**Autenticação biométrica contínua com dinâmica de toque em aplicação bancária de dispositivo móvel**

Júlio César de Sousa Cabral\*;Thiago Gentil Ramires

\*autor correspondente: jcscabral@gmail.com

**Autenticação biométrica contínua com dinâmica de toque em aplicação bancária de dispositivo móvel**

**Resumo**

**Palavras-chave:** autenticação contínua; dinâmica de toque; biometria comportamental; autenticação passiva; aplicação bancária

**Introdução**

O uso de aplicativos instalados em dispositivos móveis tem sido um desafio para a área de segurança da informação. A sensibilidade de sistemas específicos como o internet banking demanda contínuo aperfeiçoamento da verificação do usuário, sobretudo com a crescente disponibilidade de acesso aos dados biométricos coletados por sensores dos dispositivos.

O emprego da biometria na identificação de usuários de sistemas tem se mostrado eficiente em diversas aplicações. Essa tecnologia se fundamenta na particularidade das características do indivíduo, sejam fisiológicas ou comportamentais. A biometria fisiológica é medida diretamente de parte do corpo humano, como o reconhecimento de leitura facial, de impressão digital e de íris. Por outro lado, a biometria comportamental é obtida de forma indireta, passiva, com a obtenção de dados peculiares do comportamento durante a ação de um usuário (Nanavati et al, 2002).

Os sistemas biométricos são classificados em sistemas de identificação ou de verificação (Nanavati et al, 2002). A identificação envolve a determinação da identidade de uma pessoa a partir de um banco de dados sem o prévio conhecimento de qual pessoa se trata. Dessa forma, o processo de validação precisa percorrer toda a base de dados em busca de uma correspondência. O sistema de verificação, por sua vez, almeja identificar a correspondência de uma pessoa com a identidade de um determinado indivíduo já conhecido. A rigor, esse tipo de comparação 1:1 (um para um) relacionando um nome de usuário a dados é chamado de autenticação (Nanavati et al, 2002). Em dispositivos móveis, o conceito de autenticação contínua se refere à certificação da identidade de um usuário durante a navegação no sistema em uma sessão inteira (Georgiev et al. 2022a).

Na biometria comportamental, o termo genérico dinâmica de toque (touch dynamics) compreende a digitação de teclas (keystroke dynamics) e o toque de gesto (touch gesture). A tecnologia da dinâmica de digitação utiliza a maneira distinta de uma pessoa digitar padrões para realização de verificação mediante dados do intervalo de tempo que envolve o acionamento das teclas (Nanavati et al, 2002). Em contextos de digitação mais prolongada, também é possível avaliar parâmetros como a frequência de caracteres, uso de símbolos especiais, quantidade de acionamento de tecla de conserto de erro digitado (Kolakowsa, 2020). Em sistemas de autentificação por pressionamento de teclas, oito caracteres do nome do usuário e da senha são o tamanho mínimo para de dados coletados (Nanavati et al, 2002).

A dinâmica de gesto foi inicialmente inspirada no movimento do mouse dos computadores de mesa (Georgiev, 2022a) e no contexto da mobilidade refere-se ao processo de medir e acessar o ritmo do toque humano nas telas de dispositivos. Os dados das características do toque como tamanho, pressão e coordenadas espaciais podem ser considerados uma assinatura digital do usuário (Teh et al, 2016). Dados de movimento coletados por sensores como acelerômetro, giroscópio e magnetômetro também podem ser considerados características particulares adicionais na entrada de construção de modelos (Kolakowsa et al, 2020).

A viabilidade da autenticação contínua em aplicação em dispositivos móveis por meio de biometria comportamental tem sido objeto de diversos estudos (Vaishnav et al, 2022). Esses trabalhos consideram os dados de um ou mais sensores como variáveis explicativas para o fenômeno da classificação de um usuário com diferentes metodologias. No contexto bancário, uma notável proposta com as modalidades dinâmica de pressionamento de tecla e movimentos do telefone teve como princípio a maneira o usuário a informa uma senha padronizada de oito dígitos (Buriro et al, 2017). Foram usados os classificadores Naive Bnz ayes, Rede Neural e Floresta Aleatória (Random Forrest) em situações com resultados melhores à medida que a amostra de treinamento crescia.

Outra proposta de autenticação contínua em aplicação bancária considerou os dados de texto digitados nesse contexto limitados a ponto de serem desprezados (Incel et al, 2021). Esse sistema coletou não apenas os dados dos gestos de toque na tela, mas também do acelerômetro, giroscópio e magnetômetro. O estudo se valeu apenas dos eventos de rolagem por serem os mais comuns no aplicativo original. Os participantes realizaram cinco ações mais frequentes no aplicativo modificado em três posturas, totalizando 15 sessões por usuário.

