

Keystroke-Based Biometric Authentication in Mobile Devices

E. V. C. Urtiga and E. D. Moreno

Abstract— This paper presents a model for biometric authentication of users based on dynamic typing (keystroke) in mobile devices. In our model, the user accesses the system by typing a string (password) that is monitored in real time. The model captures the time needed for the user to press and release the keys when he is typing and we compare this information with those previously stored into the model. If the result is the same as registered, the user is considered authentic and has access to the system, otherwise the user is regarded as an impostor and has its access blocked by the system. Após a realização de todo o experimento usando o método da distância euclidiana foi possível chegar a uma taxa de. After performing the experiment using the method of Euclidean distance we reached a FAR = 12,97% and FRR = 2,25%.

Keywords— Biometric Authentication, Keystroke, Mobile Devices, Embedded Systems, Euclidean Distance, Cell Phones.

I. INTRODUÇÃO

O CONTROLE de acesso dos sistemas automatizados está cada vez mais importante nos dias atuais e para garantir a segurança nos sistemas computacionais a autenticação de usuários (por meio de uma senha de acesso) tem sido cada vez mais evidente. Contudo, o uso de senhas de acesso pode comprometer a segurança do sistema quando se utilizam senhas frágeis. Porém, o baixo custo e a acessibilidade deste tipo de mecanismo de segurança justificam sua utilização e em várias localidades e aplicações ainda permanece como principal mecanismo de segurança [19, 20].

O objetivo deste trabalho é apresentar o uso da abordagem para o processo de autenticação por senha usando características biométricas em dispositivos móveis, especificamente em um aparelho da Nokia, modelo 5200.

Características biométricas são aquelas observadas em cada ser humano e permitem distinguir uma pessoa da outra e podem ser classificadas em função das características fisiológicas ou características comportamentais [19, 20].

Segundo Costa [1][4] as características fisiológicas devem ser descritas como características estáveis de cada ser humano, tais como, a impressão digital, geometria da mão, padrão de retina e padrão de íris ou características faciais. As características comportamentais são aquelas que refletem o estado emocional do ser humano, que não é considerada estável e que pode ser afetada por fatores externos como estresse, cansaço. Exemplos desta classe são a dinâmica da digitação, reconhecimento da assinatura e voz [5-9].

A característica biométrica utilizada para a elaboração deste trabalho é conhecida como **dinâmica da digitação**, a qual analisa o ritmo de digitação de cada usuário em sistemas computacionais. Esta técnica usa um sistema de monitoramento e captura de amostras, as quais definem o modelo de digitação nesses dispositivos.

Este trabalho focaliza na forma como se usa a dinâmica da digitação, pois utiliza o teclado numérico de um aparelho de telefonia móvel (celular – *cell phone*) como meio de captura das amostras para observar o ritmo da digitação de um usuário quando insere sua respectiva senha nesses dispositivos.

O uso do teclado numérico do aparelho de telefonia móvel resulta num menor intervalo de tempo de digitação e numa melhor precisão dos resultados, pois os usuários seguem o ritmo da digitação com mais naturalidade e o processo é mais uniforme do que em outros teclados numéricos usados por diversos sistemas computacionais. 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 determinado momento.

O restante do trabalho está organizado da seguinte maneira. Na seção 2 são definidos os conceitos básicos de biometria e apresenta as características biométricas mais relevantes. Na seção 3 são apresentadas as abordagens relacionadas à Dinâmica da Digitação. A seção 4 apresenta os métodos utilizados para a realização do trabalho e os resultados obtidos com a metodologia adotada, com respectiva análise. A seção 5 finaliza com as conclusões e algumas propostas para trabalhos futuros.

II. CONCEITOS DE BIOMETRIA

Segundo Araújo [2], a biometria é o ramo da ciência que estuda as características dos seres vivos. Sistemas biométricos são métodos automatizados de verificação e identificação da identidade de um ser vivo e se baseiam nas características fisiológicas ou comportamentais de cada usuário.

Segundo Costa [4], os sistemas biométricos possuem três componentes fundamentais para a sua implementação: (i) mecanismos de captura de sinais que caracterizam um indivíduo, (ii) processamento e classificação dos sinais e (iii) interface capaz de fazer a verificação e a identificação automática dos dados no sistema.

Existem duas maneiras de se classificar os dados de entrada do sistema: um para um e um para muitos. Em sistemas um para um, a informação biométrica fornecida por um indivíduo é comparada com a informação biométrica armazenada na base de dados do indivíduo, ou seja, é feita

E. V. C. Urtiga, Universidade do Estado do Amazonas, Manaus (UEA), Manaus, Brasil, ediminha@gmail.com

E. D. Moreno, Universidade Federal de Sergipe, Aracaju – SE, Brasil, edwdavid@gmail.com

uma verificação se a informação fornecida é a mesma que foi armazenada anteriormente. Em sistemas um para muitos, a informação biométrica fornecida por um indivíduo é comparada com todas as informações que existem de um grupo de indivíduos e está armazenada na base de dados, ou seja, é feita uma identificação da informação fornecida na tentativa de verificar se há igualdade das informações armazenadas para o determinado grupo de indivíduos.

Nesta área devem-se levar em consideração as características fisiológicas e comportamentais dos seres vivos. A característica fisiológica é uma propriedade do corpo do indivíduo que dificilmente irá se modificar, como por exemplo, íris dos olhos, impressão digital, geometria da mão, etc. Já a característica comportamental é uma visão psicológica do estado em que o indivíduo se encontra em um momento, ou seja, a maneira como o indivíduo fala, o modo de o indivíduo digitar ou até mesmo a maneira do indivíduo escrever, entre outros exemplos.

