Entropy of audio fingerprints for unobtrusive device authentication

Stephan Sigg, Matthias Budde, Yusheng Ji, and Michael Beigl

National Institute of Informatics,
Tokyo, Japan
sigg@nii.ac.jp, kei@nii.ac.jp
http://www.nii.ac.jp

KIT, TecO,
Karlsruhe, Germany
budde@teco.edu, michael@teco.edu
http://www.nii.ac.jp

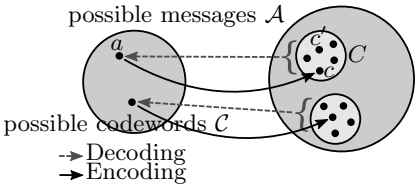
Abstract. Context-based authentication methods enable the unobtrusive establishment of authentication or even secure keys. While several context-based authentication methods have been proposed recently, often the entropy of the seed for the cryptographic keys is not exploited. We study the entropy of audio fingerprints which can be utilized to pair devices in close proximity. In this work, for 600 audio fingerprints from five distinct audio classes recorded at three different locations, we applied 7490 statistical tests from the dieHarder battery of statistical tests.

Key words: Audio fingerprinting, Entropy, Pervasive computing, Unobtrusive device authentication

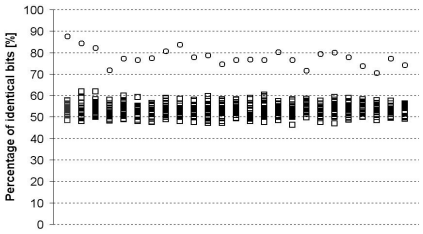
Secure communication is an important issue in pervasive environments. It is required to prevent an adversary from entering a network, to protect confidential information, or to prevent an adversary from manipulating network operations. With mobile devices, communicating entities might change frequently. Also, communications are frequently conducted with previously unknown devices. This makes it complicated to provide a mostly unobtrusive security mechanism.

We exploit ambient audio as a seed for the creation of shared cryptographic keys among devices in physical proximity. Two devices willing to establish a secure channel take a synchronized audio recording and compute audio fingerprints based on this sample. We show how secure key exchange is possible with fuzzy cryptography schemes and that the entropy of audio fingerprints is high.

The establishing of a secure communication channel between devices has been studied by several authors in the literature [1]. In [2], for instance, the authors propose two protocols for secure authentication that utilize accelerometer data on devices that are shaken together. For sufficient entropy in the generated feature vectors that are utilized as a basis to the key generation, the devices have to be shaken together for a sufficiently long time [3]. This procedure is feasible only for small mobile devices and requires direct access to both devices in order to pair them. Furthermore, the shaking of devices can not be considered unobtrusive. This Candidate Key Protocol generates feature sets with identical hash values on both devices and concatenates them to create a shared secret key among devices [4]. In [5], a false negative rate of more than 11 percent was reported for this approach and for features extracted from accelerometer data.



(a) Encoding maps message $a \in \mathcal{A}$ to codeword $c \in \mathcal{C}$. Decoding maps several $\tilde{c} \in C, C = \{c, c', c'', \dots\}$ to one $a \in \mathcal{A}$



(b) Similarity of fingerprints from different (squares) and identical (circles) audio

1 Key generation by audio fingerprinting

Through audio fingerprinting short characteristic bit sequences of an audio stream are created [6, 7]. We propose to extract audio fingerprints from synchronized recordings on mobile devices, possibly correct their Hamming distance with fuzzy cryptography schemes, and then to utilize these fingerprints as the seeds for cryptographic keys among devices.

Figure 1(a) illustrates this concept. A message a of length m is mapped uniquely to a specific codeword c of length n: $a \mapsto c : a \in \mathcal{A}, c \in \mathcal{C}$. We utilize error correcting codes for this mapping, so that similar codewords c' are decoded to the same massage a.

The protocol is able to generate secure keys based on ambient audio if and only if the audio samples are synchronized and taken in close proximity. In various case studies we experienced a good performance and robustness against differing noise levels for this protocol. It was implemented using parts of the Fuzzy Commitment scheme detailed in [8] together with Reed-Solomon codes [9]. The codes were applied in conjunction with the Secure Hash Algorithm with 256 bit (SHA-256).

However, a possible weakness are the audio fingerprints that are used as seeds for the secret keys. Clearly, if the entropy of the audio fingerprints were low, an adversary could use additional knowledge to increase the probability to correctly guess the secret key.

1.1 Similarity of audio fingerprints

In our implementation we split $r = 44100$Hz audio recordings of 6.375 seconds into $n = 17$ frames of identical length of 0.375 seconds. On each frame a discrete Fourier transformation weighted by a Hann window is applied. Frames are divided into $m = 33$ frequency bands on which the sum of the energy values is calculated and saved in an energy matrix E with energy per frame per frequency band. Using E, a 512 bit fingerprint f is generated, where every bit is calculated

as $\forall i \in \{1, \ldots, n-1\}, \forall j \in \{0, \ldots, m-2\}$

$$f(i,j) = \begin{cases} 1 & \text{if } E(i,j) - E(i,j+1) - \\ & (E(i-1,j) - E(i-1,j+1)) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

To classify the suitability of audio fingerprints as the basis for a classical key generation scheme we recorded 25 different samples with five samples for five distinct activities each. The activities considered are a music sample, a person clapping her hands, snipping, speaking and whistling. These samples were played back by an audio source and sampled again through two microphones in various distances from the source. We installed the microphones at 1.5m, 3m, 4.5m and 6m distances. The similarity of fingerprints is depicted in figure 1(b) exemplary for microphones at 1.5m and 3m distance. The image clearly shows, that the similarity is higher for identical audio samples (circles) than for dissimilar samples (squares). However, also identical samples produce fingerprints that differ in a considerable number of bits. These errors are then corrected by Reed-Solomon error correcting codes.

