Security Analysis of Quantization Schemes for Channel-based Key Extraction

Christian T. Zenger, Jan Zimmer, Christof Paar
Horst Görtz Institute for IT-Security (HGI), Ruhr-University Bochum, Germany
{christian.zenger, jan.zimmer, christof.paar}@rub.de

ABSTRACT

The use of reciprocal and random properties of wireless channels for the generation of secret keys is a highly attractive option for many applications that operate in a mobile environment. In recent years, several practice-oriented protocols have been proposed, but unfortunately without a sufficient and consistent security analysis and without a fair comparison between each other. This can be attributed to the fact that until now neither a common evaluation basis, nor a security metric in an on-line scenario (e.g., with changing channel properties) was proposed. We attempt to close this gap by presenting test vectors based on a large measurement campaign, an extensive comparative evaluation framework (including ten protocols as well as new on-line entropy estimators), and a rigorous experimental security analysis. Further, we answer for the first time a variety of security and performance related questions about the behavior of 10 channel-based key establishment schemes from the literature.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: Security and Protection

Keywords

Channel-based key extraction, physical layer security, practice-oriented protocols, quantization schemes, on-line entropy estimation, security analysis

1. INTRODUCTION

Key agreement systems based on correlated observations, such as common channel estimations, have received much attention in recent years due to their information-theoretical security properties. Compared to traditional public-key based approaches they can have a lower complexity and are resistant against (future) quantum computer attacks. Channel-based key establishment (CBKE) introduced by Hershey et al. [9] is a new approach for generating symmetric secret keys for two (or more) wireless communicating parties, where-by the key material is unique for a given point in time and space, and cannot be easily obtained by others. One major advantage of CBKE is the new possibility for dynamic key management without a complex infrastructure (e.g., public-key infrastructure or Kerberos-like key servers). In addition to the reduced system complexity, CBKE also has the advantage of not having a single point of failure. It has many possible uses for wireless communication devices and has also been investigated as a potential lightweight solution for cyber-physical and Internet of Things (IoT) devices, which are primarily formed by (small) embedded systems. However, it has been found that the security features of some CBKE designs are too weak or are based on broad channel abstractions which are not realistic.

The classical system model for channel-based symmetric key extraction schemes is based on the following scenario. Two *keying nodes*, Alice A and Bob B, plan to extract a symmetric key for secure communication while an eavesdropper, Eve E, capable of observing information, e.g., information for error correction, and measuring the channels between herself and the two communicating nodes, tries to recover the secret key (please refer to Figure 1 for illustration). We assume that A and B do not share any mutual information (e.g., shared keys) a priori.

The generic four-phase security architecture for generating secret symmetric keys from correlated random channel measurements is illustrated in Figure 2. Commonly measured channel profiles are quantized into bit vectors to obtain initial preliminary key material. The non-perfect reciprocity in measurement in addition with noise leads to errors in the bit vectors of the preliminary key material. These errors are detected and corrected during the information reconciliation stage by using error correcting techniques. Since information for error correction has to be exchanged over the channel during the information reconciliation stage, the remaining entropy is distributed over the key material in the privacy amplification stage.

A common way to evaluate the security of practice-oriented

MOBIQUITOUS 2015, July 22-24, Coimbra, Portugal

DOI 10.4108/eai.22-7-2015.2260194

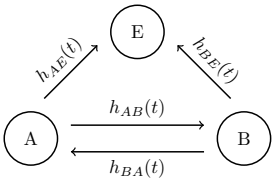


Figure 1: System model: Legitimate nodes A and B measure reciprocal properties of the physical channel, denoted by $h_{BA}(t)$ and $h_{AB}(t)$. A passive attacker E's observations $h_{AE}(t)$ and $h_{BE}(t)$ are dependent on its relative position and usually correlate less to $h_{BA}(t)$ and $h_{AB}(t)$ than $h_{BA}(t)$ to $h_{AB}(t)$.