Costa [4] afirmou que para o desenvolvimento dos métodos biométricos de identificação é necessário ter três elementos fundamentais: (1) Posse de um objeto de autenticação; (2) Conhecimento de uma string (senha) de acesso; (3) Uma característica biométrica (seja ela fisiológica ou comportamental).

De posse desses três elementos, pode-se fazer a identificação de várias maneiras, as quais variam de acordo com a complexidade do resultado que se quer alcançar, do custo para realização da identificação, do acesso que o indivíduo terá ao objeto de identificação e do tempo que levará para atingir o resultado [16][18]. Assim, o acesso ao sistema só é permitido se o usuário estiver de posse de um objeto de autenticação, se o indivíduo tiver o conhecimento da senha de acesso e se a característica biométrica apresentada pelo indivíduo for a mesma que o sistema espera que seja apresentada.

É necessário salientar que por mais que o indivíduo possua o objeto de autenticação e tenha o conhecimento da senha de acesso, a identificação não será efetuada se as características biométricas do indivíduo não coincidirem com as cadastradas no sistema. O uso das características biométricas não descarta totalmente os usuais métodos de autenticação pessoal, porém é de grande importância para a qualidade do serviço de identificação.

Métodos Biométricos de Identificação são aqueles utilizados para confirmar a identidade do indivíduo, podendo ser características fisiológicas (face, olhos, mãos, impressão digital) ou comportamentais (assinatura, voz e dinâmica da digitação) [12, 13, 14]. A seguir uma melhor descrição de cada alternativa [1][5][6][10]:

- **Face:** O processo de reconhecimento facial começa na captura da imagem do rosto de um ser humano que é comparado com modelos previamente armazenados numa base de dados. Para uma identificação completa podem ser feitas análises da cor da pele e detecção das linhas de expressão.
- **Olhos:** Os olhos fornecem dois métodos de identificação, o padrão da íris e o padrão dos vasos sanguíneos da retina. Na utilização do padrão da íris não é necessário focalizar os olhos do indivíduo que está sendo testado em um

determinado ponto, pois, pode-se obter a imagem da íris usando um dispositivo de captura. Na utilização do padrão dos vasos sanguíneos da retina uma luz infravermelha de baixa intensidade é direcionada na pupila e na parte posterior do olho então é refletida para a câmera que faz a captura dessa imagem.

- **Mãos:** A autenticação baseada no reconhecimento da geometria da mão resulta de uma análise das características da mão como a forma, o tamanho, as linhas e o comprimento dos dedos.
- **Impressões Digitais:** O resultado das características das impressões digitais consiste na localização da posição de minúsculos pontos (tais como pontos de finalização de linhas e pontos de junção de linhas, presentes nas digitais).
- **Assinaturas:** Existem dois sistemas de reconhecimento da assinatura: dinâmico e estático. Os sistemas dinâmicos utilizam técnicas baseadas nas pequenas diferenças da dinâmica da escrita como, por exemplo, o número de vezes que um usuário levanta a caneta do papel, a pressão com que a caneta é segurada e a velocidade da escrita, entre outros. Os sistemas estáticos utilizam características da imagem da escrita como a inclinação das letras da escrita, a altura da assinatura, o número de palavras.
- **Voz:** Para reconhecer uma pessoa usando sua voz são utilizadas características tais como: entonação, sotaque, velocidade da fala, dentre outros. Como esse tipo de característica é difícil de ser adquirida de maneira automática, então são empregados parâmetros de medidas acústicas do sinal de voz, frequência e energia.
- **Dinâmica da Digitação:** A característica biométrica dinâmica da digitação relaciona-se com a maneira como um indivíduo digita, exemplo um nome e senha de acesso, em um sistema. Características como o intervalo de tempo entre pressionar uma tecla, soltá-la e pressionar a tecla seguinte podem ser extraídas, processadas e armazenadas em um banco de dados; para posterior comparação, quando requisitado, identificando se o usuário é legítimo ou impostor.

Um fator importante em um sistema biométrico é o seu desempenho. Basicamente existem duas métricas utilizadas para medir o **desempenho dos sistemas biométricos** [4, 21]: FAR (Falsa Aceitação) e FRR (Falsa Rejeição) [9][14]:

- **FAR (false acceptance rate)** é definida como a probabilidade de um usuário impostor ser aceito quando solicita autenticação utilizando uma identidade falsa e então o sistema caracteriza o usuário como autêntico.
- **FRR (false rejection rate)** é definida como a probabilidade de um usuário autêntico ser rejeitado quando solicita autenticação no sistema biométrico.

Como se observa na Fig. 1, EER (*Equal Error Rate*) é o ponto de cruzamento entre as curvas de falso resultado negativo e falso resultado positivo, isto é, é o ponto onde se igualam os valores de FAR e FRR. O valor de EER mostra a diferença que o sistema oferece entre os acessos autênticos e os impostores. O ponto EER pode ser obtido a partir de uma curva denominada ROC (*Receiver Operating Characteristic*)

que interliga FRR e FAR, resultando na análise da metodologia de autenticação biométrica [7].

