Autenticação de Senhas Faladas Dependente de Usuário

Lucas Correia Villa Real*
Farlei José Heinen
Luiz Paulo Luna de Oliveira

[1]Programa de Pós-Graduação em Computação Aplicada – PIPCA
Centro de Ciências Tecnológicas e Exatas
Universidade do Vale do Rio dos Sinos
São Leopoldo – RS – Brasil

{lucasvr, farlei, luna}@exatas.unisinos.br

Abstract. *This paper presents* VoiceAuth[1]*, a system developed to allow voice authentication for GNU/Linux. By using it, the user is recognized based on the similarity between a previously spoken reference (and stored) password, and a new speech of the same password. The comparison is done using correlation analysis between the coefficients taken from the Windowed Fast Fourier Transform of both passwords (reference and the new speech).*

Resumo. *Este artigo apresenta* VoiceAuth[1]*, um sistema para permitir autenticação de usuários a partir de senhas faladas no GNU/Linux. No seu uso, o usuário é reconhecido pela comparação entre uma senha falada de referência previamente adotada (e armazenada), e uma nova execução (pronúncia) da mesma senha. A comparação se baseia na correlação entre os coeficientes obtidos através da Transformada Rápida de Fourier janelada de ambas as senhas (de referência e executada).*

1. Introdução

O uso de métodos biométricos para autenticação permite que o usuário tenha uma maneira mais natural de garantir a sua identidade. Alguns destes métodos incluem reconhecimento através da íris, de impressões digitais, reconhecimento da face e reconhecimento da voz [Matyas and Riha, 2000]. Um sistema deste tipo é desejável nos dias atuais, onde sistemas de autenticação baseados em biometria são utilizados para aumentar o acesso seguro a dispositivos, além de permitir a pessoas deficientes uma interface efetiva para realizar acessos a suas contas de usuário ou a outros recursos com uso de senhas. É ainda importante destacar que, pela disponibilidade do GNU/Linux, o software proposto pode ser utilizado em um largo conjunto de sistemas embarcados, além das estações de trabalhos e computadores pessoais.

A escolha de um método baseado em reconhecimento de voz tem, como motivação, a disponibilidade do hardware necessário em uma grande parte dos computadores atuais (microfones, placas de som, etc.). Isto permite que o sistema seja utilizado de forma mais ampla do que caso fosse necessário o uso de hardware caro e restrito. Neste trabalho, propomos um método de reconhecimento de usuário por senha falada dependente de locutor, com o uso de software livre e bibliotecas de grande generalidade.

*ITI-CNPq
[1]http://cscience.org/ lucasvr/projects/voiceauth.php

Uma senha falada pode se constituir de uma palavra ou frase que, em última análise, é uma seqüência temporal de sons específicos. Por isto, é desejável que um método de reconhecimento de senhas faladas seja sensível tanto às mudanças temporais dos sons quanto ao seu conteúdo de freqüências. Por outro lado, é desejável também que ele seja sensível ao locutor, o que se traduz novamente em uma sensibilidade do método aos conteúdos de freqüências. Motivados por isso, propomos um método que se baseia na comparação por correlação dos coeficientes da Transformada Rápida de Fourier (FFT) janelada [Oppenheim and Schafer, 1980], o que dá a este método a característica de ser sensível ao vocábulo propriamente dito (sílabas e suas durações) bem como ao locutor, com o conteúdo de freqüências características de sua voz.

2. Método Proposto

VoiceAuth foi desenvolvido como um módulo para a biblioteca Linux-PAM (*Pluggable Authentication Modules*), estendendo as funcionalidades desta. O Linux-PAM é uma implementação de software livre baseada na proposta original do PAM da SUN Microsystems [Samar and Schemers, 1995]. O Linux-PAM é uma biblioteca que fornece um mecanismo flexível para a autenticação de usuários. Utilizando PAM, não há a necessidade de modificar o código fonte de aplicações para que estas façam o uso deste método biométrico (desde que elas já tenham suporte às facilidades oferecidas pelo PAM). Assim, tem-se um método de autenticação que permite a uma grande gama de aplicativos aproveitar-se deste recurso.

