# Autenticação Biométrica via Dinâmica da Digitação em Teclados Numéricos

Carlos Roberto do N. Costa, Glauco F. G. Yared, Ricardo Nagel Rodrigues, João B. T. Yabu-Uti, Fábio Violaro, Lee Luan Ling

***Resumo*— Este artigo apresenta nova abordagem para autenticação biométrica via dinâmica da digitação em teclados numéricos. O sinal de entrada é obtido em tempo real durante a digitação pelo usuário da *String* alvo. Cinco características são extraidas do sinal (código ASCII da tecla e quatro durações associadas) e quatro experimentos usando amostras de usuários e impostores foram analisados comparando-se dois classificadores de padrões. Obteve-se melhores resultados com *HMM* (EER=4,5%). Esta nova abordagem traz melhorias ao processo de autenticação pois permite que a senha não seja mais segredo, assim como permite incluir autenticação biométrica em dispositivos móveis, como celulares.**

***Palavras-Chave*— Processamento digital de sinais, reconhecimento de padrões, segurança, biometria, dinâmica da digitação.**

***Abstract*— This paper presents a new approach for biometric authentication using keystroke dynamics in numerical keyboards. The input signal is generated in real time when the user enters the target string. Five features are extracted from this input (key ASCII code and four associated durations) and four experiments using samples for genuines and impostors users were performed using two pattern classification techniques. The best results were achieved by the *HMM* (EER=4.5%). This new approach brings improvements to the process of user authentication since it allows the password not to be a secret anymore, as well as it allows to include biometric authentication in mobile devices, such as cell phones.**

***Keywords*— Digital signal processing, pattern recognition, security, biometrics, keystroke dynamics.**

## I. Introdução

Controlar o acesso a sistemas computacionais torna-se cada vez mais importante nos dias de hoje, e o mecanismo mais conhecido e usual para garantir segurança em sistemas de informação é através da autenticação do usuário por uma senha. Porém este tipo de mecanismo é frágil pois existem usuários descuidados que comprometem a segurança quando utilizam-se de senhas frágeis e de contexto normalmente familiar, como por exemplo uma data de nascimento. Por outro lado, o custo e a simplicidade deste tipo de mecanismo clássico de segurança justifica sua adoção, e em várias situações permanece como mecanismo principal ao lado de outras políticas de segurança. O propósito deste trabalho é melhorar o processo de autenticação por senha usando características biométricas. Características biométricas são padrões observados no ser humano que permitem criar algoritmos capazes de distinguir uma pessoa da outra, e podem ser classificadas como sendo características fisiológicas ou características de natureza comportamental [1]. A inclusão de características biométricas em sistemas de autenticação pessoal aumenta o grau de confiança dos usuários na segurança do sistema, pois características biométricas são únicas para cada pessoa e não podem ser roubadas, perdidas ou esquecidas.

A tecnologia biométrica tratada neste trabalho é conhecida como biometria da digitação ou dinâmica da digitação. A biometria da digitação é o processo de analisar o ritmo com que o usuário digita em um terminal através do monitoramento das entradas do teclado durante as tentativas de identificação. Autenticação por dinâmica da digitação pode ser classificada em estática ou contínua. A abordagem estática analisa as entradas em um momento particular, enquanto que na abordagem contínua são analisadas todas as entradas no teclado durante a sessão do usuário [2].

Este trabalho inova na forma como usa a dinâmica da digitação, pois utiliza somente um teclado numérico para observar a dinâmica com que o usuário digita uma senha exclusivamente numérica. O uso do teclado numérico diminui o intervalo de tempo entre teclas e aumenta ainda mais a exigência de precisão, além de praticamente forçar o usuário a usar somente uma das mãos. Além disso, permite que sejam propostos soluções para funcionar em telefones celulares, em sistemas de caixa automático para bancos ou mesmo no acesso a áreas restritas. A metodologia adotada neste trabalho tem um baixo custo de processamento, é não-intrusiva e verifica o usuário de maneira estática, ou seja, apenas considera a entrada digitada em um dado momento.

O resto do trabalho está organizado da seguinte maneira. Na seção 2 são apresentados resumidamente os trabalhos publicados e relacionados com a área. Na seção 3 é discutida a metodologia proposta na extração das características e são apresentados os classificadores utilizados. Na seção 4 são apresentados os experimentos conduzidos e os resultados obtidos; e finalmente na seção 5 são apresentadas as conclusões e as propostas de trabalhos futuros.

## II. Trabalhos Relacionados

Autenticação biométrica por dinâmica da digitação é uma área de pesquisa ativa desde 1990 [2]-[14]. Alguns aspectos sobre sistemas desta natureza são discutidos neste trabalho e resumidamente apresentados a seguir.