1.2 Entropy of audio fingerprints

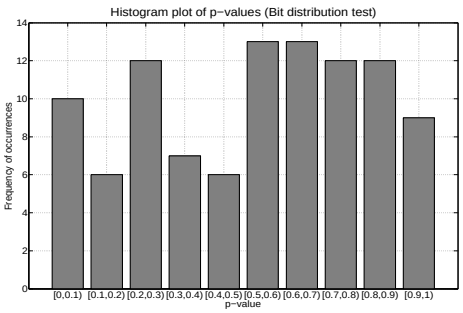
In order to test the entropy of generated fingerprints we applied the dieHarder [10] set of statistical tests. This battery of tests calculates the p-value of a given random sequence with respect to several statistical tests. The p-value denotes the probability to obtain the provided input sequence given a truly random bit generator. We applied all 24 statistical tests in the set on fingerprints of 480 bits length. In 7490 test-batches of 100 repeated applications of one specific test each, only 173 batches, or about 2.31% resulted in a p-value of less than 0.05.[1] Each specific statistical test was repeated at least 70 times. The p-values are calculated according to the statistical test of Kuiper [11].

Figure 1 depicts the proportion of p-values that pass a sequence of tests and an exemplary histogram for the RGB bit distribution test. For each of the 34 test series conducted, figure 1(d) depicts the fraction of tests that did not pass a sequence of 100 consecutive runs for all 107 distinct tests in the DieHarder battery of statistical tests.

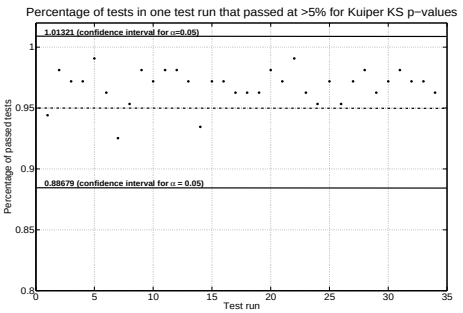
2 Conclusion

We detailed the unobtrusive creation of audio-fingerprint based cryptographic keys and presented a subset of our results achieved on the entropy of audio fingerprints. Generally, we observed that no significant weakness could be observed for any of the 24 different statistical tests utilized.

[1] All results are available at `http://www.ibr.cs.tu-bs.de/users/sigg/StatisticalTests/TestsFingerprints_110601.tar.gz`



(c) Histogram plot of p-values for the RGB Bit Distribution test



(d) Proportion of sequences passing a test

Fig. 1. P-Values for the fingerprints obtained

Acknowledgments. This work was supported by a fellowship within the Postdoc-Program of the German Academic Exchange Service (DAAD)

References

1. Mayrhofer, R., Gellersen, H.: Spontaneous mobile device authentication based on sensor data. information security technical report **13**(3) (2008) 136–150
2. Mayrhofer, R., Gellersen, H.: Shake well before use: Intuitive and secure pairing of mobile devices. IEEE Transactions on Mobile Computing **8**(6) (2009)
3. Mayrhofer, R.: The Candidate Key Protocol for Generating Secret Shared Keys from Similar Sensor Data Streams. Security and Privacy in Ad-hoc and Sensor Networks (2007) 1–15
4. Mayrhofer, R., Gellersen, H.: Shake well before use: Authentication based on accelerometer data. Pervasive Computing (2007) 144–161
5. Bichler, D., Stromberg, G., Huemer, M., Loew, M.: Key generation based on acceleration data of shaking processes. In Krumm, J., ed.: Proceedings of the 9th International Conference on Ubiquitous Computing. (2007)
6. Wang, A.: An Industrial Strength Audio Search Algorithm. In: International Conference on Music Information Retrieval (ISMIR). (2003)
7. Wang, A.: The Shazam music recognition service. Communications of the ACM **49**(8) (2006) 48
8. Juels, A., Wattenberg, M.: A Fuzzy Commitment Scheme. Sixth ACM Conference on Computer and Communications Security (1999) 28–36
9. Reed, I., Solomon, G.: Polynomial codes over certain finite fields. Journal of the Society for Industrial and Applied Mathematics (1960) 300–304
10. Brown, R.G.: Dieharder: A random number test suite. http://www.phy.duke.edu/ rgb/General/dieharder.php (2011)
11. Kuiper, N.: Tests concerning random points on a circle. In: Proceedings of the Koinklijke Nederlandse Akademie van Wetenschappen. Volume Series a 63. (1962) 38–47