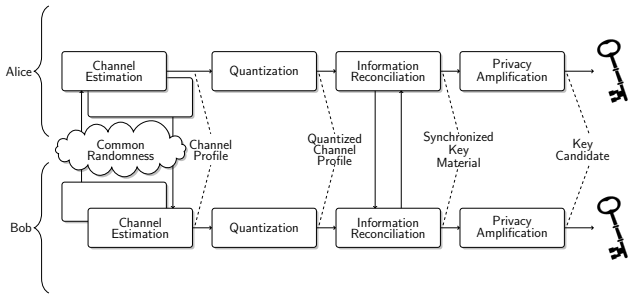


Figure 2: Overview of the core components involved in the security architecture for key agreement systems from correlated observations.

schemes in the past has been to perform (extensive) channel measurement campaigns with respect to legitimate participants and potential passive adversaries, succeeded with an analysis of the goodness of the preliminary key material applying off-line statistical tests [11, 10, 12, 1, 15]. However, although the low probability of success for potential passive attacks on quantization schemes was demonstrated based on individual experimental measurements (e.g., by [2, 3, 11, 10, 8, 12, 1, 15, 16]), little attention has been paid to a secure CBKE scheme against active attackers (i.e. active control over the environment, not actively altering the channel, e.g., by jamming) exploiting low entropy environments.

1.1 Contributions

We apply a security analysis of all (to the best of the authors' knowledge) published works on practice-oriented protocols for channel-based key extraction. We extend the classical adversary model [1] of two passive attackers (P1,P2) by an active adversary: (P1) The attacker measures corresponding channel profiles between legitimate parties and itself. Usually the assumption is given that (partial) access to the random source is dependent on the physical position of the attacker. (P2) The attacker exploits possible statistical defects in combination with public information (i.e., parameters of the scheme as well as eavesdropped communication). The distance from which the attacker can eavesdrop the communication on the channel has to be assumed big and may even work outside the connection range of network specifications. (A) The attacker is capable of manipulating the environment or to force one or both legitimate parties into an artificial environment, e.g., using a Faraday cage to artificially build a static scenario. The aim of the attacker is to determine the symmetric key material by exploiting statistical defects of the used key extraction scheme. The attacker thereby does not alter the channel directly, e.g., by jamming techniques.

Therefore, we present a framework to evaluate and compare channel-based key extraction schemes. Further, we present simulation results related to performance and security trade-offs and compare those with results from real-world measurements, which might help to select an appropriate quantization scheme. For substantiating analyses, an extensive experimental measurement campaign, covering several real-world use cases, was performed. We show that it needs to be assumed that changes in the environments are both statically and dynamically, which in turn influences the statistical agility (and so the goodness) of the source of randomness. We introduce security validation based on on-line entropy estimation as a possible approach to cover those security issues by constant evaluation of the goodness of the source of randomness as an extension to the off-line entropy estimation and statistical tests proposed in the literature.

1.2 Related Work

Jana et al. [10] and Premnath et al. [13] proposed the first analysis of several quantization schemes introducing entropy estimations. Both evaluated several schemes [11, 2, 14, 3, 10]. They analyzed the statistical properties off-line by calculating the Shannon entropy. As the entropy of the key material may decrease due to changes in the environment, solely applying such estimations off-line may not be sufficient for security applications.
A theoretical analysis of the 'BDR[2] vs. PEARSON correlation coefficient' behavior of several quantization schemes [14, 2, 11, 10, 1] was introduced by Guillaume et al. [7].

Our work extends those of [10, 13, 7] by (1) analyzing the 'BDR vs. PEARSON correlation coefficient' behavior of further recent practice-oriented protocols via simulation, (2) introducing a real-world-based evaluation strategy of 10 protocols, (3) demonstrating experimental validation of indoor channels decorrelation behavior, and (4) introducing results of on-line entropy estimation for security level verification.

[1] We do not consider impersonating or *man-in-the-middle* (MITM) attacks, which are based on unauthenticated relationships. We also do not consider *random number generator* (RNG)-manipulation attacks or *denial of service* (DoS) attacks. The adversary knows the whole key generation protocol and the values of the parameters.

[2] The *bit disagreement rate* (BDR) indicates the percentage of bits that are in disagreement between the initial key material of Alice and Bob. With decreasing BDR, the effort needed to detect and correct errors decreases as well. BDR is evaluated after quantization by the relation: $\mathrm{BDR} = \frac{b_e}{b}$, where b_e is the number of bits in the sequence that disagree and b is the length of the initial key.

2. PRACTICE-ORIENTED QUANTIZATION PROTOCOLS UNDER TEST

As mentioned above, both A and B quantize their channel profiles (usually a vector of *received signal strength indicator* (RSSI) samples) into digital vectors to obtain preliminary key material. A number of algorithms for quantization have been shown in the literature and are divided into two categories: lossless and lossy quantization schemes. Lossless quantization maps every sample to a n-bit symbol whereas lossy quantization schemes may drop certain samples in favour of a more robust key generation and to maintain a high bit entropy. The original intention was that the output stream could be used directly as a shared symmetric key without using posterior information reconciliation and privacy amplification.