- ***String* Alvo**: É a *string* que será digitada pelo usuário e monitorada pelo sistema. Em [3] são usadas quatro

Os autores pertencem ao Departamento de Comunicações, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, São Paulo, Brasil, E-mails: {ccosta, glauco, ricardonagel, yabuuti, fabio, lee}@decom.fee.unicamp.br.

*strings* como alvo durante a autenticação (usuário, senha, primeiro nome e último nome). Porém, em alguns trabalhos, somente a senha é suficiente. Outro aspecto importante diz respeito ao tamanho da *string*. Em [4] os autores perceberam que o sistema fica sujeito a mais erros durante o processo de classificação quando *strings* de entrada menores que dez são adotadas pelos usuários.

- **Número de amostras**: Amostras são coletadas durante o processo de cadastramento dos usuários, e irão compor o conjunto de treinamento do classificador. O número de amostras varia muito, sendo que a menor quantidade foi relatada em [5], onde somente três amostras foram usadas, e o maior número de amostras foi relatado em [6], onde pediu-se ao usuário que digitasse trinta amostras. Em [2] os autores observaram que o número mínimo de amostras para não comprometer o desempenho do sistema é de seis amostras por usuário.
- **Extração das características**: Duas das características mais observadas durante a digitação são o tempo em que a tecla permanece pressionada e o intervalo de tempo entre teclas sucessivas [5]. Em [7]-[9] os autores combinam estas características, obtendo melhores resultados do que usando isoladamente cada uma delas. De Ru *et al.* [10] analiza uma característica diferente baseada na distância das teclas no teclado alfa-numérico e a combinação de teclas consideradas difíceis, porém obriga o usuário a decorar uma *string* à qual não está habituado. Como a maioria das características observadas são temporais, a precisão com que se deve observar o tempo onde uma determinada tecla é pressionada ou solta torna-se importante. Os intervalos entre teclas podem variar entre 0.1ms [7] e 1000ms (1s) [11].
- **Tentativas de autenticação**: Em [12] os autores observaram que usuários legítimos normalmente falham na primeira tentativa de autenticação, sendo normalmente autenticado na segunda tentativa. Em [6], cada usuário deve digitar sua *string* de entrada duas vezes, usando uma técnica de embaralhamento.
- **Mecanismo de adaptação**: Características biométricas podem sofrer pequenas mudanças com o passar do tempo. Portanto, faz-se necessário em sistemas biométricos a presença de mecanismos de adaptação, ou seja, um recadastramento pode ser feito para manter o modelo do usuário sempre atualizado. A maioria dos pesquisadores não menciona este aspecto importante em seus trabalhos, porém em [13] os autores citam um mecanismo de adaptação onde sempre que um usuário é autenticado de maneira correta o sistema calcula outro modelo acrescentando a entrada atual como uma nova amostra e descartando a mais antiga.
- **Classificador**: Em [2]-[4], [6], [7] e [11], os autores adotaram classificadores estatísticos em seus experimentos, como por exemplo *k-means*, Bayes, etc. Em [9] e [10], lógica nebulosa foi aplicada usando como saída um categorizador de usuários. Finalmente, em [5], [8] e [14], redes neurais artificiais foram usadas para identificar o usuário.

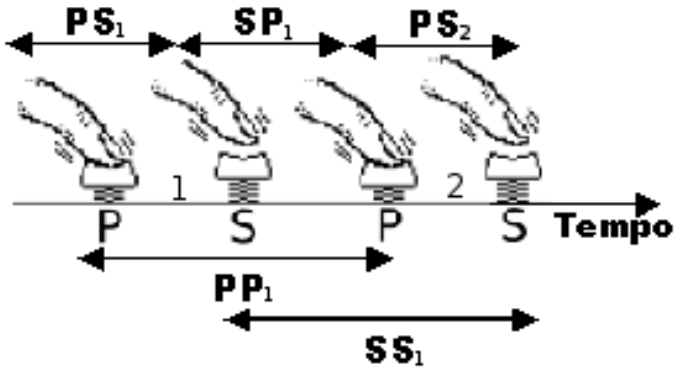


Fig. 1. Representação das características observadas durante a digitação dos caracteres 1 e 2. *PS* é o tempo em que a tecla permanece pressionada, *SP* é o intevalo até a próxima tecla ser pressionada, *PP* é o intervalo de tempo que o usuário leva para pressionar duas teclas consecutivas e *SS* é o intervalo de tempo que o usuário leva para soltar duas teclas consecutivas.