No seu uso, o usuário pronuncia pelo microfone uma senha falada de referência, que então é pré-processada pela Transformada de Fourier Janelada. São então gravados os coeficientes de Fourier de cada uma das J janelas de Fourier utilizadas. Para autenticar seu login, o usuário executa novamente a senha de referência, falando-a pelo microfone. Da execução da senha, que será então pré-processada da mesma maneira que a senha de referência, extraem-se os coeficientes de Fourier de cada janela, os quais serão comparados por correlação com aqueles da senha de referência, armazenados anteriormente. Abaixo, são dados mais detalhes do uso do método proposto.

Registro da senha de referência:

- **Calibragem de Silêncio:** usada para determinar o nível de ruído do ambiente. O conjunto de dados que será processado pelo software inicia após a primeira ocorrência de uma amplitude superior ao nível de silêncio determinado pelo sistema, e termina quando este é reestabelecido. Tal calibragem atualmente dura 3 segundos.
- **Registro da Senha:** o usuário expressa sua senha através da voz quando solicitado pelo software. Fazendo uso da amplitude definida na etapa de calibragem de silêncio, apenas o sinal relevante é processado pelo software.
- **Armazenamento Físico:** o nome do usuário é adicionado a uma entrada em um arquivo, onde têm-se controle de quais usuários estão fazendo uso deste sistema de autenticação biométrico. Um arquivo contendo as partes *reais* dos coeficientes de Fourier da senha nas J janelas utilizadas é armazenado para futuras comparações com as tentativas de autenticação no sistema.

Entrada no sistema (login):

- **Calibragem de Silêncio:** bem como no registro da senha de referência, este passo é necessário para determinar o início e o término da fala do usuário.
- **Tentativa de Autenticação:** o usuário expressa através da voz sua senha previamente registrada.
- **Normalização do Sinal:** como o sistema precisa comparar a senha falada com a senha armazenada, os sinais precisam estar representados na mesma escala. Este passo realiza

a normalização do sinal em função do tempo (comprimento dos sinais) e da amplitude (intensidade do sinal). Caso o comprimento dos sinais tenha uma diferença significativa (onde esta diferença pode ser especificada pelo administrador), o processamento é cancelado neste ponto, e a autenticação não é bem sucedida pelo usuário.

- **FFT Janelada:** a senha falada é dividida em J janelas. O final do *buffer* referente às janelas é preenchido com *zeros*, caso o sinal não ocupe-o completamente. Uma rotina de FFT é aplicada nestas J janelas, modificando a representação do sinal: ele passa da representação em função do tempo para uma representação em função das freqüências, permitindo obter dados relativos à identidade do locutor.
- **Correlação:** após processadas pela FFT, as freqüências são analisadas em relação à similaridade das partes *reais* de seus coeficientes. Para isto, uma função de correlação é aplicada na saída gerada pela FFT para cada uma das J *janelas*. Existe um limite X definido pelo módulo em um arquivo de configuração: caso 2 janelas subseqüentes tenham uma correlação inferior a X, as senhas são definidas como diferentes, e o reconhecimento é interrompido. Caso não hajam janelas subseqüentes com valores com valores inferiores a este limite X estabelecido, a média das somas obtidas no processo de correlação para as J janelas é comparada com um outro limite Y. Caso esta média seja maior que o valor especificado por Y, as senhas são ditas equivalentes, e o usuário é autenticado.

3. Interfaces

Para o uso deste sistema, alguns programas foram desenvolvidos para auxiliar a administração do sistema. O módulo contendo as rotinas de autenticação do PAM é apresentado na biblioteca compartilhada `voiceauth.so`. Para validar este módulo, um programa denominado `voicepasswd` foi criado, de forma a manter um registro de usuários que têm permissão para o uso deste serviço. Para todo usuário adicionado a este arquivo, um arquivo extra, salvo sob o mesmo nome do *username*, é gerado contendo a senha deste usuário. Um programa denominado `voicelogin` foi ainda criado para interagir com o módulo, permitindo que o login padrão do sistema seja realizado através da voz.