Several *lossy* schemes are based on a guard interval $[q+, q-]$, where $q+$ and $q-$ denote the upper and lower threshold, respectively. Samples within the interval are dropped. The selection of scheme-parameters affects the output rate and probability of errors. Also, lossy schemes need to communicate to select samples and increase the robustness of their *key* generation. However, Edman et al. [6] stated that such reconciliation requires interaction that may result in potentially zeroing the conditional entropy [3] and Eberz et al. [5] successfully attacked a guard-band-based quantization scheme, demonstrating that strong characteristics may be observed by a passive adversary.

Lossless schemes do not lose valuable information due to the processing of the whole measurement and are not only based on strong channel characteristics. Therefore, possible channel prediction attacks do not compromise the entire key material. Further, some lossless schemes do not require communication and therefore impede the work of an attacker since no information on the internal process is revealed. Table 1 provides an overview on current schemes for physical-layer key extraction. The table is sorted chronologically.

[3]Here the conditional entropy is defined as the leftover entropy contained in the initial key material after revealing some information by communication of the quantization scheme.

Table 1: Details of the covered quantization schemes.

Name	Category	Comm.	Bit/sample
Tope et al. [14]	guard-band	$\rightarrow$	≤ 1
Aono et al. [2]	guard-band	$\leftrightarrow$	≤ 1
Azimi et al. [3]	lossless	-	1
Mathur et al. [11]	guard-band	$\leftrightarrow$	≤ 1
ASBG [10]	guard-band	$\leftrightarrow$	≤ 1
ASBG-multibit [10]	lossless	-	> 1
Hamida et al. [8]	guard-band	-	≤ 1
Patwari et al. [12]	lossless	$\rightarrow$	> 2
Ambekar et al. [1]	guard-band	$\leftrightarrow$	$0/2$
Zenger et al. [16]	lossless	-	> 1

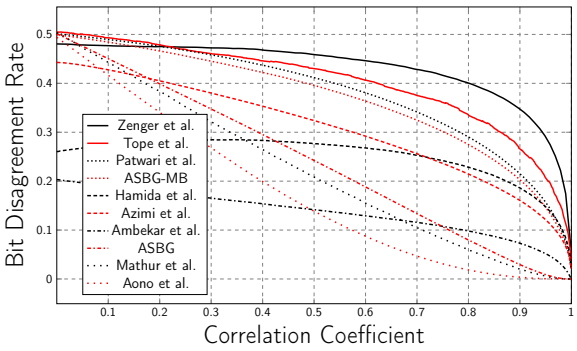


Figure 3: Bit disagreement rate versus correlation coefficient for different quantization schemes.

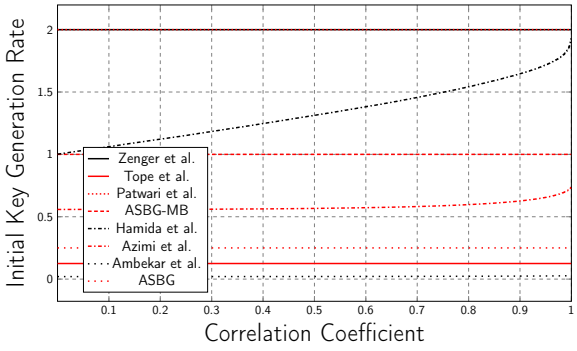


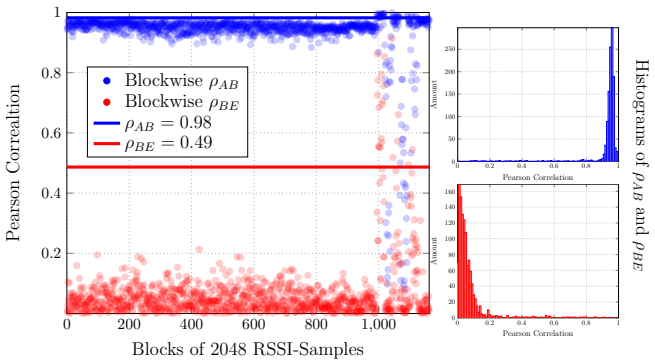
Figure 4: Initial key generation rate versus correlation coefficient for different quantization schemes.