Os dois estudos citados voltados para autenticação de sistemas bancários, com dinâmica de digitação e de gestos fornecem direções conceituais, metodológicas e de desempenho para uma nova abordagem que considera os dados de ambos procedimentos. O uso de senha alfanumérica ainda é um meio comum de autenticação de usuários em aplicativos de bancos no Brasil, mesmo com a disponibilidade de outras modalidades. Embora a quantidade de caracteres possa variar conforme a instituição, uma avaliação em condições restritas pode ajudar na compreensão do desempenho de um sistema de segurança que recebe dados oriundos da digitação e a da dos gestos.

Espera-se que um usuário intruso em um sistema se comporte de maneira diferente do usual. A existência de dados modelados considerados normais faz que uma anomalia seja reconhecida como um desvio suficiente desse padrão (Aggarwal, 2017). Virtualmente, todos os algoritmos de detecção de anomalia criam um modelo do padrão normal dos dados e computa uma pontuação (score) do desvio de um determinado dado comparado. A avaliação do modelo é feita pela qualidade do ajuste entre o dado e o modelo. Dessa forma é possível identificar qual metodologia analisada apresenta melhor desempenho final.

**Material e Métodos**

Devido à dificuldade de disponibilidade dos dados específicos a serem estudados, foi realizada pesquisa de campo para coletar as informações geradas durante o uso de aplicativo bancário. Como não foi possível utilizar software de um banco real no dispositivo conforme realizado na produção do modelo Dakota (Incel et al, 2021), foi desenvolvido um aplicativo simulador de ações presentes no contexto bancário de um hipotético banco chamado “Meu Banco”.

O sistema operacional Android foi escolhido por questão de facilidade de desenvolvimento e de ser o mais presente nos celulares no Brasil. A API atual utiliza lingagem Kotlin com o paradigma de desenvolvimento *Jetpack Compose.* Todos os dados foram armazenados localmente, no próprio celular. A biblioteca SQLite foi usada no controle das informações dos pesquisados, enquanto que os registros dos eventos das ações foram gravados em arquivos no formato csv.

Apenas um aparelho celular foi manipulado na coleta de dados, o modelo Galaxy A30 da marca Samsung, provido da 11 versão do Android, com interface de usuário One UI 3.1. e tela de 6.4 polegadas. Com a restrição de um único dispositivo, buscou-se evitar variações nos resultados decorrentes das características dos aparelhos e, ainda, considerar que, de maneira geral, um invasor usa o mesmo celular da vítima em um ataque (Georgiev et al, 2022b).

Os dados foram coletados por meio de pesquisa de campo em ambiente coorporativo de um banco público e na praça de alimentação de um shopping center da cidade de Jundiaí, São Paulo. Em todos os lugares o participante permaneceu na mesma postura durante a pesquisa, sentado em uma cadeira diante de uma mesa. Dessa forma, cada participante não apenas ficou livre para usar a mesa como suporte auxiliar se fosse julgasse conveniente, mas também foi controlado o efeito da postura que impacta diretamente nos resultados (Syed et al, 2019). A postura pode ser definida como a combinação da orientação do aparelho (retrato ou paisagem) e a posição de uso (seguro em mão ou sobre uma mesa) (Syed et al, 2019). A tela foi fixada na orientação retrato que espelha o funcionamento dos aplicativos bancários.

A coleta foi dividida em momento de cadastro, treinamento e teste. Inicialmente, o pesquisador explicou de maneira informal todo o procedimento da pesquisa, auxiliado por uma folha impressa entregue ao participante (anexo 1), fornecendo os esclarecimentos necessários acerca das tarefas solicitadas. Na fase de treino os dados inseridos não foram validados, pois serviu apenas para ajudar o participante a enetender o funcionamento da aplicação. No momento de teste, quando os dados foram efetivamente gravados, a interferência no procedimento foi interrompida.

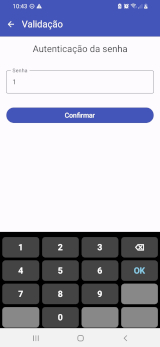
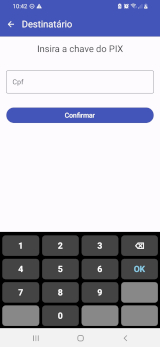
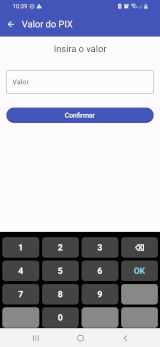
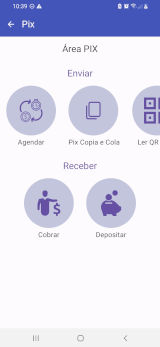
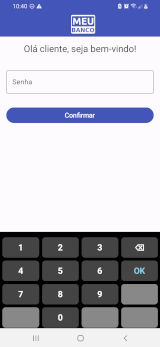
Os sistemas de dinâmica de digitação costumam requerer longo treinamento com repetição de digitação de texto definido por aproximadamente 15 vezes (Nanavati et al, 2002), o que pode ser bastante desencorajador para um participante voluntário da pesquisa. Para contornar esse problema, a quantidade de repetições foi reduzida para sete transações de transferência de PIX com operações de digitação de senha, valor e chave de PIX (códigos válidos de CPF). Apenas uma senha de de oito dígitos foi fixada, enquanto cada transação recebeu código de CPF e valores distintos. Dessa forma, totalizaram oito ações de inserção de senha, uma do login inicial e sete para cada autenticação de transação, enquanto que foram sete as ações de valores de transferência e chaves PIX.