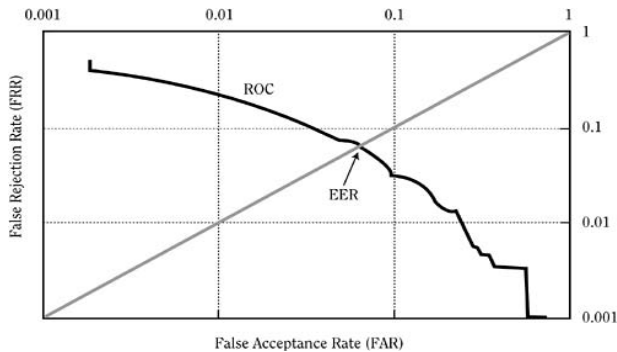


Figura 1. Exemplo de uma Curva de ROC [4]

De modo a facilitar a comparação, a tabela I sintetiza os valores encontrados para FRR e FAR para vários sistemas biométricos, seguindo a experiência de Hwang (2009), Santos (2005) e Clarke (2006) e as referências [4-13].

III. DINÂMICA DA DIGITAÇÃO

A. Definição da Dinâmica da Digitação

Atualmente, o modo mais comum de identificação e autenticação de um usuário em caixas-eletrônicas, computadores, sistemas e até mesmo em aparelhos de celular é a digitação de uma senha.

Quando digita sua senha o usuário executa suas próprias características por meio do ato de “pressionar” e “soltar” uma tecla. Contudo, devido à dinâmica própria desta característica é possível criar um sistema capaz de identificar o usuário através da sua dinâmica de digitação.

Assim, define-se dinâmica da digitação como a maneira de observar e modelar como um usuário digita sua senha em um determinado sistema, levando em consideração o tempo em que “pressiona” e “solta” as teclas, permitindo então a identificação e autenticação usando seu ritmo de digitação [3].

O funcionamento desse sistema baseia-se na verificação de um perfil criado para cada usuário e que deve estar armazenado no sistema. Esse perfil pode registrar a velocidade de digitação que consiste em definir o tempo em que se mantém pressionada uma determinada tecla, o tempo que existe entre o pressionamento e a soltura dela e etc.

TABELA I

COMPARAÇÃO ENTRE OS MÉTODOS BIOMÉTRICOS VIA O FRR VS FAR [19][5]

FAR	Iris	Impressão digital	Face (Masc)	Face (Fem)	Retina	Geometria da mão
0	0,002					
4,48E-08					0,0055	
0,00000417					0,0022	
0,00000338					0,00081	
0,000238					0,00271	
0,0001			0,725	0,71		
0,001			0,91	0,88		
0,01			0,91	0,89		
0,1			0,96	0,95		
0,19		0,19				
0,2	0,0014					
0,33		0,33				
0,41		0,41				
0,6	0,0008					
0,77		0,77				
0,92		0,92				
0,99		0,99				
5,29						8,34

Segundo [1][4][5][19], um sistema baseado no reconhecimento da dinâmica da digitação pode gerar resultados bastante satisfatórios, já que as características citadas acima são de difícil imitação por usuários impostores.

A tabela II apresenta alguns resultados usando dinâmica da digitação. Conforme observado também por Peacock (2004) e Santos (2005), os valores de FAR variam de 0% até 50%, e o FRR de 25% a menos de 1%, e o número de pessoas que ajudaram nos testes está entre 10 e 100.

TABELA II
FAR E FRR NA DINÂMICA DA DIGITAÇÃO

Autor	FAR (%)	FRR (%)	Usuários
Monrose [17]	0	4	-
Joyce [23]	0.25	16.67	33
Cho [24]	0.0	1.0	21
Cavalcanti [25]	0.0	6.04	
Cavalcanti [25]	0.0	30.99	
Sang 1 classe [26]	10	-	15
Sang 2 classes [26]	20	-	15
Modi [27]	0.33	94.87	
Pavezi 5 repetições [28]	10	26	20
Pavezi 10 repetições [28]	16	30	20
Costa [1][4]	2.1	1.2	25

B. Definições usadas nos Modelos da Dinâmica da Digitação

A seguir alguns termos úteis nos modelos da dinâmica da digitação:

- **String:** é definida como o conjunto de caracteres digitados pelo usuário. É uma informação monitorada pelo sistema de autenticação que a coleta, verifica e autoriza.
- **Número de amostras:** Fator importante para o resultado final que varia de acordo com o trabalho adotado.
- **Extração das características:** Para a representação final da amostra, as características armazenadas no sistema devem ser repetitivas para cada usuário e diferente para os demais usuários.
- **Latência da Digitação:** Representa o intervalo de tempo entre a digitação de teclas sucessivas (t_n e t_{n+1}). Essa característica é extraída calculando a diferença entre o pressionamento das teclas t_n e t_{n+1} , gerando um valor positivo, e calculando a diferença entre o pressionamento da tecla t_{n+1} e a soltura da tecla t_n o que resulta em um valor negativo, pois irá depender da maneira como cada usuário digita.
- **Duração da Digitação:** Representa o intervalo de tempo entre o pressionamento e a soltura de uma mesma tecla.
- **Pressão da Tecla:** Representa o nível de pressão aplicada por cada usuário a uma determinada tecla.
- **Precisão do Tempo:** A precisão do tempo indica a medida exata de captura das amostras digitadas. Assim, quanto maior for a precisão do tempo, maior a quantidade de acertos na captura.
- **Tentativas de autenticação:** Para que a autenticação seja completa é necessário escolher a menor latência entre as duas amostras coletadas. Os valores de FRR e FAR também são essenciais para a autenticação do usuário cadastrado no sistema.
- **Mecanismo de adaptação:** O mecanismo de adaptação

deve manter o ritmo de digitação do usuário para que não haja interferências na determinação do resultado.