To evaluate and compare quantization schemes, we implemented the 10 protocols in Matlab and applied a Monte-Carlo simulation environment, as proposed in [7]. Two independent random sequences of length 1 000 000 are modeled as temporally correlated Rayleigh-distributed random variables. The maximum Doppler shift specifies the assumed Jakes Doppler spectrum. To achieve a quantitive measure for the grade of reciprocity, we define $\rho_{\alpha\beta} \in [0, 1]$ as the correlation coefficient between the channel measurements of two nodes. Figure 3 and 4 present the simulation results of the performance analysis of the quantization schemes summarized in Table 1. Thereby, the change of BDR over correlation coefficient is shown as well as the change of the *initial key generation rate*[4] (IKGR) as the two major performance metrics. The IKGR $\mathrm{IKGR}_{\mathrm{quantizer}} = \frac{|q|}{|h|}$ is defined as the average ratio $fracqk$of the length of the emitted bit stream q and the number of samples provided as an input h.

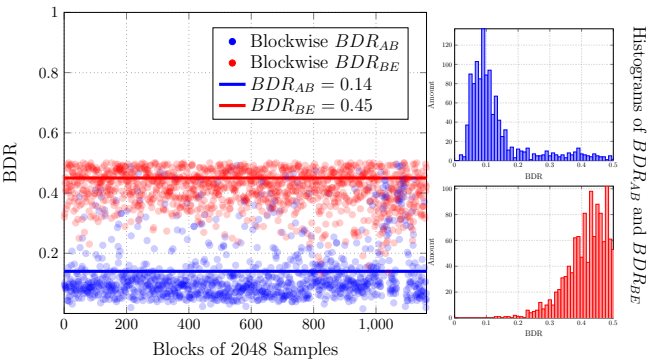
3. SECURITY ANALYSIS

We implemented a common channel measurement process on the hardware platform Raspberry Pi. This tiny computer is universally deployable with a Linux-based operat-

[4]We would like to note that *key* might be a miss-leading terminology. In our opinion the term *potential key material* comes closer to the real meaning.



(a) Correlation over time and its histograms.



(b) BDR over time and its histograms.

Figure 5: Evaluation results for: (a) correlation coefficient over time. Alice and Bob with strong correlated signals. After block 1000, the motion of the measurement platform stopped, with the result of low correlation. Eve and Bob achieve weak correlated signals. (b) BDR over time, applying ASBG-multibit quantization [10]. Alice and Bob with relatively low BDR; Eve and Bob with BDR around 0.5.

ing system and flexible expansion options. We equipped the computer with a TP-Link TLWN722N wireless USB adapter, utilizing IEEE802.11g and providing RSSI values on a per packet basis, as well as a battery for mobility. The setup is specifically designed to obtain synchronized measurements between three parties ($3\times$ bidirectional estimations) within $r_p^{-1} \leq 5\,\mathrm{ms}$ with a sampling interval of $r_s^{-1} \approx 10\,\mathrm{ms}$.

In the present paper, we focused on the following indoor setups: A was positioned on a *randomly moving robotic platform* (RMRP). A and B are separated by a distance of approximately $10\,\mathrm{m}$ with a standard deviation of $4.6\,\mathrm{m}$. To analyze the performance of an attacker in relation to his proximity in a continuous manner, we let a potential passive attacker E constantly move towards B from a distance of $30\,\mathrm{cm}$ to $0\,\mathrm{cm}$ with a speed of 7 centimeters per hour. Thereby, the change of correlation over distance is evaluated.

3.1 Results

To compare the performance of different algorithms, we agree on mutual metrics. In this section, we will provide an overview on the metrics that will be used for our evaluation.

In the past, researchers (e.g., [2, 12, 8]) have estimated the reciprocity of their gathered channel data using the PEARSON correlation coefficient ρ. Because in our opinion one single correlation value over the entire measurement is not convincing (please refer to Figure 5(a)), we evaluate channel reciprocity by computing the Pearson correlation coefficient blockwise. This gives us the ability to evaluate the reciprocity over time. This is advantageous since channel data may include stationary as well as dynamic sequences which is reflected using this visualization. The protocol processing is as well done blockwise - therefore the metric of choice for practice oriented evaluation may be in blocks.

