Emanações Acústicas de Teclado – Um Experimento Básico

Eduardo Leivas Bastos
UNISINOS
elbastos@unisinos.br

Felipe Teixeira de Cidade Moura
UNISINOS
felipe.cidade@linqsistemas.com.br

Resumo—**Os equipamentos eletrônicos podem emitir sinais de modo não intencional que podem estar relacionados com os dados que estão sendo processados. Tais emanações de energia podem ser utilizadas por usuários mal intencionados para o acesso a informações confidenciais sem que haja uma intrusão direta no sistema. Esse artigo oferece uma visão geral e atualizada das técnicas utilizadas para a exploração de emanações acústicas de teclado e apresenta os resultados de um experimento de captura e análise dos sons de duas teclas de um teclado convencional para desktop padrão IBM/PS2. Os resultados demonstram que, mesmo com a utilização de recursos de fácil obtenção e operação, é possível distinguir facilmente e de modo eficaz os sons das duas teclas.**

I. Introdução

Os equipamentos eletrônicos podem emitir sinais de modo não intencional e que podem estar relacionados com os dados que estão sendo processados. Tais emanações de energia podem ser utilizadas por usuários mal intencionados para o acesso a informações confidenciais sem que haja uma intrusão direta no sistema, usando a apenas a correlação direta desses sinais com os dados processados. Dessa forma, emanações acústicas, eletromagnéticas e óticas podem ser interceptadas e analisadas de forma "indireta" e os dados a elas associados podem ser gerados.

As emanações acústicas de teclado (*keyboard acoustic emanations*) representam um risco sério que não deve ser menosprezado pelos profissionais de segurança. Em artigo seminal, Asonov e Agrawal (2004) propuseram uma técnica para a captura e análise de sons de teclado com o auxílio de redes neurais e amostras rotuladas [1]. Posteriormente, Zhuang et al. (2009) mostraram que é possível associar os sons das teclas a classes acústicas da língua inglesa com base apenas na frequência das letras na linguagem de origem, sem a necessidade de amostras rotuladas [8].

A aplicação da técnica de Zhuang et al. (2009) em outros idiomas, como o Português, é uma questão de pesquisa em aberto e que pode trazer resultados relevantes para a segurança dos sistemas computacionais brasileiros, especialmente para aqueles dependentes de autenticação por teclado convencional. Além disso, o uso da técnica em outro idioma oferece a possibilidade de testar sua eficácia em outro cenário, aumentando sua validade externa. Um passo importante para a concretização desse objetivo é o estudo de como os sons das teclas podem ser capturados e os padrões resultantes dessa captura.

Esse artigo está dividido em quatro partes. Na primeira parte são descritos de modo geral os principais avanços obtidos com técnicas de captura de emanações acústicas, em especial aquelas oriundas de teclados comuns. A segunda parte descreve o experimento básico para a captura dos sons de duas teclas e a geração dos respectivos espectros, estas escolhidas a fim de limitar a análise visual das amostras. A terceira parte apresenta e discute os resultados obtidos no experimento. Por fim, são apresentadas as limitações do experimento e indicações para trabalhos futuros na área.

II. Emanações Acústicas de Teclado

Os equipamentos eletrônicos de transmissão e processamento de dados recebem e emitem energia de várias formas, tais como eletricidade, calor, luz, ondas eletromagnéticas, vibração e sons. A maioria dessa energia é dissipada em forma de calor ou é utilizada para a transmissão dos dados nos meios de transmissão. Uma parte dessa energia, entretanto, possui alguma correlação com os próprios dados que estão sendo processados e pode ser utilizada por usuários mal intencionados para a obtenção de acesso a informações confidenciais e comprometer a segurança dos sistemas computacionais [1] [5]. Tais ataques são denominados de ataques por canal lateral (*side-channel attacks*), pois se baseiam na criptoanálise das informações que vazam dos dispositivos de modo não intencional [7].