- **Classificador:** O classificador é o responsável pelo processo de autenticação do usuário, é capaz de verificar e identificar, validando ou não a entrada do usuário em um determinado sistema. Existem diversos tipos de classificadores: estáticos, baseados em redes neurais e os nebulosos. Para um melhor resultado, o classificador deve estar de acordo com o sistema.
- **Distância Euclidiana:** Um método de classificação na dinâmica da digitação é a distância euclidiana, que entre dois vetores n -dimensionais x e y é definida como o escalar:

$$d_e(x, y) = \|x - y\| = \|y - x\| = \left[(x_1 - y_1)^2 + \dots + (x_n - y_n)^2 \right]^{1/2}$$

A seguir apresentamos uma metodologia aplicada para a realização de uma autenticação biométrica baseada na dinâmica da digitação.

Quando a senha é digitada no sistema, o sistema constrói um vetor. Esse vetor é composto de posições que representam a unidade de tempo calculada em função de “pressionar” e “soltar” as teclas.

Para dar início à autenticação, o usuário digita uma *string* (senha), com isso capturam-se as informações dos tempos “pressionar” e “soltar” teclas. Com essas informações obtém-se uma amostra. Para um melhor resultado, uma série de amostras é armazenada no sistema, o que permite gerar um modelo “ideal” para cada usuário. Com as amostras calculadas, o resultado é enviado a um classificador que é o responsável pela autorização de acesso do usuário ao sistema através da identificação e autenticação da sua identidade.

Com base na teoria pesquisada na área da dinâmica da digitação, aplicamos as seguintes características:

- A *string* deverá conter oito caracteres, sendo todos compostos de números e colhidos no espaço que localiza o teclado do aparelho de telefonia móvel onde se encontram somente teclas numéricas.
- Em nosso trabalho analisamos várias características, entre as quais estão as latências entre teclas e a duração de pressionamento e soltura de cada tecla.
- Outro aspecto analisado foi a utilização de apenas uma mão para a digitação de cada usuário, o que facilitou a pesquisa, uma vez que diminuiu o tempo que cada usuário precisou para coleta de suas amostras.
- Em conjunto com o aparelho de telefonia móvel, foi utilizado um computador, onde estava armazenada a base de dados das coletas dos usuários.
- Outro recurso utilizado para facilitar a coleta das amostras foi a tecnologia *bluetooth*, que permitiu ligar um fone *bluetooth* diretamente ao aparelho de telefonia móvel, que coletava os sons emitidos por cada tecla, enviando-os à base de dados por meio de um microfone ligado ao computador.

Nos testes de validação do modelo, trabalhamos com um grupo de usuários com características diferentes entre si,

formado por homens e mulheres com idades variando entre 14 e 55 anos. Na necessidade de um espaço amostral comum aos usuários, escolheu-se a amostragem de convivência para formar a base de dados do trabalho de autenticação biométrica baseada na dinâmica da digitação em aparelhos de telefonia móvel.

Os elementos que formam a nossa base de dados são as informações digitadas por cada usuário, em conjunto com as características extraídas de cada caractere que compõe a *string* digitada pelo usuário. Para composição das amostras, foi utilizado o teclado numérico do aparelho de telefonia móvel (o aparelho Nokia 5200), digitou-se uma *string* composta por oito caracteres, a qual foi armazenada no banco de dados. Para obtenção dos resultados, foram coletadas amostras de todos os usuários do sistema.

O método utilizado para obtenção do limiar de cada usuário foi a Distância Euclidiana, que consiste em calcular de forma simples a distância entre dois vetores. Depois de encontrado os limiares de cada usuário autêntico utilizaram-se as amostras dos usuários impostores para a obtenção das taxas de FAR (Falsa Aceitação) e as amostras dos usuários autênticos para obter as taxas de FRR (Falsa Rejeição) para que, finalmente, pudéssemos determinar a curva de ROC que caracteriza a dinâmica da digitação de cada usuário.

A Fig. 2 mostra o processo de cadastro e autenticação de um usuário via dinâmica da digitação, proposto inicialmente por Costa [1][4]. O sistema pode ser visto em duas fases: a primeira, ver parte superior da Figura 2, refere-se ao cadastramento do usuário no sistema, enquanto que o segundo, ver parte inferior da Fig. 2, corresponde à autenticação de um usuário.

Inicialmente, um usuário informa a sua senha para a respectiva extração das características, se o número de amostras digitadas for suficiente, em nossos testes foi igual a 50, é composto o modelo do usuário no Banco de Dados.

Os dados são armazenados, e o modelo “ideal” é obtido com o cálculo dos tempos de latência entre cada dois caracteres digitados e o tempo em que cada tecla fica pressionada, assim, se capturam todos os tempos da digitação pertencente ao conjunto de caracteres que forma a senha de acesso da digitação de cada usuário. De posse desse modelo, é feita a segunda fase, a autenticação do usuário.

O classificador determina se o usuário é autêntico ou impostor. Neste trabalho usamos um método simples, a distância euclidiana.

Conforme se observa na Fig. 2, o processo de autenticação é do ponto de vista do usuário, igual à autenticação em qualquer sistema que solicite uma senha de acesso, tendo apenas que digitá-la quando solicitado.



Figura 2. Fluxograma do Processo de Cadastramento e Autenticação de Senhas via Dinâmica da Digitação [1][4]

O sistema compara os tempos de latência propostos (recolhidos na dinâmica da digitação) com os armazenados anteriormente na base de dados, classificando como autênticos se e somente se o valor for menor que o limiar encontrado, senão o usuário é classificado como impostor e é bloqueado pelo sistema.