## III. METODOLOGIA

Este trabalho inova pois, ao invés de considerar o teclado alfa-númerico e uma *string* que pode conter tanto letras como números, restringe o usuário ao teclado númerico, com senhas numéricas limitadas em oito caracteres e escolhidas pelo usuário. Teclados desta natureza estão presentes em telefones celulares, caixas automáticos de banco ou em controle de acesso a áreas restritas. Além disso, também utiliza uma nova abordagem no processo de classificação que se baseia em Modelos Ocultos de Markov (*Hidden Markov Models*, *HMM*).

Fundamentalmente a dinâmica da digitação necessita adquirir os padrões do digitador de forma precisa no tempo [14]. Neste trabalho usamos a função *Time Stamp Counter* para adquirir o número preciso de ciclos do processador, tal qual descrita em [14]. A precisão adotada deve manter-se constante para toda a base de dados, onde 98% das amostras coletadas estão entre 10 e 900ms, portanto, 1ms de precisão foi adotado neste trabalho.

Nas próximas seções são discutidos os aspectos relacionados com a extração de características, construção do modelo do usuário e classificação do padrão.

### *A. Extração das características e base de dados*

Para a construção da base de teste usou-se um teclado numérico, onde dez amostras com oito caracteres cada são coletadas em cada sessão e para cada usuário, totalizando 40 amostras (4 sessões por usuário), sendo que vinte usuários foram convidados a participar do experimento. Em [15] observou-se que mais de dez amostras por sessão no cadastramento incomoda os usuários e em [2] os autores observaram que quanto menor o número de amostras, pior será o desempenho do classificador.

Cada amostra de *n* caracteres pode conter várias características de digitação, onde uma determinada característica de digitação para a amostra *w* da conta *a* pode ser difinida por $K_{a,w}$ = ($K_1(a,w)$, $K_2(a,w)$,...,$K_n(a,w)$). Cada característica $K_i(a,w)$, onde $i \leq n$, representa uma das seguintes características observadas:

- Código ASCII: O caracter digitado possui uma entrada na tabela ASCII. $C_a$ = ($C_1(a)$, $C_2(a)$,...,$C_n(a)$) representa os códigos das respectivas teclas para o teclado numérico e que estão contidos no modelo do usuário *a* e $C_{a,w}$ = ($C_1(a,w)$, $C_2(a,w)$,...,$C_n(a,w)$) representa os códigos das teclas da amostra *w* para a conta do usuário *a*;
- O intervalo de tempo em que a tecla permanece pressionada (PS) [5]. Esta característica é representada pela expressão $PS_{a,w}$ = ($PS_1(a,w)$, $PS_2(a,w)$,...,$PS_n(a,w)$), onde $PS_i(a,w)$ = $T_{i.solta}(a,w)$ - $T_{i.pressiona}(a,w)$ está relacionado com $K_i$. $T_{i.solta}(a,w)$ é o instante onde a tecla *i* é solta e $T_{i.pressiona}(a,w)$ é o instante onde a tecla *i* é pressionada;
- O intervalo de tempo até a próxima tecla ser pressionada (SP). Esta característica é representada pela expressão $SP_{a,w}$ = ($SP_1(a,w)$, $SP_2(a,w)$,...,$SP_{n-1}(a,w)$), onde $SP_i(a,w)$ = $T_{i+1.pressiona}(a,w)$ - $T_{i.solta}(a,w)$ está relacionado com ($K_i$,$K_{i+1}$);
- O intervalo de tempo que o usuário leva para pressionar duas teclas consecutivas (PP). Esta característica é representada pela expressão $PP_{a,w}$ = ($PP_1(a,w)$, $PP_2(a,w)$,...,$PP_{n-1}(a,w)$), onde $PP_i(a,w)$ = $T_{i+1.pressiona}(a,w)$ - $T_{i.pressiona}(a,w)$ está relacionado com ($K_i$,$K_{i+1}$);
- O intervalo de tempo que o usuário leva para soltar duas teclas consecutivas (SS). Esta característica é representada pela expressão $SS_{a,w}$ = ($SS_1(a,w)$, $SS_2(a,w)$,...,$SS_{n-1}(a,w)$), onde $SS_i(a,w)$ = $T_{i+1.solta}(a,w)$ - $T_{i.solta}(a,w)$ está relacionado com ($K_i$,$K_{i+1}$);

A figura 1 mostra um exemplo de extração de durações associadas à digitação dos caracteres um e dois.

Além da base de usuários descrita anteriormente, três usuários foram convidados a participar como impostores de cada uma das vinte contas. Ao final de dez tentativas, 600 amostras de impostores foram coletadas. A figura 2 apresenta as distribuições entre autênticos e impostores usando o método descrito em [15]. Pode-se observar que a sobreposição entre as classes é visível, o que ocasiona altas taxas de erro. Esta sobreposição ocorre pois em teclados numéricos os usuários digitam de maneira similar, ou seja, usando apenas uma mão e com ritmos semelhantes.