The potential success of a passive attacker E (estimating the channel from A to B) was evaluated for a distance of $30\,\mathrm{cm}$, as described before. For the metric of success we

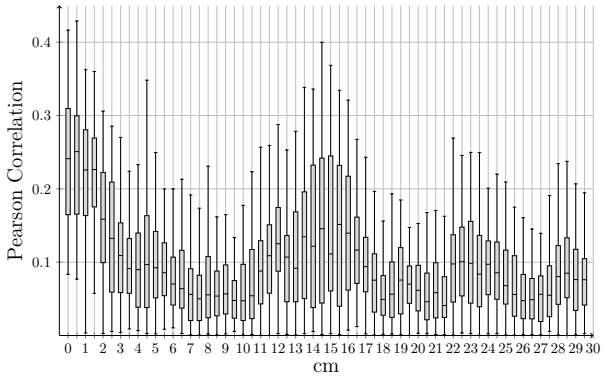
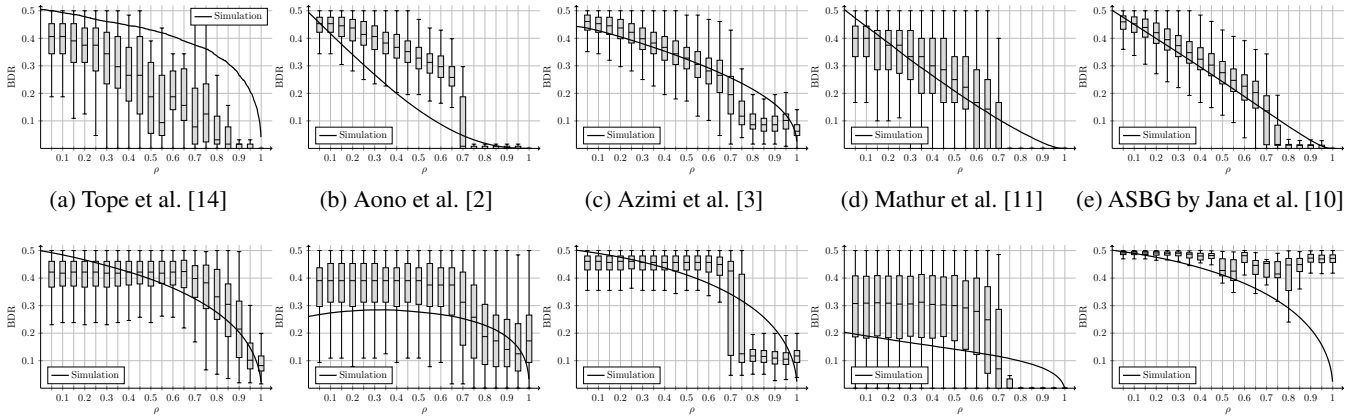


Figure 6: Correlation over distance for a potential attacker.

again chose the blockwise PEARSON coefficient. The evaluation results are illustrated in Figure 6. Very characteristic are the recurring correlation rises at a distance of approximately $\lambda/2$ and λ, where $\lambda = 12.5\,\mathrm{cm}$ is the wavelength of the carrier. Nevertheless, a potential attacker in our scenario achieves a correlated observation of the channel profiles with a PEARSON coefficient not higher than 0.45.

To achieve initial key material with a maximum of secret information concerning a passive adversary, the BDR between a potential eavesdropper and Alice or Bob should be 0.5. We present a blockwise evaluation approach in Figure 5(b). Based on all data of the large measurement campaign (with approximately $112\,\mathrm{hours}$ of measurement and $40\,000\,000$ RSSI values per channel) we evaluated the BDR versus the correlation coefficient. Therefore, we calculated the blockwise correlation as well as the corresponding blockwise BDR and sorted those by correlation value. Further, we calculated the BDR distribution for the following subgroups: $[0 : 0.05, 0.05 : 0.1, ..., 0.95 : 1]$. Figure 7(a)-(j) displays the distribution of the blockwise BDR vs. PEARSON correlation



(a) Tope et al. [14] (b) Aono et al. [2] (c) Azimi et al. [3] (d) Mathur et al. [11] (e) ASBG by Jana et al. [10]

(f) ASBG-multibit [10] (g) Hamida et al. [8] (h) Patwari et al. [12] (i) Ambekar et al. [1] (j) Zenger et al. [16]

Figure 7: Evaluation results based on real-world measurements for different quantization schemes. The bit disagreement rate versus correlation coefficient ρ is presented.

coefficient of the preliminary key material. For better comparison we also plotted the simulation results from Section 2.

