,52%	16,66 %	0 %	0 %
usuário 8	2,07 %	0 %	0 %	0 %
usuário 9	0 %	0 %	0 %	0 %
usuário 10	1,55 %	0 %	0 %	0 %
usuário 11	1,04 %	0 %	0 %	0 %
usuário 12	0 %	0 %	0 %	0 %
usuário 13	2,07 %	0 %	0 %	0 %
usuário 14	3,63 %	0 %	2,13 %	0 %
Total	1,04%	1,18 %	0,15 %	0 %

Como no primeiro experimento, quando o classificador utilizou os dados de treinamento oriundos do texto fixo, seu desempenho aumentou significativamente. A Tabela 2 mostra que 13 usuários foram classificados corretamente e que outros usuários poderiam se fazer passar apenas por um dos usuários cadastrados (usuário 14).

Comparando com o primeiro experimento, observa-se uma grande melhora no desempenho do sistema, passando-se de uma FAR de 1,67% para uma FAR de 0,15% quando utilizando o texto fixo, e de uma FRR de 0,58% para uma FRR de 0%.

6. Conclusões

Este trabalho procurou demonstrar a utilização de RNAs para resolver problemas complexos de classificação. Deste estudo podemos concluir que é possível desenvolver um método seguro, barato e contínuo para autenticação de usuários através da característica biométrica da dinâmica de digitação. Este método utiliza como classificador uma variação da arquitetura de redes neurais organizadas em máquina de comitê.

Podemos observar que a utilização de apenas uma RNA não é apropriado para este tipo de tarefa, devido ao elevado número de atributos, superposições e variabilidade do conjunto de treinamento. Entretanto, ao utilizarmos um conjunto de redes neurais organizadas como um conjunto de "especialistas", os resultados obtidos são comparáveis aos melhores resultados apresentados na literatura utilizando métodos estatísticos de classificação. Também é importante ressaltar que, como método de classificação, a metodologia proposta poderia ser utilizada com outras características biométricas comportamentais, como, por exemplo, a utilização do "mouse".

Os experimentos de classificação utilizando a máquina de comitê mostram que utilizando um par de limiares fixo para todos os usuários o sistema apresentou uma FRR=4,11% e FAR=3,77% quando as RNAs foram treinadas e testadas através da frase

fixa. Quando a informação alvo utilizada foi o texto fixo os resultados foram FRR=0,58% e FAR=1,67%. Quando os pares de limiares não são fixos e cada usuário é analisado por limiares específicos, o sistema apresenta FRR=1,18% e FAR=1,04% para análise da frase fixa, e FRR=0% e FAR=0,15% para análise do texto fixo. Que demonstra um desempenho superior utilizando o segundo método.

Muitos trabalhos nesta área de pesquisa mostram como principal fator negativo da utilização de RNAs a necessidade de re-treinamento das redes quando um novo usuário é adicionado ao sistema. Neste ponto está uma das grandes vantagens da metodologia proposta, pois com a utilização da máquina de comitê adaptada, este trabalho de re-treinamento não é necessário. Quando um novo usuário é adicionado no sistema de autenticação proposto basta treinar as RNAs que constituirão a máquina de comitê do usuário em questão e uma nova RNA para cada máquina de comitê dos usuários existentes. Após treinar todas as RNAs basta modificar a configuração do combinador da máquina de comitê para que este leve em consideração mais uma RNA. Neste trabalho, utilizamos um comitê de 13 especialistas para classificar um usuário válido, entretanto, existe a conjectura que o número efetivo de especialistas necessários para fazer a classificação, à medida que novos usuários forem sendo adicionados ao sistema, não necessita crescer na mesma proporção que o número de novos usuários, o que representa um ganho na questão da escalabilidade do sistema. Esta possibilidade necessita ser comprovada em trabalhos futuros.