### *B. Classificador estatístico*

Em [15] os autores sugerem que o modelo do usuário seja gerado a partir da média e do desvio padrão entre as amostras observadas no cadastramento. Sempre que o usuário tenta autenticar sua senha, o sistema calcula a distância da *string* alvo para o modelo e, se ela for maior que um limiar pré-definido, então o usuário é autenticado. O modelo é gerado a partir de *N* amostras, podendo conter as características *K*={PP, SP, PS, SS} descritas anteriormente, e de acordo com as equações (1) e (2), re-escritas a seguir:

$$\mu_{K_i(a)} = \frac{1}{N}\sum_{j=1}^{N} K_i(a,j) \qquad (1)$$

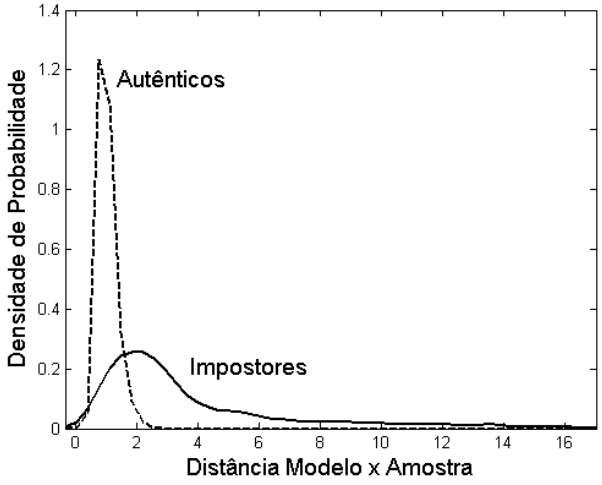


Fig. 2. Distribuição entre autênticos e impostores para a base de dados usada nos testes.

$$\sigma_{K_i(a)} = \frac{1}{N-1}\sum_{j=1}^{N} |K_i(a,j) - \mu_{K_i(a)}| \qquad (2)$$

Na autenticação, o código ASCII identifica inicialmente o usuário que deseja ser autenticado como dono da conta *a*. Conhecendo inicialmente quem clama por sua identidade, o sistema conduz a uma comparação um-para-um, onde a distância total de cada uma das características observadas para o modelo da conta *a* é calculado, através da equação (3):

$$D_K(a,w) = \frac{1}{n}\sum_{i=1}^{n} \frac{K_i(a,w) - \mu_{K_i(a)}}{\sigma_{K_i(a)}} \qquad (3)$$

onde *n* é o número de elementos da característica *K*, para *K*={PP, SP, PS, SS}. Se $D_K(a,w) \leq \tau_K(a)$, para todo *K*, o usuário é considerado autêntico para a conta *a*. $\tau_K(a)$ é o limiar de decisão da conta *a* definido empiricamente.

Ainda em [15] foi proposto também um mecanismo de adaptação do modelo em que, se o usuário foi autenticado positivamente, então a *string* alvo é adicionada ao modelo e o sistema recalcula a média e o desvio padrão entre as amostras armazenadas e a nova amostra, usando uma metodologia semelhante a uma fila: descarta a amostra mais antiga e acrescenta a amostra recente. A figura 3 ilustra o processo de atualização do conjunto de treinamento ao longo do tempo.

Os autores reportaram taxas de falsa-aceitação e falsa-rejeição menores que 2% usando esta abordagem em teclados alfa-numéricos com *strings* contendo letras e/ou números.

### *C. Classificador usando HMM*

Os sistemas baseados em *HMM* têm sido amplamente utilizados em reconhecimento de padrões [17],[16]. A utilização de tais sistemas está, em grande parte, associada à necessidade de se modelar a variabilidade temporal dos padrões analisados. Além disso, a utilização de misturas Gaussianas permite a modelagem de distribuições complexas, cujas fronteiras são de grande importância do ponto de vista da classificação.

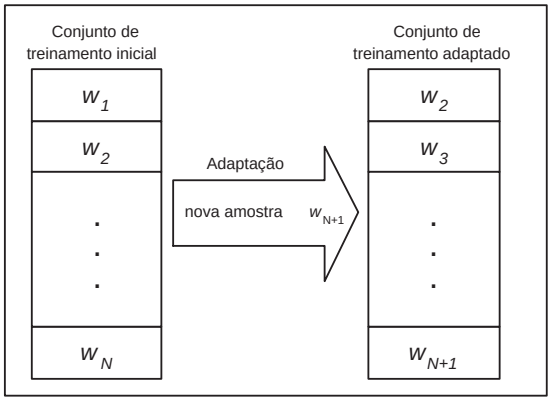


Fig. 3. Mecanismo de adaptação do modelo.