Um ponto de destaque relaciona-se ao número de amostras necessárias para se obter uma boa precisão. Costa em [4] fez uma análise mudando o número de amostras para 10, 15, 20, 25 e 30 usuários, e obteve que o ERR foi de 10, 7.8, 5.9, 4.6 e 5, respectivamente, e então concluiu que para o caso analisado por ele o número com menor ERR, de 4.6, foi obtido para amostras = 25. Em nosso caso, usamos 50, pois com esse valor se obteve uma confiança de 95%, não obstante, se faz necessário mais pesquisa para obter o erro em dispositivos móveis quando se muda o número de amostras.

IV. EXPERIÊNCIA E RESULTADOS

Na análise da metodologia adotada usamos a classificação por distância euclidiana para a identificação dos usuários, verificou-se que o classificador é o responsável em determinar a autenticidade do usuário. Assim, para a realização da autenticação, podem ocorrer quatro situações na realização da autenticação, a saber:

1. O usuário é autêntico e o classificador o descreve como autêntico.
2. O usuário é impostor e o classificador o descreve como impostor.
3. O usuário é impostor e o classificador o descreve como autêntico.
4. O usuário é autêntico e o classificador o descreve como impostor.

Nas situações 3 e 4 o classificador comete erros de classificação, ou seja, descrever um usuário impostor como autêntico e um usuário autêntico como impostor,

respectivamente. Nesse caso, o desempenho é medido pelas taxas de FAR e FRR, que são calculadas da seguinte maneira:

$$FAR = \frac{\text{Quantidade de testes enquadradas na situação 3}}{\text{Quantidade total de testes}} (\%)$$

$$FRR = \frac{\text{Quantidade de testes enquadradas na situação 4}}{\text{Quantidade total de testes}} (\%)$$

Em nossos experimentos, o grupo voluntário de participantes foi composto por 20 pessoas, das quais 10 pessoas eram do sexo masculino e 10 pessoas eram do sexo feminino, e com idade variando de 14 a 55 anos e, com diferentes níveis de conhecimento na manipulação do aparelho de telefonia móvel.

Neste trabalho, 15 participantes representam os utilizadores autênticos que são atacados por todos os demais participantes que não estavam autorizados a entrar no sistema, ou seja, usuários impostores, em nosso estudo de caso são 5.

Os dados solicitados no momento do cadastro foram: nome, idade, informação se o participante possuía familiaridade com qualquer aparelho de telefonia móvel, e a data do nascimento no formato [d d m m a a] composta por 8 caracteres, que foi utilizada como senha.

Este valor para a senha foi determinado pensando na facilidade em lembrar essa informação, além desse ser um dado “secreto” para cada usuário, aumentando ainda mais a segurança. Soma-se a isso o grau de aproximação da senha com as características na própria digitação, fazendo com que usuários impostores sejam então bloqueados pelo sistema.

Outro ponto relevante com relação ao tamanho da senha vem da pesquisa feita por Costa [1][4], que mostrou que a taxa de ERR melhora quando se aumenta o número de caracteres na senha. Em [4], ele mostrou o ERR para senhas mudando de 1 até 8 caracteres, e o ERR diminui respectivamente de 16 para 3.8. Senhas pequenas mantêm um alto ERR, e este cai consideravelmente após 4 caracteres [1][4][22]. Apesar dessas

observações seria interessante fazer uma análise mais completa mudando o número de caracteres na senha usada em dispositivos móveis e definir um valor adequado que satisfaz precisão do método com facilidade dada aos usuários, pois senhas grandes aumentam a precisão, mas causam desconforto e podem ser facilmente esquecidas pelos usuários.

Com relação ao número de amostras, cada participante digitou sua senha 50 vezes, constituindo 750 amostras de usuários autênticos (são 15 participantes) e 250 amostras de usuários impostores (pois são 5 impostores), o que resultou em 1000 amostras cadastradas no sistema. Os resultados foram descritos da seguinte forma e são sumarizados na Tabela III:

TABELA III
RESULTADO DAS AMOSTRAS COLETADAS DE CADA USUÁRIO AUTÊNTICO

Usuários	Limiar (α)	FAR	FRR
1	0,2468	0,1333	0,0044
2	0,3625	0,1556	0,0000
3	0,2090	0,1600	0,0133
4	0,1265	0,1956	0,0178
5	0,2124	0,1511	0,0533
6	0,1109	0,2044	0,0000
7	0,2514	0,0311	0,0267
8	0,3375	0,0622	0,0622
9	0,0990	0,1111	0,0000
10	0,1187	0,1200	0,0000
11	0,1122	0,1600	0,0000
12	0,2775	0,1467	0,0533
13	0,0975	0,1111	0,0489
14	0,1965	0,0933	0,0489
15	0,1892	0,1111	0,0089

- O número de participante mostrado nos resultados foi o número de participantes caracterizados como autênticos, ou seja, 15. Os demais foram utilizados para tentar “quebrar” a senha dos usuários autênticos.
- Os dados dos limiares foram determinados usando a média dos tempos de latência de 80% das amostras digitadas por cada usuário.
- Os 20% das amostras restantes foram utilizadas para a comparação e determinação das taxas de Falsa Rejeição (FRR).
- Com 100% dos dados das amostras dos usuários impostores e as latências médias de cada usuário autêntico, foram determinadas as taxas de Falsa Aceitação (FAR).