The BDR distributions of the real-world measurements are almost always similar to the pattern of the simulation. Stronger differences are given for the scheme of Aono et al. [2] and Zenger et al. [16]. Note that increasing the number of evaluated blocks may further improve the results. Further, our results show that the schemes by Hamida et al. [8] and Ambekar et al. [1] do not lead to a BDR of 0.5 for low correlated channel profiles. This leads to a serious security problem, in so far as even with totally uncorrelated measurements an adversary may recover partial information of the initial key material. The schemes introduced by Jana et al. [10], Azimi et al. [3] and Mathur et al. [11] show almost linear decrease of BDR with increasing correlation. As this linear behavior lets an adversary learn a lot of secret information already with reasonably good correlations, this could corrupt the security of the scheme. The scheme of Zenger et al. [16] shows strong security properties for potential attackers with low correlated observations, however it also shows very bad performance for highly correlated channel profiles. The scheme by Tope et al. [14] shows good simulation results, however, it produces wide-spreaded real-world values which prevent the tight definition of boundaries. The scheme by Aono et al. [2] shows different behavior between simulation and real-word evaluation. In the real-world scenario, the scheme seems to have pretty good performance. This is also true for the schemes by Jana et al. [10] and Patwari et al. [12] The shape of the distributions of schemes show the wanted results: high BDR for low correlations and low BDR for high correlations without a fluent transition between both. The scheme by Patwari et al. [12] seems to have the best property with its abrupt change of BDR at a high correlation as it means that an adversary with lower correlation will not learn any information while both parties gather almost the same information if they have relatively correlating channel measurements.

The results of the on-line bit entropy estimation of the preliminary key material over time are given in Figure 8(a). Here the draft $800 - 90$B of NIST [4] is applied and the worst-case estimation result of the five tests is used, as recommended by the draft. We investigated if those tests are applicable in an on-line scenario in terms of performance and implemented them. As we already stated, the on-line estimation ensures the current security level by responding to statistical defects of the material. As an example we evaluated the influence of the channel sampling rate.
Because of the high channel sampling rate r_s of $100\,\mathrm{Hz}$, the key material is highly correlated in time and therefore the estimated entropy is relatively low. Reducing the sampling rate (applying downsampling in our framework) helps to find the optimal sampling rate at the maximum estimated entropy as demonstrated in Figure 8(b). Interesting to mention are the different behaviors of the different quantisation schemes. This may be because of statistical defects of the schemes itself or because, the amount of input material, and/or the sample selectivity of the (lossy) scheme.

4. CONCLUSION

Prior work has documented the effectiveness of CBKE systems in improving key generation rates and reducing bit disagreements; Jana et al. [10], for example, report the first evaluation results of five key extraction approaches. Although these practical works are security-motivated, their evaluation strategies follow rules which are based on channel models for (robust) communication engineering and miss important security requirements, such the on-line evaluation of the goodness of the entropy source with respect to the continuously differing environment. In this study we give a comparative overview and demonstrate for the first time evaluation results of all (to the best of the authors' knowledge) practice-oriented key extraction systems from the lit-

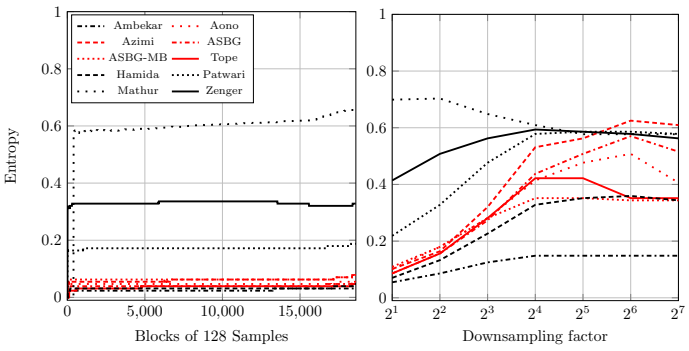


Figure 8: Evaluation results based on the real-world setups $1 - 12$ of all quantization schemes for (a) estimated min-entropy over time (b) on-line entropy estimation for different downsampling factors.

erature. Therefore, we introduced on-line entropy estimation as an evaluation basis. Our results include 40 million common channel estimations for a three party model from an extensive measurement campaign. In addition, we stated general problems of key agreement systems, presented a set of criteria for selecting quantization schemes, and demonstrated the results of a rigorous evaluation strategy. We found that in two cases the quantization schemes lead to a significant statistical defect and named promising schemes.