A coleta de dados de rolagem horizontal precisou ser adaptada ao *layout* da aplicação para aproveitar ao máximo a interação do participante. Concretamente, para alcançar a opção final de transferência PIX, o participante foi obrigado a realizar a ação de deslizar botões em uma lista (em Compose, *lazyrow*) em duas sessões consecutivas. Os participantes que concluíram as sete transações contabilzaram assim, catorze ações nas áreas de rolagem, com quantidade de toques conforme a interação do usuário.

Devido à limitação dos dados dos eventos disparados no teclado virtual nativo do sistema operacional disponibilizados pela API, foi desenvolvido um componente de teclado capaz de capturar características de pressão do toque. Esse teclado considerou a digitação de forma contínua como qualquer outro toque, capturando assim, variações da pressão desde o início até o fim da intervenção na tecla.

Apenas os participantes que concluíram devidamente as sete transações foram considerados, totalizando trinta e um entrevistados, com média de duração de 8,45 minutos , incluindo o cadastro e o treino incial. Contabilizou-se 13 pessoas do sexo feminino e 18 do masculino, com média de idade aproximada de 38 anos para ambos.

Todas as telas do emulador bancário responsáveis pelas ações de captura de dados estão dispostas na Figura 01: teclado do login (I), a rolagem da entrada (II), rolagem da área PIX (III), teclado do valor (IV), teclado da chave pix (V) e teclado de autenticação (VI). Os botões circulares dispostos horizontalmente em II e III sob o título de “Serviços”(II) e “Enviar”(III) correspondem à rolagem horizontal com captura de dados.



I

II

III

IV

V

V

VI

**Figura 01.** Telas do aplicativo com captura de dados

Toda ação de toque foi capturada com um tipo de evento do gesto na biblioteca da API: 1 para pressionamento inicial, 2 quando o pressionamento é liberado e 3 para o movimento intermediário. De forma ordenada, toda intervenção começa em 1, termina em 2 e, dependendo da extensão do gesto, haverá de zero a vários registros com a categoria 3. Por exemplo, um toque bastante breve, sem característica de movimento, não recebe o valor 3. Os demais dados brutos são apresentados na tabela 1, de acordo com a ação da tela.

**Tabela 01.** Dados brutos extraídos por tela do aplicativo:

|  |  |
| --- | --- |
| Telas | Dados |
| Login, valor, chave e autenticação | Tipo de evento, caractere digitado, pressão, offset posição x, offset posição y e carimbo de tempo. |
| Tela de rolagem inicial e Pix | Tipo de evento, pressão, offset posição x, offset posição y e carimbo de tempo. |
| Todas as telas (sensores) | Tipo de sensor, coordenadas x, y e z, e carimbo de tempo |

O dado pressão representa a ação em milímetros de mercúrio (mmHg), as posições offset x e y estão relacionadas com o tamanho da tela e o carimbo de tempo em milisegundos.

Seguindo estudo referência (Incel et al, 2021), apenas foram coletados dados tridimensionais de movimento dos sensores acelerômetro, giroscópio e magnetômetro na frequência de 100 Hz. O acelerômetro e o giroscópio são sensores de movimento. O primeiro mede em m/s² a força de aceleração ao longo dos eixos, com a gravidade inclusa, enquanto que o segundo afere a taxa de rotação ao redor dos eixos em m/s2. O magnetômetro é um sensor de posição que calcula a instesidade do campo magnético ao longo dos eixos em microtesla (μT).

Adicionalmente, para melhor controle dos dados, cada registro recebeu um identificador do usuário, um dado categórico da ação conforme a tela e o número da ação da transação (de um a sete).

**Adição de variáveis**

Os dados de rolagem e do teclado receberam variáveis da diferença do tempo em relação a um registro anterior. Assim, foram adicionados os valores de tempo referente a um toque único e entre ações na tela, permitindo a mensuração dos intervalos entre os gestos, sejam esses de digitação ou de rolagem. Semelhantemente, adicionou-se variáveis do deslocamento nos eixos horizontais e verticais com as respectivas velocidades.