A sobreposição das distribuições de classes distintas dificulta a determinação das fronteiras e é um problema comum no processo de modelagem, podendo ser tratado pela escolha apropriada do número e localização das componentes de cada mistura. No entanto, se a quantidade de dados disponíveis para o processo de modelagem for insuficiente, torna-se difícil a determinação do número apropriado de Gaussianas por mistura para modelar cada classe.

O modelo probabilístico empregado para representar cada usuário com sua senha específica pode ser compreendido como um caso particular de um *HMM* contínuo com 15 estados e topologia *left-right*. Cada estado está associado ao tempo em que uma tecla fica pressionada ou ao intervalo para o digitador se deslocar entre uma tecla e outra. Assim:

- Estado 1, associado ao tempo em que a primeira tecla fica pressionada;
- Estado 2, associado ao tempo para o digitador se mover entre a primeira e a segunda tecla;
- Estado 3, associado ao tempo em que a segunda tecla fica pressionada;

.
.
.

- Estado 14, associado ao tempo para o digitador se mover entre a sétima e a oitava tecla;
- Estado 15, associado ao tempo em que a oitava tecla fica pressionada;

As probabilidades de transição são também dadas por $a_{i(i+1)} = 1$, ou seja, não ocorrem auto-transições ($a_{ii} = 0$). Cada novo parâmetro de entrada $K_i(a, w)$, *i* de 1 a 15, está associado ao estado *i* e é modelado através de uma mistura de 6 Gaussianas unidimensionais. A cada novo parâmetro $K_i(a, w)$, o sistema avança do estado *i* para o estado *ì+1*.

A idéia consiste basicamente em modelar cada duração por um estado do modelo, uma vez que o número de durações é fixo para todos os digitadores (8 dígitos resultando em 15 durações do tipo tecla pressionada e intervalo entre teclas). A figura 4 ilustra um exemplo de digitação da senha constituída pela sequência de dígitos "2-5-0-5-7-8-9-4" e os estados do *HMM* utilizados na modelagem. Dessa forma, para uma dada senha digitada, utiliza-se o modelo *HMM* correspondente a fim

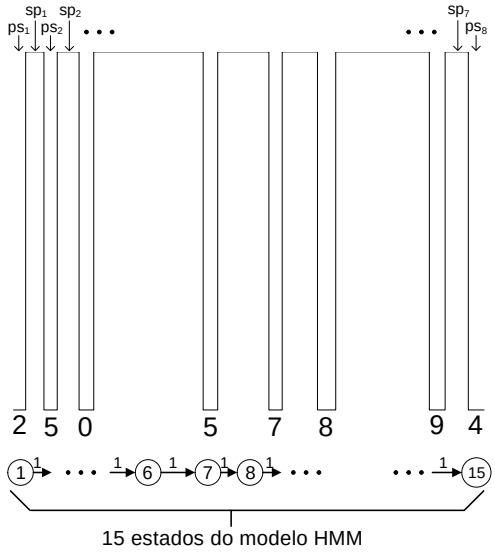


Fig. 4. Digitação da sequência "2-5-0-5-7-8-9-4". O HMM utilizado para a modelagem das durações correspondentes e as probabilidades de transição estão indicados na figura.

de se obter o valor de verossimilhança $P(O_{a,w}|\lambda_a)$ através do algoritmo de Viterbi [16], em que

$$O_{a,w} = \{PS_1(a,w), SP_1(a,w), PS_2(a,w), SP_2(a,w), \ldots, PS_{n-1}(a,w), SP_{n-1}(a,w), PS_n(a,w)\},$$

e $\lambda_a$ é o conjunto de parâmetros do modelo associado à conta *a*.

Se tal valor for superior a um limiar $\tau_K(a)$ pré-definido para a conta *a*, então o usuário é autenticado, caso contrário é considerado impostor e então rejeitado. O sistema foi treinado utilizando-se as ferramentas do HTK [18], através do algorítmo de Baum-Welch [17]. Os modelos são gerados inicialmente a partir de *N* amostras de cada digitador. Na sequência, para cada uma das amostras da base de teste, realiza-se a autenticação e, caso o usuário seja autenticado, tal amostra é adicionada à base de treinamento de maneira semelhante ao adotado para o classificador estatístico, e o sistema é treinado novamente.

Na próxima seção discute-se os experimentos conduzidos usando os classificadores propostos, bem como demonstra-se os resultados obtidos com a base de teste.