Diversas técnicas têm sido propostas na literatura para a obtenção de informações de emanações eletromagnéticas e óticas [4][6]. As emanações acústicas também podem ser utilizadas como fontes de dados para ataques. Pesquisas demonstram que emanações acústicas de uma impressora matricial podem carregar informações substanciais a respeito do texto que está sendo impresso [2]. Shamir e Tromer (2004) sugerem a possibilidade de descobrir as operações executadas pela CPU pela análise de suas emanações acústicas [7].

Asonov e Agrawal (2004) mostraram que é possível recuperar o texto que está sendo digitado analisando-se as emanações acústicas do teclado (*keyboard acoustic emanations*) [1]. A técnica empregou redes neurais para o treinamento prévio dos diversos pares de teclas-sons gravados de um mesmo digitador e de um mesmo teclado. A taxa de reconhecimento de caracteres obtida pela técnica chegou a 79% [9]. Apesar da alta taxa de reconhecimento obtida, a técnica possui a desvantagem de necessitar de

uma amostra identificada (pares teclas-sons) de bom tamanho antes da execução do ataque, o que limita a sua aplicabilidade de utilização (o atacante teria que previamente associar a tecla com o som correspondente).

Recentemente, Zhuang et al. (2009) propuseram uma técnica que não requer a utilização de uma amostra identificada previamente [8]. Os autores basearam-se no fato de que os textos digitados seguem normalmente uma gramática e, por isso, não são randômicos. Dessa forma, os sons das teclas podem ser agrupados em determinadas classes acústicas. Dada uma quantidade suficiente de amostras de treinamento não identificadas, um mapeamento provável entre tais classes acústicas e os caracteres realmente teclados pode ser obtido usando-se as restrições intrínsecas da linguagem de origem (certas sequências de letras são mais ou menos prováveis de existirem em determinadas linguagens). Os experimentos realizados com a técnica aplicada a textos digitados na língua inglesa apresentaram taxas de reconhecimento de 87,6% para palavras e 95,7% para caracteres [10].

III. EXPERIMENTAÇÃO

A fim de verificar a complexidade envolvida na coleta das emanações acústicas de teclado e testar os primeiros passos da técnica de Zhuang et al. (2009) para estendê-la para a língua portuguesa, um experimento básico foi desenhado para a coleta e análise visual dos sons de um conjunto de teclas alfanuméricas. A Figura 1 mostra a sequência de etapas seguidas nesse experimento.

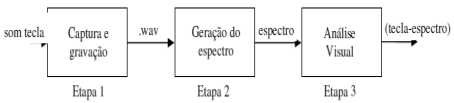

Figura 1 – Sequências de etapas do experimento.

Para a realização da captura dos sons das duas teclas foi utilizado um microfone de pedestal convencional para computadores e um teclado de 101 teclas padrão IBM. Para a gravação dos sons foi utilizado o próprio gravador do sistema operacional Windows XP SP3. Foi utilizado o padrão de arquivo `.wav` com 01 (um) canal de captação. Esse padrão foi escolhido em função de sua ampla aceitação por diversos aplicativos no mercado.

Cada tecla alfanumérica foi pressionada (03) três vezes por um mesmo digitador em um ambiente sem qualquer tipo de isolamento acústico. Os sons gerados foram capturados e gravados em arquivos sequenciais no disco rígido da máquina coletora. O padrão utilizado para a nomeação dos arquivos foi o X00Y, onde a posição "X" foi utilizada para representar a tecla pressionada e a posição "Y" o número da amostra gerada. Desse modo, o arquivo `A001.wav` representa a primeira amostra gerada para a tecla "A". Tomou-se cuidado para não haver variação na pressão exercida na tecla pelo digitador a fim de permitir a comparabilidade das amostras (apesar disso não representar o modo convencional de digitação por um usuário normal).

Os arquivos `.wav` foram importados para o programa MATLAB através da função `wavread()`. Cada arquivo foi transformado do domínio tempo para o domínio frequência com o uso do algoritmo FFT (*Fast Fourier Transform*) através do uso da função `specgram()`. A FFT é uma otimização do algoritmo convencional FT (*Fourier Transform*), utilizado para a obtenção das frequências características (espectrograma) de um determinado sinal de entrada. A técnica *cepstrum* [3] também poderia ter sido utilizada nessa fase por ser mais eficaz na geração dos espectros, mas optou-se pela FFT em função de sua disponibilidade imediata no pacote MATLAB e pela farta documentação disponível em livros e artigos especializados.