Após a realização de todo o experimento se obtém os valores médios, e foi possível obter que o Limiar médio = 19.6%, a taxa de FAR = 0.1297 (ou 12.97%) e uma taxa de FRR = 0.0225 (ou 2.25%).

As taxas obtidas são compatíveis com aqueles obtidos em outros trabalhos relacionados à dinâmica da digitação, em teclados numéricos [1-5] e [13-17], conforme mostrado na tabela II.

Assim, foi possível ter as seguintes observações:

- O classificador é ponto essencial para gerar bons resultados quando aplicado à dinâmica da digitação em teclados de aparelhos de telefonia móvel. Os experimentos foram conduzidos usando a Distância Euclidiana, que oferece um modelo simples e de fácil entendimento;
- A combinação da *string* e das características

“pressionar-pressionar”, “pressionar-soltar”, “soltar-pressionar”, resulta em um bom resultado sobre a maneira com que cada usuário digita no aparelho de telefonia móvel;

- A escolha de uma *string* é importante, pois permite que o usuário tenha familiaridade com a senha utilizada e possibilita que os impostores sejam recusados pelo sistema. Adotou-se um valor de 8 caracteres para *string*, representando a data de nascimento do usuário, por se tratar de um número de fácil memorização para o usuário e um número com menores chances de ser quebrado pelos usuários impostores. Foi possível observar que aspectos como familiaridade do usuário com a *string* (senha) a ser digitada, o número de amostras coletadas para cada usuário, os mecanismos de adaptação do modelo do sistema e a precisão desejada dos dados são essenciais para se chegar ao objetivo desejado.
- A quantidade de amostras é essencial na escolha da metodologia adotada. Neste trabalho utilizamos uma quantidade de 50 amostras para compor o modelo, e tivemos uma confiança de 95%.
- O mecanismo de adaptação deve manter o ritmo de digitação do usuário para que não haja interferências na determinação do resultado.
- A determinação do limiar permite que o sistema determine e identifique o usuário como autêntico ou impostor, de acordo com as características cadastradas no ato da entrada no sistema.

As Fig. 3-6 mostram o comportamento de FAR e FRR dos usuários classificados por idade e sexo. Ao lado de cada figura aparece o valor de α (valor que caracteriza cada usuário, o qual foi calculado pelo classificador – usando o modelo da Distância Euclidiana) que utilizou todas as amostras armazenadas para cada usuário.

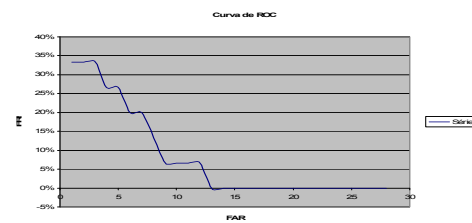


Figura 3. $\alpha = 0,3625$, Masculino, Idade = 35 a 55

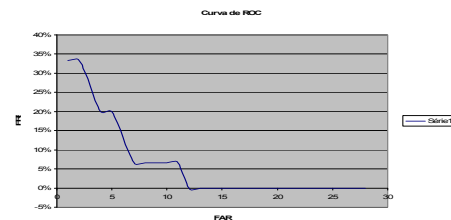


Figura 4. $\alpha = 0,1187$, Feminino, Idade = 35 a 55

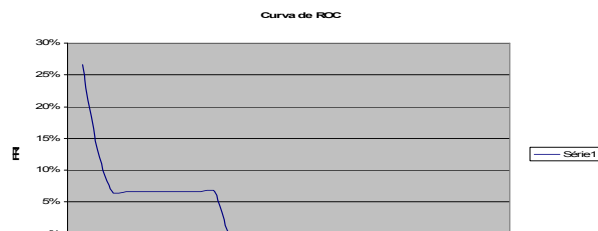
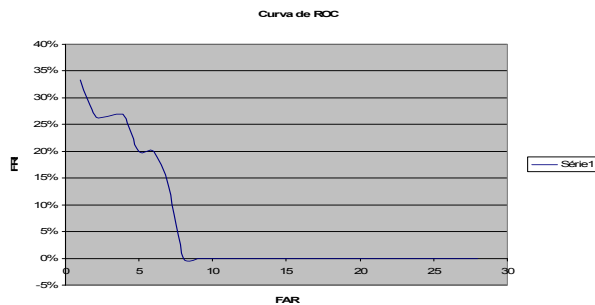


Figura 5. $\alpha = 0,3375$, Masculino, Idade = 15 a 30Figura 6. $\alpha = 0,1965$, Feminino, Idade = 15 a 30

As Fig. 3 e Fig. 4 são representativas do ritmo de digitação do sexo masculino e feminino com idade entre 35 a 55 anos. Em nossos experimentos notamos que os usuários obtiveram um gráfico parecido e com intervalos bem definidos. Por serem usuários que possuem pouca familiaridade com o aparelho, mostram a preocupação do usuário em digitar cuidadosamente sua senha no aparelho de telefonia móvel.

As Fig. 5 e 6 mostram o comportamento para pessoas com idade entre 15 a 30 anos, para masculino e feminino respectivamente. Nota-se que os usuários obtiveram um gráfico parecido com curvas quase uniformes, com intervalos mais longos, devido à maior familiaridade com o aparelho.

Interessante perceber que foi possível detectar algumas mudanças da curva ROC para diferentes idades dos usuários, similar ao obtido em Hwang (2009)[30], que fez um trabalho similar, mas com 25 pessoas com idade entre 22 e 33 anos, e senhas de 4 dígitos, e obteve um ERR entre 4 e 16%.