## IV. Experimentos e Resultados

Os experimentos foram conduzidos em um computador *Pentium IV* e utilizando a região do teclado composta de teclas numéricas. Vinte usuários, entre homens e mulheres de várias faixas etárias e com diferentes níveis de familiaridade com o teclado numérico participaram do experimento, usando como *string* alvo uma combinação de oito números a sua escolha. Duas situações foram observadas:

- *Usuários autênticos*: O usuário tenta autenticar-se em sua conta. Foram coletadas dez amostras do usuário em cada uma das 4 seções, totalizando 800 amostras de usuários autênticos.

- *Usuários impostores*: Impostores foram convidados a tentar autenticar-se nas contas dos usuários autênticos. Cada conta foi atacada 30 vezes, resultando em 600 amostras de impostores.

As amostras dos digitadores foram coletadas no Laboratório de Reconhecimento de Padrões e Redes de Computadores (LRPRC) da UNICAMP. As amostras dos usuários foram coletadas em diferentes períodos, e nunca ao mesmo tempo, intercalando um período médio de uma semana entre as seções.

Foram feitos três experimentos com o classificador estatístico:

1) utilizando dez amostras ($N$=10) para gerar o modelo com $K$={PP, SP, PS, SS};
2) utilizando vinte amostras ($N$=20) para gerar o modelo com $K$={PP, SP, PS, SS};
3) utilizando trinta amostras ($N$=30) para gerar o modelo com $K$={PP, SP, PS, SS}.

Além dos experimentos com o classificador estatístico, foi feito um experimento com *HMM* utilizando trinta amostras ($N$=30) para a obtenção do modelo.

### A. Resultados

O desempenho de sistemas biométricos é geralmente medido por três tipos de taxas de erros [1]:

- Taxa de falsa-aceitação (*FAR*): A probabilidade do sistema falhar na rejeição de usuários impostores.
- Taxa de falsa-rejeição (*FRR*): A probabilidade do sistema falhar quando verifica se um usuário legítimo é quem diz ser.
- Taxa de erro igual (*EER*): É o valor assumido quando *FAR* e *FRR* são iguais.

Para representar estas taxas, utiliza-se o gráfico *ROC* (*receive operating curve* [1]), que representa a *FAR* pela *FRR* para diferentes limiares, e apresenta graficamente o ponto onde os erros são iguais (*EER*).

As figuras de 5 a 7 representam as curvas *ROC* para o classificador estatístico nos experimentos (1), (2) e (3). Pode-se observar que o melhor resultado foi alcançado quando o valor de *N* empregado na construção do modelo foi de 20 amostras, resultando uma taxa de *EER* igual a 6.2%. Vale ressaltar que em [15] os autores obtiveram taxa de *EER* de 1.6%, o que é bastante plausível pois seus experimentos foram conduzidos em um teclado alfa-numérico, os usuários utilizam senhas com *string* alvo contendo mais de 8 caracteres e com durações mais representativas do que as obtidas em teclados numéricos.

A figura 8 mostra o desempenho do sistema quando o classificador usado é o *HMM* com $N$=30. Podemos observar que os resultados obtidos com *HMM* superam os obtidos com o classificador estatístico. O método de atualização do modelo utilizado permitiu representar de uma forma mais eficiente a dinâmica do digitador, o que possívelmente é causado pela não-estacionariedade dos dados. Futuramente, deve-se utilizar técnicas de seleção de topologia [19] para o modelo *HMM* afim de se obter modelos mais consistentes. Vale ressaltar que a utilização de menos amostras para gerar o modelo *HMM* dificulta a estimação dos parâmetros dos usuários no teclado numérico, pois a quantidade de dados torna-se insuficiente, o que foi comprovado também no classificador estatístico.

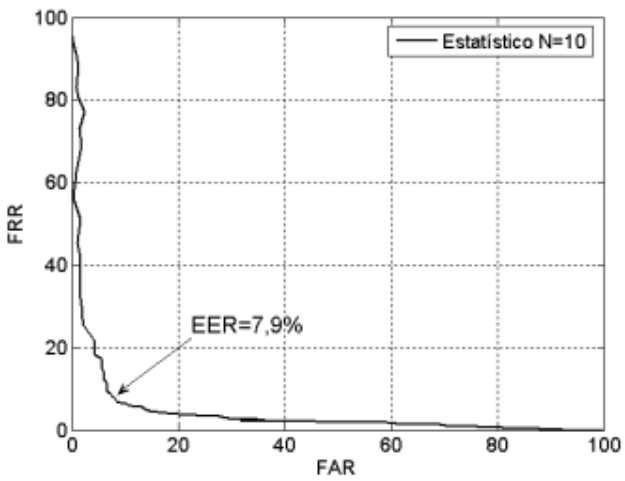


Fig. 5. *ROC* para o classificador estatístico quando $N$=10.