O uso das técnicas em conjunto com algoritmos de classificação de padrões (sejam eles visuais ou analíticos) torna possível a geração de classes acústicas que possuem associações mais prováveis com determinadas teclas dentro de uma mesma linguagem fonte. Essa é a base da técnica proposta por Zhuang et al. (2009). Um exemplo de espectro obtido com o uso da técnica FFT pode ser visto na Figura 2, onde a tecla "A" é digitada três vezes.

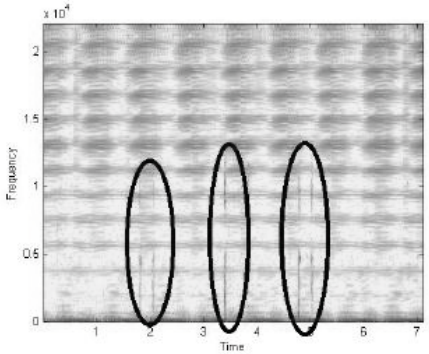

Figura 2 – Espectro resultante da digitação do caractere "A" três vezes consecutivas.

Os espectros obtidos foram então categorizados de acordo com o caractere correspondente em um banco de dados. Apesar de a técnica de Zhuang et al. (2009) não necessitar de amostras rotuladas para a sua execução, optou-se pela geração de um banco de dados identificado a fim de facilitar a execução de novos experimentos e possibilitar um maior controle nesse estágio inicial.

IV. RESULTADOS

A Figura 3 apresenta os resultados obtidos depois da aplicação do método para as teclas "A" e "B", respectivamente. Uma análise visual preliminar mostra os instantes de tempo em que as teclas foram pressionadas e liberadas (*push peak* e *release peak*). De acordo com Zhuang et al. (2009), o tempo transcorrido entre esses dois eventos é tipicamente da ordem de 100ms. Esse fato pode ser verificado visualmente nos dois espectros gerados.

Zhuang et al. (2009) afirmam que a energia dos toques de teclas está concentrada principalmente entre 400Hz e 12KHz. A análise visual dos espectros da Figura 3 comprova esse fato. Uma análise mais detalhada dos respectivos espectros na janela de tempo de 100ms é necessária para se obter um maior entendimento dos padrões de energia na ocorrência do toque.

Os espectros também mostram uma energia constante na faixa entre 7000Hz e 8000Hz. Apesar de as escalas de tempo estarem defasadas, é possível perceber que em um caso (tecla "A"), a energia está contida dentro do intervalo de pressão/liberação da tecla, o que não acontece no espectro do caractere "B". Tal fato pode ser decorrência do processo de gravação ou decorrente da presença de algum equipamento emitindo sinais nessa frequência.

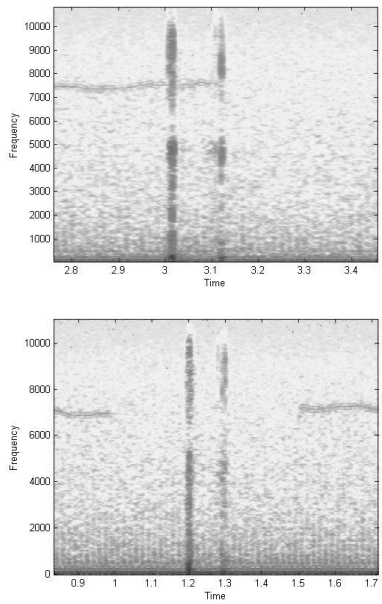

Figura 3 – Espectros gerados com os sons das teclas "A" e "B" respectivamente.