É bastante difícil a comparação de desempenho deste trabalho com outros apresentados na literatura, uma vez que cada autor utiliza um conjunto de dados e uma sistemática diferente. Apesar dos resultados iniciais parecerem bastante promissores, seria necessário aumentar o número de usuários analisados, tanto do grupo 1 como do grupo 2: para este trabalho foram analisados 20 usuários, uma pesquisa com uma maior quantidade de usuários poderia demonstrar se a eficácia do sistema permanece estável na utilização em larga escala.

Referências

[1] ALEXANDRE, T. J. Biometrics on smartcards: An approach to keyboard behavioral signature. In Second Smart Card Researsh & Advanced Applications Conference, 1996.

[2] ANAGUN, A. S. & CIN, I. A Neural Network based Computer Access Security System for Multiple Users. Computers & Industrial Engineering, Vol. 35, Nº. 1-2, pp. 351-354, 1998.

[3] BERGADANO, F.; GUNETTI, D. & PICARDI C. User authentication through Keystroke Dynamics. ACM Transactions on Information and System Security (TISSEC), Vol. 5, Nº. 4, pp. 367-397, 2002.

[4] BLEHA, S.; SLIVINSKY, C. & HUSSIEN, B. Computer-Access Security Systems Using Keystroke Dynamics. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, Nº.12, pp. 1217-1222, 1990.

[5] BROWN, M. & ROGERS, S. J. User identification via keystroke characteristics of typed names using neural networks. International Journal of Man-Machine Studies, Vol. 39, Nº. 6, pp. 999-1014, 1993.

[6] CAPUANO, N.; MASELLA, M.; MIRANDA, S. & SALERNO, S. User authentication with neural networks. Proceedings of the 5th International Conference on Engineering Applications of Neural Networks EANN 99, Warsaw, Poland, 1999.

[7] COLTELL, O.; BADÍA, J.M. & TORRES, G. Biometric Identification System Based in Keyboard Filtering. Proceedings of XXXIII Annual IEEE International Carnahan Conference on Security Technology, IEEE Pub., pp. 203-209, 1999.

[8] GAINES, R.; LISOWSKI, W.; PRESS, S. & SHAPIRO, N. Authentication by keystroke timing: Some preliminary results. Rand Report R-256-NSF. Rand Corporation, Santa Mônica, CA, 1980.

[9] HAYKIN, S. Redes Neurais - Princípios e Práticas. Segunda Edição, Porto Alegre, Editora Bookman, 2001, 900 p.

[10] ILONEN, J. Keystroke Dynamics. Disponível em: http://www.it.lut.fi/kurssit/03-04/010970000/seminars/Ilonen.pdf. Acesso em: 24 jul. 2005.

[11] JOYCE, R. & GUPTA G. Identity Authentication Based on Keystroke Latencies. Communications of the ACM, Vol. 33, Nº. 2, pp. 168 - 176, 1990.

[12] MONROSE, F.; REITER, M. K. & WETZEL, S. Password Hardening Based on Keystroke Dynamics. Proceedings of the 6th ACM conference on Computer and communications security, 1999.

[13] MONROSE, F. & RUBIN, A. D. Authentication via Keystroke Dynamics. Proceedings of the 4th ACM conference on Computer and communications security, pp. 48-56, 1997.

[14] OBAIDAT, M. S. A verification Methodology for Computer Systems Users. Proceedings of the 1995 ACM symposium on Applied computing, Nashville, Tennessee, United States, pp. 258-262, 1995.

[15] OBAIDAT, M. S. & SADOUN, B. Verification of Computer Users Using Keystroke Dynamics. IEEE Transactions on System, Man and Cybernetics, Vol. 27, Nº. 2, pp. 261-269, 1997.

[16] PEACOCK, A. Learning User Keystroke Latency Patterns (Preliminary Report). Disponível em: http://pel.cs.byu.edu/ alen/personal/CourseWork/cs572/KeystrokePaper/. Acesso em: 24 jul. 2005.