Language Building Model: A language model of size n assigns a probability to each sequence of n words. The probability distribution can be estimated by computing the frequencies of all n-grams from a large text corpus. n-grams words with probability 0 will never be selected by the Viterbi algorithm; we smooth the probabilities by assigning a small probability to each unseen n-gram. The length of an n-gram determines how many words of context are taken into account by the language model. Higher values for n can lead to better models but also require exponentially larger corpora for an accurate estimation of the n-gram probabilities. The higher the value of n, the larger the likelihood that some n-grams never appear in the corpus, even though they are valid word sequences and thus may still appear in the text.

Reordering words based on obtained language model: Having built thelanguage model, we can reorder the candidate words using the model to select the most likely word sequence [4].This task is addressed by the Viterbi algorithm [13], which takes as input an HMM (Q, O, A, B, I) of order d and a sequence of observations $a_1, \ldots, a_T \in O^T$. Its state consists of $\psi = T \times Q^d$. First, the d-th step is initialized (the earlier are unused) according to the initial distribution, weighted with the observations:

$$\Psi d,i_{1,..}i_d = Ii_{1,..}i_d \prod_{k=1,..d} Bi_k,a_k \; \forall \; 1 \leq i,j \leq N$$

In the recursion, for increasing indices s, the maximum of all previous values is taken:

$$\Psi s,i_1,..i_d = Bi_d,a_s \; max_{i0 \epsilon Q} (Ai_0,i_{1,..}i_d \; \psi s\text{-}1,i_{0,..}i_{d-1}) \; \forall_{s>d,\, 1 \leq i,j \leq N}$$

The sequence of hidden states finally can be obtained by back tracking the indices that contributed to the maximum value in the recursion step.

4.Devices under Threat and its Countermeasures

Secret information leakage caused by emanations from electronic devices has been a topic of concern for a long time. Emanations such as sound produced by electronic devices can be from different sources. [5] Sound as a wave carries information in the form of frequency, wavelength and amplitude which can bemeasured by audio capturing device like microphone. The powerful acoustic attacks sources have been keyboard, keypad of ATM machine and key strokes of printer machine and application of such attacks are mainly for capturing login detail, passwords and other secret information recovery.

An obvious idea for countermeasuring acoustic attack is a silent keyboard, which do not produce more sound. It can be a keyboard made of rubber or touchpad [13], or a keyboard based on a touchscreen or touchstream technologies [15]. Nowadays, virtual keyboards have appeared that can be projected on a flat surface [16] and printers with acoustic shielding foam which minimize sound of keys pressed. These choices are more expensive than the standard mechanical keyboard. Typing on a standard keyboard is much comfortable than typing on a touchscreen or a rubber keyboard.

The above mentioned ways are useful in avoiding emanation of sound from devices. But there are also some other methods by which we can prevent this attacks to take place. They are: *Distance*- the recognition rate drops substantially if the distance between the device and the microphone is increased. *Obstacle*- any obstacles between thedevice and microphone can prevent the sound reaching the recording device (microphone). *Avoiding contact with microphone:* the absence of microphones near emanation device is sufficient to protect privacy.

5. Conclusion

This paper describes the overview of acoustic side-channel attack and provides different techniques like HMM (Hidden Markov Model), triangulation method reordering words using Viterbi algorithm to recognize the data that is been recorded. At last, some of the countermeasures to avoid and overcome the attack.

Reference

[1] Side-Channel Attacks: Ten Years after Its Publication and the Impacts on Cryptographic Module Security Testing YongBin Zhou, DengGuo Feng State Key Laboratory of Information Security, Institute of Software, Chinese Academy of Sciences, Beijing, 100080, China.

[2] Power analysis attack Countermeasures and their weaknesses. Thomas S. Messerges, Ph.D., Security Technology Research Laboratory Motorola Labs, Motorola.

[3] Meinard M¨uller. Information Retrieval for Music and Motion. Springer, 2007.

[4] Acoustic Side-Channel Attacks on Printers. Michael Backes, Markus D¨urmuth1, Sebastian Gerling1, Manfred Pinkal3, Caroline Sporleder, Saarland University, Computer Science Department, Saarbr¨ucken, Germany. Saarland University, Computer Linguistics Department, Saarbr¨ucken, Germany.

[5] side-channels, compromising emanations and surveillance: current and future technologies. Richard Frankland.

[6] Dmitri Asimov and Rakesh Agarwal, "Keyboard Acoustic Emnations", IBM.

[7] Lawrence R. Rabiner. "A tutorial on hidden markov models and selected applications in speech recognition."

[8] Biing-Hwang Juang and Lawrence R. Rabiner. "Hidden markov models for speech recognition."

[9] Frederick Jelinek. "Statistical Models for Speech Recognition". MIT Press.

[10] Kenneth W. Church. "A stochastic parts program and noun phrase parser for unrestricted text".

[11]R. Nag, Kin HongWong, and Frank Fallside. Script recognition using HiddenMarkovModels.

[12] Steven DeRose. Grammatical category disambiguation by statistical optimization. Computational Linguistics.

[13]Hidden Marckov Models - http://cs.brown.edu/research/ai/dynamics/tutorial/Documents/HiddenMarkovModels.html

[14] The virtually indestructible keyboard. http://www.grandtec.com/vik.html

[15] TouchStream keyboards. http://www.fingerworks.com/.

[16] Canesta keyboards. http://www.canesta.com/products.html.