V. LIMITAÇÕES

O uso de um microfone convencional em um ambiente sem isolamento acústico também são fatores que podem ter influenciado os resultados obtidos. No entanto, considerando-se que todas as amostras foram obtidas em um mesmo ambiente e em um mesmo período do dia, pode-se considerar o ruído externo como uma constante que pode ser desprezada nas análises. Mesmo com a presença de um razoável ruído de fundo, é possível perceber com clareza os momentos de pressão/liberação das teclas. Experimentos futuros deverão ser realizados com microfones mais sensíveis e ambientes menos sensíveis a ruídos de forma a melhorar as análises gráficas e matemáticas.

Outro fator limitante do experimento foi a impossibilidade de se analisar com maior nível de detalhamento a faixa de espectro até 1000Hz. Essa faixa contém a maior energia espectral nos tempos de pressão/liberação e deve ser analisada com maior resolução como formal de auxiliar o processo de comparação não visual dos espectros. Essa limitação deve-se ao entendimento ainda incipiente da ferramenta MATLAB. Novos experimentos deverão ser gerados com uma resolução maior nessa faixa do espectro.

VI. CONSIDERAÇÕES FINAIS

A captura de emanações de energia oriundas de dispositivos eletrônicos tem sido alvo de muitas pesquisas nos últimos anos. Diversas técnicas têm sido propostas na literatura para a exploração desses 'vazamentos' não intencionais de informação. Usuários mal-intencionados podem se valer de tais vulnerabilidades para acessarem informações confidenciais que podem comprometer a segurança dos sistemas computacionais. As emanações acústicas representam uma classe de emanações de energia na qual os sons emitidos pelo dispositivo (teclado, impressora matricial, etc) estão associados aos dados processados (teclas digitadas, texto impresso). Zhuang et al. (2009) propuseram uma técnica para o reconhecimento de teclas digitadas em teclados convencionais baseada apenas na linguagem de origem do texto, não havendo a necessidade de um treinamento prévio do algoritmo com amostras rotuladas.

Esse artigo demonstrou a possibilidade de captura e análise visual prévia dos sons oriundos de um teclado convencional apenas com a utilização de ferramentas de coleta de baixo custo e um software de processamento matemático de uso geral para a geração do espectro dos sinais. Os espectros obtidos no pressionamento das duas teclas (Figura 3) mostram que é possível diferenciá-las preliminarmente em modo visual (nas frequências mais baixas, menores do que 1000Hz). Além disso, é notável a existência de dois 'picos' de energia, mostrando o momento em que as teclas foram pressionadas e liberadas, sendo o pressionamento da tecla o ponto de diferenciação das teclas, essa diferença indicada graficamente pela a

largura dos espectros, representado energias diferenciadas para cada uma das teclas A e B. Os resultados obtidos mostram que tais técnicas de ataque por 'canais laterais' devem ser levadas em consideração nas políticas de segurança das organizações. A obtenção de resultados significativos, mesmo com o uso de ferramentas de fácil acesso, deve servir de alerta para a definição de perímetros de segurança mais abrangentes.

Apesar de demonstrar que a captura e análise visual prévia dos sons gerados pelas teclas de teclados convencionais são processos que podem ser executados com a utilização de ferramentas simples e de baixo custo, o presente trabalho apresenta algumas limitações. A primeira delas está relacionada com o próprio método de geração das amostras. Apesar de ter sido tomado o cuidado para não haver diferenças significativas de pressão de toque entre as três amostras de cada tecla, é praticamente impossível garantir essa premissa. Apesar de Zhuang et al. (2009) mencionarem que teclas iguais pressionadas de modo diferente podem ser catalogadas em classes acústicas distintas, os autores utilizam as próprias características da linguagem de origem e dicionários para minimizarem os erros prévios de coleta. Além disso, o tempo decorrido entre o início da gravação e o pressionamento de cada tecla mostrou-se variável, o que pode impactar negativamente em uma análise visual preliminar. Como podem ser observados na Figura 2, os tempos em que ocorrem os pressionamentos das duas teclas estão defasados entre si. Apesar de o fenômeno de interesse ser especificamente o pressionamento/liberação da tecla, essa defasagem pode dificultar análises mais complexas da estrutura do espectrograma.