Nossos resultados são preliminares e destacamos que é necessário realizar mais pesquisas e testes para chegar a afirmações concretas, como por exemplo, que o método oferece menos precisão quando os usuários são idosos, que é necessário aumentar o tamanho da senha, ou usar senhas com valores alfanuméricos para aumentar a segurança, assim como verificar o impacto no número de amostras na precisão e conforto dos usuários, assim como incluir outros métodos nos classificadores e realizar respectivas análises um melhor entendimento da dinâmica da digitação em dispositivos móveis.

V. CONCLUSÕES E TRABALHOS FUTUROS

Cada método de autenticação biométrica tem suas particularidades. Escolhemos a dinâmica da digitação por se tratar de uma área em expansão. A autenticação biométrica

baseada na dinâmica da digitação em aparelhos de telefonia móvel pode aumentar a segurança nos sistemas de informação que usam dispositivos móveis.

Este trabalho apresentou e verificou que a autenticação biométrica por meio das características obtidas usando um modelo da dinâmica da digitação dos usuários pode efetivamente melhorar o controle de acesso aos aparelhos de telefonia móvel aumentando a segurança.

Sendo a dinâmica da digitação uma característica pessoal e intransferível podemos verificar que por mais que um “impostor” saiba a senha de um usuário autêntico, o sistema poderá bloquear o acesso, pois verificará que o modo como o usuário digita a senha é diferente do modo cadastrado no sistema.

Considerando as limitações dos aparelhos de telefonia móvel utilizamos um computador para armazenar o sistema de base de dados dos usuários “ligado” ao aparelho de telefonia móvel por meio de um microfone que captura a frequência do som que cada tecla do aparelho de telefonia móvel possui caracterizando a tecla digitada e assim identificando e autenticando o usuário quando este for autêntico e recusando qualquer usuário impostor que tente acessar o sistema.

Nos testes contamos com a colaboração de 20 pessoas. Cada pessoa fazia o seu cadastro no sistema e usamos 50 amostras, ou seja, cada usuário digitava sua *string* (senha) 50 vezes, para que pudesse ser capturado o tempo enquanto o usuário “pressionava” e “soltava” as teclas. Desses 20, 15 usuários estavam caracterizados como usuários autênticos e 5 usuários estavam caracterizados como usuários impostores, que tinham como objetivo “tentar quebrar” a senha do usuário autêntico. Após a captura das amostras fomos ao processo de aquisição do modelo ideal de cada usuário, ou seja, foi necessário determinar um limiar que caracterizava o usuário diferenciando-o dos demais usuários. Para isso foi utilizada a Distância Euclidiana que calcula a distância entre os vetores para obtenção do limiar.

De posse dos dados, fizemos comparações entre os limiares obtidos pelos usuários autênticos e os limiares obtidos pelos usuários impostores, e obtivemos um FAR (Falsa Aceitação) = 12,95% e FRR (Falsa Rejeição) = 2,25%.

Como trabalhos futuros é possível visualizar os seguintes:

- Obter um novo método de determinar o conjunto de limiares que irão identificar os usuários utilizando outros tipos de classificadores, tais como, redes neurais [8][11].
- Analisar o impacto quando se aumenta a quantidade de caracteres que compõe a senha, e verificar se aumenta a dificuldade de “quebra” da senha por usuários impostores.
- Analisar o número de amostras coletadas, e verificar o impacto na diminuindo o tempo de aquisição das amostras com uma coleta de dados menos cansativa para os participantes.
- Verificar o impacto da autenticação biométrica baseada na dinâmica da digitação em outros tipos de teclados, como por exemplo, outros dispositivos móveis e caixas eletrônicos, observando a eficiência da técnica baseada em requisitos.