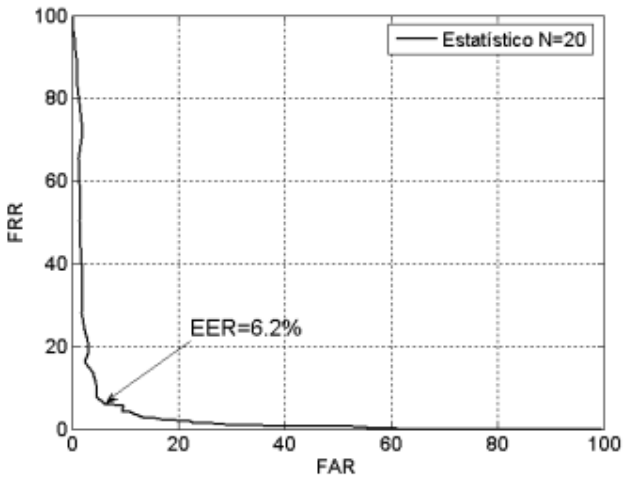


Fig. 6. *ROC* para o classificador estatístico quando $N$=20.

## V. Conclusões e trabalhos futuros

Este trabalho apresentou uma metodologia de autenticação biométrica através das características extraídas da dinâmica da digitação que permite melhorar o processo de controle de acesso a áreas restritas ou aumentar a segurança de transações bancárias. Alguns experimentos foram conduzidos e o melhor resultado foi alcançado usando um classificador baseado em modelos ocultos de Markov e combinando três características (código ASCII e as durações SP e PS), onde foi obtida *EER* de 4.5%. Esta taxa é competitiva, pois foi comparada com o classificador estatístico proposto em [15] e obteve melhores resultados para uma base de dados que restringe o usuário ao teclado numérico e usando apenas uma *string* alvo no cadastramento. Essa abordagem não havia sido feita até o momento na literatura mundial.

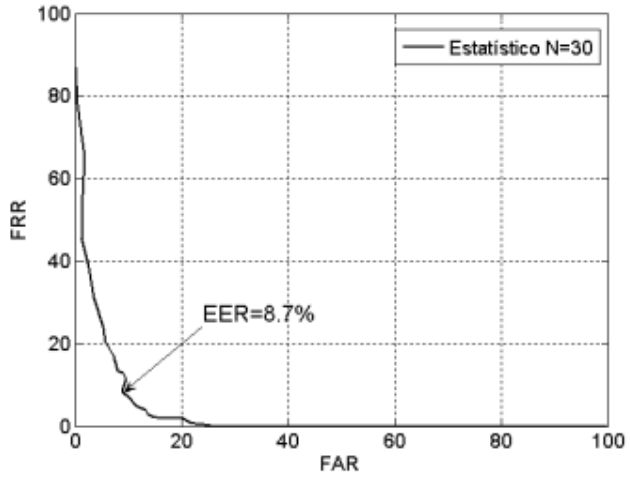


Fig. 7. *ROC* para o classificador estatístico quando $N$=30.

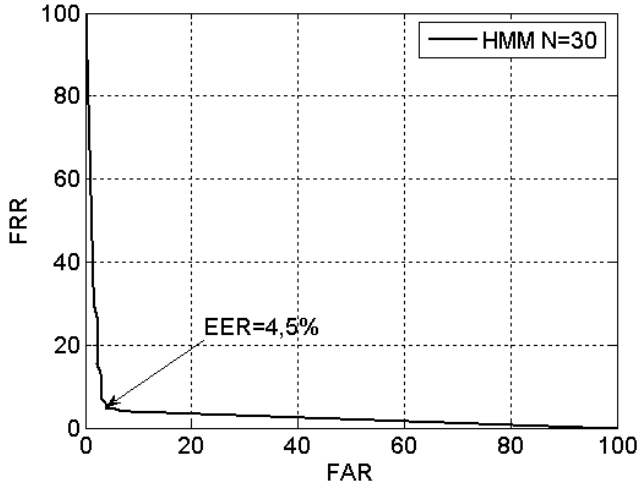


Fig. 8. *ROC* para o classificador *HMM* quando $N$=30.

Foi observado também a influência de alguns aspectos práticos, os quais foram testados e comprovados, mostrando que eles definitivamente devem ser considerados no desempenho de sistemas desse tipo. Estes aspectos são a familiaridade do usuário com a *string* alvo, o mecanismo de adaptação do modelo adotado no sistema, a precisão como os dados são adquiridos e, principalmente, o número de amostras usadas no cadastramento.

Para trabalhos futuros, pretende-se aumentar a população de usuários e capturar mais seções dos mesmos, pois observou-se a não-estacionariedade da dinâmica ao longo do tempo. Além disso, um mecanismo de adaptação do modelo *HMM* será proposto afim de baixar ainda mais a *EER* do sistema. Outro aspecto a ser melhorado é a forma de obtenção da topologia do *HMM*, na qual deve-se utilizar técnicas de seleção de estrutura de modelos [19].