A geração de espectros de arquivos sonoros de teclas é apenas um passo para o domínio total da técnica de Zhuang et al. (2009). Diversas outras etapas devem ser cumpridas a fim de se obter um protótipo funcional da técnica adaptado para a língua portuguesa. Entre os processos que devem ser aprimorados incluem-se: a) melhorar o entendimento das diversas técnicas de análise acústica, entre elas a FFT e *cepstrum*; b) comparar as facilidades de uso e os resultados obtidos pelas técnicas de análise acústica com outras ferramentas; c) realizar as análises visuais para as outras teclas do mesmo teclado; d) realizar captura e análise das teclas de outro teclado e comparar os resultados e; e) desenvolver software para a captura e análise integrada do espectro acústico das teclas.

As emanações acústicas de teclado podem servir de ponto de entrada para diversos ataques por 'canal lateral' a sistemas computacionais. Teclados sem ruído, baseados em toque (*touch*) ou projetados por luz podem ser utilizados para impedir a realização de tais ataques. No entanto este tipo de ataque, devido a sua alta complexidade no desenvolvimento de um algoritmo de captura, e a necessidade um treinamento de uma rede neural como realizado por Asonov e Agrawal (2004) ou gravação previa do teclado como no experimento de Zhuang et al. (2009), acaba se tornando muito difícil onde não temos nenhum caso registrado desse tipo de ataque, mas esta ameaça não pode ser descartada, pois com essa técnica aprimorada é possível ser aplicada em caixas automáticos, urnas eletrônicas ou até mesmo dispositivos moveis.

VII. AGRADECIMENTOS.

Gostaríamos de agradecer o empenho e a dedicação de todos os integrantes que já fizeram parte e daqueles que hoje formam o Grupo de Estudos de Emanações Acústicas (GE-EAT) da UNISINOS. Sem o apoio incansável de todos, esse trabalho não seria possível. Em especial, gostaríamos de agradecer ao Prof. MSc. Leonardo Lemes pela oportunidade de pesquisa em um campo ainda pouco explorado da área de Segurança da Informação e pelo constante incentivo em aprimoramento científico e inovação tecnológica.

REFERÊNCIAS

[1] Dimitri Asonov and Rakesh Agrawal, "Keyboard Acoustic Emanation", IEEE Symposium on Security and Privacy. pp.3, 2004.

[2] R. Briol, "Emanation: How to Keep Your Data Confidential", Symposium on Electromagnetic Security for Information Protection, SEPI 91, Rome, 1991.

[3] D. Childer, D. Skinner and R. Kemerati, "The Cepstrum: A Guide to Processing", Proceeding of the IEEE, v.65, n.10, pp. 1428-1443, October, 1997.

[4] Markus Khun and Ross Anderson, "Soft Tempest: Hidden Data Transmission Using Eletromagnetic Emanations", Proceedings 2nd Workshop on Information Hiding. 1998.

[5] Markus Khun, "Compromising emanations: eavesdropping risks of computer displays", Technical Report Number 577. University of Cambridge, 2003.

[6] Joe Loughry and David Umphress, "Information Leakage from Optical Emanations", ACM Transactions on Information and System Security, v.5, n.3, pp. 262-289, August, 2002

[7] Adi Shamir and Eran Tromer, "Accoustic crytptanalysis. Eurocrypt 2004 rump session", Maio, 2004. Disponível em http://people.csail.mit.edu/tromer/acoustic/. Acesso em: 15 AGO 2011.

[8] Li Zhuang, Feng Zhoug and J. Tygar, "Keyboard Acoustic Emanations Revisited", ACM Transactions on Information and System Security, v.13, n.1, Article 3, 2009.

[9] Dimitri Asonov and Rakesh Agrawal, "Keyboard Acoustic Emanation", IEEE Symposium on Security and Privacy. pp.5, 2004.

[10] Li Zhuang, Feng Zhoug and J. Tygar, "Keyboard Acoustic Emanations Revisited", ACM Transactions on Information and System Security, v.13, n.1, Article 3, pp.16 2009 .