REFERÊNCIAS

- [1] COSTA, R. N.; YARED, G.F.; RODRIGUES, R.N.; YABU-UTI, J.B.T.; VIOLARO, F.; LING, L.L. Autenticação Biométrica via Dinâmica da Digitação em Teclados Numéricos. *Anais do XXII Simpósio Brasileiro de Telecomunicações - SBrT'05*, Campinas, Brasil, 2005.
- [2] ARAÚJO, Livia Cristina Freire. "Uma metodologia para autenticação pessoal baseada em dinâmica da digitação". Campinas, SP. Dissertação de Mestrado – UNICAMP, Fevereiro 2004.
- [3] GAVA, Érika Aparecida. "Sistemas biométricos com ênfase na técnica da dinâmica da digitação". Trabalho de Conclusão de Curso–Faculdade de Tecnologia de Americana, Americana, SP. Dezembro 2006.
- [4] COSTA, Carlos R. do Nascimento. "Autenticação Biométrica via Teclado Numérico Baseada na Dinâmica da Digitação: Experimentos e Resultados". Campinas, SP. Dissertação de Mestrado – Faculdade de Engenharia Elétrica e de Computação, Dezembro 2006.
- [5] SANTOS, Henrique Manuel Dinis dos. "Estudo dos padrões de digitação e sua aplicação na autenticação biométrica". Portugal. Dissertação de Mestrado – Universidade do Minho, 2005.
- [6] ASSIS, Miriam. "Sistemas biométricos aplicados à telefonia móvel celular" Trabalho de Conclusão de Curso. Americana, SP. 2004.
- [7] CLARKE, N. L. "Authenticating mobile phone users using keystroke analysis". *International Journal of Information Security*, Volume 6 , Issue 1 Pages: 1 – 14, December 2006.
- [8] SANTOS E SILVA, Daniel F. dos; ABE, Narumi. Sistema Reconhecedor de Impressões Digitais Utilizando Redes Neurais. Projeto de Eng. de Graduação), Univ. Católica Dom Bosco, Campo Grande. 2002.
- [9] CAMPESTRINI, Jonas Rafael; AUTENTICAÇÃO SEGURA UTILIZANDO BIOMETRIA. Mestrado Profissional em Sistemas de Informação), Univ. da Região de Joinville - UNIVILLE, 2003.
- [10] FERRARI, M. J; MACOPPI, M. A. F. Análise de viabilidade de características biométricas como alternativa para a segurança na Internet. São Paulo: HBTec, 2001.
- [11] S. Hayken. *Redes Neurais Artificiais: Princípios e Prática*. Editora Bookman, 2001.
- [12] Wikipedia. Biometrics – Wikipedia, the free encyclopedia. <http://en.wikipedia.org/wiki/Biometric>, 2007. Acessado em 10/2009.
- [13] J. Leggett et all. Dynamic identity verification via keystroke characteristics. *Intl. Journal of Man-Machine Studies*, 35:859–870, 1990.
- [14] D. Umphress and G. Williams. Identity verification through keystroke characteristics. *Intl. Journal of Man-Machine Studies*, 23:263–273, 1985.
- [15] J. Leggett and G. Williams. Verify identity via keystroke characteristics. *International Journal of Man-Machine Studies*, 28:67–76, 1987.
- [16] BioPassword INC. BioPassword Family of Products. <http://www.biopassword.com>, 2005. Acessado em Set. 2009.
- [17] MONROSE, F.; MONROSE, F.; RUBIN, A. Keystroke dynamics as a biometric for authentication. In *Journal FGCS - Future Generation Computing Systems – Security on the WEB*, p.351-359, 2000. Em <http://www.caslon.com.au/biometricsnote6.htm#keystroke.htm>.
- [18] TEIXEIRA, Sérgio Ricardo. Projetar de forma a receber tecnologias atuais e futuras. [on line] Disponível na internet via web URL: <http://www.portaldaaautomacao.com.br>. Acessado em setembro 2009.
- [19] HWANG, Seong-seob et Al. Improving authentication accuracy using artificial rhythms and cues for keystroke dynamics-based authentication. In *Elsevier Journal - Expert Systems with Applications* No. 36, 2009.
- [20] BHATNAGAR, Jay; KUMAR, Ajay. On estimating performance indices for biometric identification. In *Elsevier Journal - Pattern Recognition* No. 42, p. 1803 – 1815, 2009.
- [21] SU, Ching-Liang. Palm-print recognition by matrix discriminator. In *Elsevier Journal - Expert Systems with Applications* No. 36, 2009.
- [22] ORD, T. and Furnell, S.M. *User authentication for keypad-based devices using keystroke analysis*. In *proc. of Second International Network Conference (INC)*, 2000.
- [23] JOYCE, R.; GUPTA, G. Identity Authentication based on Keystroke latencies. In *Communications of the ACM* Vol. 33, p. 168-176, 1990.
- [24] CHO, Sungzoon et Al. Web-based keystroke dynamics identity verification using neural network. In *Journal of Organizational Computing and Electronic Commerce* Vol 10, No. 4, p. 295-307, 2000.
- [25] CAVALCANTI, G. et Al. Um Sistema de verificação de identidade pessoal através da dinâmica da digitação. In *XXV Congresso da SBC – Sociedade Brasileira de Computação*, 2005.
- [26] SANG, Yingpeng.; SHEN, Hong; FAN, Pingzhi. Novel impostors detection in keystroke dynamics by support vector machine. In *Proc. Of the Intl. Conference on Parallel and Distributed Computing – Applications and Technologies (PDCAT)*, p. 666-669, 2004.
- [27] MODI, S.K.; ELLIOT, S.J. keystroke dynamics verification using a spontaneously generated password. In *Proc. Of the Intl. Carnahan Conference on Security Technology (ICCST)*, U.S.A., 2006.
- [28] PAVEZI, R.S. et Al. Dinâmica da digitação aplicada a ambientes WEB. Laboratório de telemedicina na Universidade Federal de Santa Catarina (UFSC), Brasil, 2009.
- [29] PEACOCK, A.; KE, X. and WILKERSON, M: *Typing Patterns: A Key to User Identification*, IEEE Security and Privacy, September/October 2004.
- [30] HWANG, Seong-Seob; CHO, Sungzoon; PARK, Sunghoon. Keystroke dynamic-based authentication for mobile devices. In *Journal of Elsevier - Computers & Security* 28, p. 85-93, 2009.

Edma Valléria Cunha Artiga, She is Bachellor in Computer Engineering, from UEA - Universidade do Estado do Amazonas, Manaus, AM, Brasil. She is interested in computer architecture, embedded systems, biometrics researches and security.



Edward David Moreno Ordoñez, Received the MSc. and PhD degrees in Electrical Engineering from University of Sao Paulo - Brazil, in 1994 and 1998. During 1996 and 1997 he stayed as invited researcher at University of Toronto, Canada, and Chalmers University of Technology, Sweden. He is professor. at the UFS (Federal University of Sergipe, Aracaju, SE, Brazil. Prof. Moreno has participated on 100 events as International Program Committee and he is editorial board of 4 important Journals: JUCS, IJCSNS, JCP, and Springer - Transactions on Computational Science. He has published five books about digital systems, FPGAs, Microcontrollers, Reconfigurable Computing and Hardware Security. The research areas are: computer architecture, reconfigurable computing, embedded systems, hardware security, power aware computing and performance evaluation.