No caso dos sensores, os dados tridimensionais de movimento do acelerômetro, giroscópio e magnetômetro recebeam a variável vetor com a magnitude, ou seja, a raiz da soma dos quadrados dos valores de cada dimensão. Essa prática tem eficácia comprovada por estudos anteriores (Buriro, 2016).

Na prática, os valores temporais são calculados com a diferença do carimbo de data/hora com as ações do início (down) e conclusão do pressionamento (up) de uma tecla. O intervalo entre ações de teclas também é considerado importante variável a ser analisada. Também será seguida a criação de variáveis com medidas média, desvio padrão, mínimo, máximo, percentil, distância euclidiana, variância, curtose e assimetria com dados temporais.

De forma semelhante, os dados disparados pelos eventos do gesto na tela também possuem características temporais, com a diferença de também apresentarem coordenadas espaciais em duas dimensões, área e pressão do toque do dedo. O enfoque da abordagem da pesquisa será o movimento de rolagem de tela que alcança a opção do botão de transação do PIX. Inicialmente os dados serão coletados na frequência de 100 Hz à semelhança do estudo similar (Incel et al, 2021).

Adicionalmente, os dados tridimensionais de movimento do acelerômetro, giroscópio e magnetômetro também serão coletados pela API com a criação de vetor com a magnitude, ou seja, a raiz da soma dos quadrados dos valores de cada dimensão. Essa prática tem eficácia comprovada por estudos anteriores (Buriro, 2016).

Após a extração de dados, haverá fusão de todas as variáveis, com posterior seleção das características mais importantes e formação de modelo (template) em banco de dados. Serão implementados algoritmos comuns na análise de dinâmica de toque: Regressão Logística, Naive Bayes, K-Vizinhos mais Próximos (K-Nearest Neighbors) , Árvore de Decisão, Floresta Aleatória (Random Forest ) e Máquina de Vetores de Supoert (SVM).

As métricas utilizadas na avaliação do desempenho são as presentes em sistemas de biometria: a) Verdadeiros Positivos, b) Falsos Positivos, c) Falsos Negativos, d) Falsos Positivos, e) Acurácia, f) Proporção de erro equilibrada (EER, equal error rate). De fato, o EER foi a medida mais encontrada nos trabalhos pesquisados, embora não seja muito aplicável a sistemas reais (Nanavati, 2002).

**Resultados Preliminares**

Uma vez definido um gesto único de toque na tela, calculou-se a média de toques por pessoa em aproximadamente 81 gestos. Esse valor aparentemente destoa do estudo da abordagem Dakota que totalizou 240 (Incel et al, 2021) em três posições, entretanto, abrange apenas um tipo de postura. As médias gerais do tempo de toque e pressão, por cada tipo de tela (início e pix), obtiveram valores próximos. No aspecto temporal foram 10,81 e 10,77 minutos enquanto que 0,002112 e 0,002142 para a pressão. Essa proximidade é esperada de telas projetadas para a mesma funcionalidade.

Coefiente de variação

De maneira preliminar, percebeu-se

**Conclusão(ões) ou Considerações Finais**

**Agradecimentos**

**Referências**

**Apêndice ou Anexo** (opcional)

Olá participante, muito obrigado por sua colaboração nessa pesquisa.

Você irá utilizar um aplicativo que emula uma aplicação bancária. O nome do banco é “MEU BANCO” e você é simplesmente identificado por “cliente”.

Nessa aplicação você possui R$30.000,00 em saldo e enviará valores por meio de transações PIX.

I. FASE DE TREINO

Nesse momento inicial você pode ficar à vontade para navegar pelo aplicativo, tirar dúvidas com o entrevistador e realizar transferências PIX, sem restrições quanto à senha, código CPF e valor.

II. FASE DE TESTE

Essa fase corresponde ao início da pesquisa, com restrições de senha, valor e CPF.

Sua senha de acesso é: **54831790**

Você realizará 7 (sete) transferências PIX, sequencialmente para os seguintes códigos de CPF e respectivos valores:

1) CPF **010.095.374-31**, valor **R$8.000,00**

2) CPF **343.268.720-63**, valor **R$1.590,00**

3) CPF **091.999.344-34**, valor **R$1.699,99**

4) CPF **068.511.503-87**, valor **R$5.000,00**

5) CPF **796.272.160-03**, valor **R$3.000,00**

6) CPF **008.465.574-79**, valor **R$1.34’0,50**

7) CPF **586.936.170-28**, valor **R$4.000,00**

Ao concluir todas essas ações você será redirecionado para a tela “Controle de pesquisa”

**Anexo 1.** Folha auxiliar usada na coleta dos dados com as informações de digitação