Existem várias configurações do sistema que podem ser ajustadas para modificar dispositivos e características na função de correlação. Os parâmetros aceitos pelo módulo são:

- `va_dir:` define o diretório a ser utilizado como repositório para os arquivos de senha;
- `audiodev:` define o dispositivo de áudio a ser utilizado, ao invés do padrão `/dev/dsp`;
- `mixerdev:` define o dispositivo de mixer a ser utilizado, ao invés do padrão `/dev/mixer`;
- `mic_level`: ajusta o volume do microfone a um nível específico;
- `igain_level:` ajusta o ganho no sinal de entrada a um nível específico;
- `silence_level:` ajusta o nível de silêncio. Este parâmetro pode ser utilizado em um ambiente onde o nível de silêncio é conhecido. Este valor pode ainda ser redefinido pelo aplicativo, no instante da calibragem do nível de silêncio do ambiente;
- `silence_threshold:` define a amplitude na qual o sistema irá passar a tratar o sinal, definindo o início e término da fala do locutor;
- `corr_threshold:` determina o grau de semelhança necessário entre os sinais para defini-los como “iguais” ou não.
- `passwd_maxlen_diff:` determina a diferença máxima permitida no tamanho entre a senha armazenada e a falada. Caso uma senha tenha um comprimento fora dos limites estabelecidos, ela já é dada como incorreta, poupando o processamento da FFT.

VoiceAuth utiliza-se das bibliotecas OSS e ALSA para interagir com o dispositivo de som. A OSS (*Open Sound System*) é a interface original para a programação de áudio no GNU/Linux,

mantido pela 4Front Technologies. A OSS suporta a maioria das placas de som atuais, no entanto alguns dos drivers mais recentes necessitam de uma licença comercial da OSS. O ALSA (*Advanced Linux Sound Architecture*) é um projeto comunitário que busca superar o OSS em todas as áreas, fornecendo um conjunto completo de drivers de placas de som, livres de licenças comerciais.

A freqüência de amostragem utilizada no VoiceAuth é de 44100 Hz, enquanto a quantização (8, 16, 24-bit) é determinada pelo módulo `voiceauth.so`, permitindo que a melhor qualidade de áudio seja empregada [Pholmann, 1995].

4. Resultados e Conclusão

O recurso apresentado neste artigo consiste em uma extensão de uma ferramenta disponibilizada segundo a política de Software Livre. O sistema propõe ser uma alternativa válida como método de autenticação para sistemas GNU/Linux, tendo como vantagem a implantação de um sistema de autenticação biométrico de baixo custo. Os resultados preliminares obtidos com os testes demonstraram que o sistema foi capaz de identificar usuários através da senha falada, com uma taxa de acerto não inferior a 80%, apesar de um pequeno tempo de treinamento ser necessário para que o usuário consiga habituar-se a expressar sua senha através da voz. Nos testes preliminares não foram utilizados equipamentos profissionais, como hardware dotado de filtros ou de conversores de alta fidelidade.

Ainda existe a possibilidade de adaptar um método de criptografia às senhas armazenadas no sistema. Entretanto, um certo grau de segurança já é garantido pelo armazenamento da senha a partir da parte *real* dos coeficientes de Fourier. Isto permite que sejam armazenados apenas os dados relevantes à função de correlação, evitando que um possível invasor no sistema tenha acesso ao sinal original. Na implementação atual, existe ainda uma precaução relacionada a evitar que a memória utilizada pela leitura da senha seja *swapped-out* para o disco rígido, através da família de funções *mlock()*. Além disso, as permissões utilizadas pelos arquivos de senhas individuais são feitas de forma a garantir que estes sejam lidos apenas pelos seus proprietários.

A idéia apresentada admite muitas possibilidades de aperfeiçoamento. Por exemplo, pode-se tentar separar os conteúdos de freqüência do locutor (*pitch*) e da senha propriamente dita. Feito isto, a validação de senhas poderia se dividir em reconhecimento de locutor e da senha propriamente dita. Isto poderia agilizar o processo e até mesmo torná-lo mais preciso. Outra possível melhoria seria a implementação do método utilizando técnicas de Wavelets para o processamento de sinais. Tais extensões estão sendo consideradas como temas de trabalhos futuros.

Referências

Matyas, V. and Riha, Z. (2000). Biometric Authentication — Security And Usability.

Oppenheim, A. V. and Schafer, R. W. (1980). *Discrete - Time Signal Processing*. Prentice-Hall International, Inc.

Pholmann, K. (1995). *Principles of Digital Audio*. McGraw-Hill, Inc., 3 edition.

Samar, V. and Schemers, R. (1995). Request for comments: Unified Login with Pluggable Authentications Modules (PAM). Technical Report 86.0, SunSoft.