## AGRADECIMENTOS

Os autores gostariam de expressar seus sinceros agradecimentos à CAPES e ao CNPQ pelo apoio financeiro e a todos os usuários que colaboraram pacientemente com a coleta de dados. Sem eles este trabalho não seria possível.

## REFERÊNCIAS

[1] A. K. Jain, A. Ross and S. Prabhakar, *An Introduction to Biometric Recognition*, IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image and Video-Based Biometrics, Vol. 14, No. 1, pp. 4-20, January 2004.

[2] F. Monrose and A. D. Rubin, *Keystroke Dynamics as a Biometric for Authentication*, Future Generation Computer Systems, Vol. 16, no. 4, pp. 351-359, March 1999.

[3] R. Joyce and G. Gupta, *Identity authentication based on keystroke latencies*, commun. ACM, vol. 33, no. 2, pp. 168-176, 1990.

[4] d. Bleha and M. Obaidat, *Dimensionality reduction and feature extraction applications in identifying computer users*, IEEE Trans. Syst., Man, Cybern., Vol. 21, no. 2, pp. 452-456, Mar.-Apr. 1991.

[5] D. T Lin, *Computer-access authentication with neural network based keystroke identity verification*, in Proc. Int. Conf. Neural Networks, vol. 1, 1997, pp. 174-178.

[6] S. Bleha, C. Slivinsky, and B. Hussain, *Computer-access security systems using keystroke dynamics*, IEEE Trans. Pattern Anal. Machine Intell., vol. 12, no. 12, pp. 1217-1222, Dec. 1990.

[7] J. A. Robison, V. M. Liang, J. A. Michael, and C. L. MacKenzie, *Computer user verification login string keystroke dynamics*, IEEE Trans. Syst., Man, Cybern., vol. 28, no. 2, pp. 236-241, Mar.-Apr. 1998.

[8] M. S. Obaidat and B. Sadoun, *Verification of computer user using keystroke dynamics*, IEEE Trans. Syst., Man, Cybern., vol. 27, no. 2, pp. 261-269, Mar.-Apr. 1997.

[9] L. C. F. Araújo, L. H. R. Sucupira Jr., M. G. Lizárraga, L. L. Ling, and J. B. T. Yabu-uti, *A fuzzy logic approach in typing biometrics user authentication*, in Proc. 1st Indian Int. Conf. Artificial Intelligence, 2003, pp. 1038-1051.

[10] W. G. de Ru and J. H. P. eloff *Enhaced password authentication through fuzzy logic*, IEEE Expert, vol. 17, no. 6, pp. 38-45, Nov.-Dec. 1997.

[11] O. Coltell, J. M. badfa, and G. Torres, *Biometric identification system based in keyboard filtering*, in Proc. IEE 33rd Annu. Int. Carnahan Conf. Security Technology, 1999, pp. 203-209.

[12] S. Haidar, A. Abbas, and A. K. Zaidi, *A multi-technique approach for user identification trhough keystroke dynamics*, in Proc. IEEE Int. Conf. Systems, Man and Cybernetics, vol. 2, 2000, pp. 1336-1341.

[13] F. Monrose, M. K. Reiter, and S. Wetzel, *Password hardening based on keystroke dynamics*, in Proc. 6th ACM Conf. Computer Security, Singapore, Nov. 1999.

[14] F. W. M. H. Wong, A. S. M. Supian, A. F. Ismail, L. W. Kin, and O. C. Soon, *Enhanced user authentication through typing biometrics with artificial neural network and k-nearest neighbor algorithm* in Conf. Rec. 35th Asilomar Conf. Signals, Syst., comput., Vol. 2, 2001, pp. 911-915.

[15] L. C. F. Araújo, L. H. R. Sucupira Jr., M. G. Lizárraga, L. L. Ling, and J. B. T. Yabu-uti,*User authentication through typing biometrics features*, IEEE Trans. on Signal Processing, vol. 53, No. 2, Feb. 2005.

[16] R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification* Wiley-Interscience Publication, 2nd Edition, Oct. 2000.

[17] L. R. Rabiner, *A tutorial on hidden Markov models and selected applications in speech recognition*. Proc. of IEEE, 77, pp. 257-286, 1989.

[18] Cambridge University Engineering Departament, *The HTK Book*, Cambridge University, 2002.

[19] M. A. T Figueiredo and A. K. Jain, *Unsupervised learning of finite mixture models*, IEEE Trans. Pattern Anal. Machine Intell., vol. 24, no. 3, pp. 381-396, Mar. 2002.