5. REFERENCES

[1] A. Ambekar, M. Hassan, and H. D. Schotten. Improving channel reciprocity for effective key management systems. In *ISSSE*. IEEE, 2012.

[2] T. Aono, K. Higuchi, T. Ohira, B. Komiyama, and H. Sasaoka. Wireless secret key generation exploiting reactance-domain scalar response of multipath fading channels. *Antennas and Propagation, IEEE Transactions on*, 53(11):3776–3784, 2005.

[3] B. Azimi-Sadjadi, A. Kiayias, A. Mercado, and B. Yener. Robust key generation from signal envelopes in wireless networks. In P. Ning, S. D. C. di Vimercati, and P. F. Syverson, editors, *Proceedings of the 2007 ACM Conference on Computer and Communications Security, CCS 2007, Alexandria, Virginia, USA, October 28-31, 2007*, pages 401–410. ACM, 2007.

[4] E. Barker and J. Kelsey. NIST DRAFT Special Publication 800-90b recommendation for the entropy sources used for random bit generation. 2012.

[5] S. Eberz, M. Strohmeier, M. Wilhelm, and I. Martinovic. A practical man-in-the-middle attack on signal-based key generation protocols. In S. Foresti, M. Yung, and F. Martinelli, editors, *Computer Security - ESORICS '12, Pisa, Italy, September 10-12, 2012. Proceedings*, volume 7459 of *Lecture Notes in Computer Science*, pages 235–252. Springer, 2012.

[6] M. Edman, A. Kiayias, Q. Tang, and B. Yener. On the security of key extraction from measuring physical quantities. *CoRR*, abs/1311.4591, 2013.

[7] R. Guillaume, A. Mueller, C. T. Zenger, C. Paar, and A. Czylwik. Fair comparison and evaluation of quantization schemes for phy-based key generation. *OFDM 2014*, 2014.

[8] S. T. B. Hamida, J. Pierrot, and C. Castelluccia. An adaptive quantization algorithm for secret key generation using radio channel measurements. In K. A. Agha, M. Badra, and G. B. Newby, editors, *NTMS 2009, Cairo, Egypt*, pages 1–5. IEEE, 2009.

[9] J. E. Hershey, A. A. Hassan, and R. Yarlagadda. Unconventional cryptographic keying variable management. *IEEE Transactions on Communications*, 43(1):3–6, 1995.

[10] S. Jana, S. N. Premnath, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy. On the effectiveness of secret key extraction from wireless signal strength in real environments. In K. G. Shin, Y. Zhang, R. Bagrodia, and R. Govindan, editors, *Proceedings of the 15th Annual International Conference on Mobile Computing and Networking, MOBICOM 2009, Beijing, China, September 20-25, 2009*, pages 321–332. ACM, 2009.

[11] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik. Radio-telepathy: Extracting a Secret Key from an Unauthenticated Wireless Channel. In *MobiCom 2008*, pages 128–139. ACM, 2008.

[12] N. Patwari, J. Croft, S. Jana, and S. K. Kasera. High-rate uncorrelated bit extraction for shared secret key generation from channel measurements. *IEEE Trans. Mob. Comput.*, 9(1):17–30, 2010.

[13] S. N. Premnath, S. Jana, J. Croft, P. L. Gowda, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy. Secret key extraction from wireless signal strength in real environments. *IEEE Trans. Mob. Comput.*, 12(5):917–930, 2013.

[14] M. A. Tope and J. C. McEachen. Unconditionally secure communications over fading channels. In *Military Communications Conference, 2001. MILCOM 2001. Communications for Network-Centric Operations: Creating the Information Force. IEEE*, volume 1, pages 54–58. IEEE, 2001.

[15] C. T. Zenger, A. Ambekar, F. Winzer, T. Pöppelmann, H. D. Schotten, and C. Paar. Preventing scaling of successful attacks: A cross-layer security architecture for resource-constrained platforms. In *Advances in Cryptology - BalkanCryptSec '14, Istanbul, Turkey, October 16-17, 2014, Proceedings*.

[16] C. T. Zenger, M.-J. Chur, J.-F. Posielek, G. Wunder, and C. Paar. A novel key generating architecture for wireless low-resource devices. In *International Workshop on Secure Internet of Things (SIoT)*, volume 3, pages 74–89.