2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications: Track 1 - PHY and Fundamentals

Blind Twins: Siamese Networks for Non-Interactive Information Reconciliation

Paul Walther* and Thorsten Strufe†
*Chair for Privacy and Security, TU Dresden
†Karlsruhe Institute of Technology (KIT) and Centre for Tactile Internet, TU Dresden (CeTI)
{<firstname>.<lastname>}@tu-dresden.de

***Abstract*—Through Information Reconciliation, two legitimate parties of Channel Reciprocity-based Key Generation assure that they extract the same key from local channel measurements. Current protocols exchange messages: Interactivity both causes delays and energy expenditure, and leaks information about the keying material to adversaries.**

We suggest non-interactive reconciliation, using a Siamese Network of CNNs that extracts reciprocal and suppresses non-reciprocal components in the measurements. Training and evaluating on real-world and synthetic data, we demonstrate that it *blindly* achieves higher correlation of the outputs at legitimate parties than the *interactive* state of the art, thus eliminating cost and information leakage at superior performance.

I. Introduction

Channel Reciprocity-based Key Generation (CRKG) is a class of approaches to agree on a shared secret based on the properties of a radio channel. It is based on the two observations that (1) the volatile channel characteristics are symmetric, so close to identical for the two legitimate users Alice and Bob, yet (2) diverge widely for any third, observing party Eve. The channel hence is suggested as a shared source of secret randomness between the two parties who extract keying material to agree on a key.

Approaches implementing CRKG comprise of four steps. Both parties first perform *Channel Probing* to measure the characteristics, second *Quantization* of the corresponding signals, third *Information Reconciliation*, which we will look at in greater detail in this paper, to agree and identical components in the measurement, and fourth *Privacy Amplification*, to reduce the chance of eavesdroppers to correctly guess the shared secret.

Channel probing and quantization does not yield identical bit sequences at Alice and Bob. They do not necessarily perform probing at the exact same moment so the extracted channel characteristics may vary. Local interference has an additional effect, and the used transceivers both are not perfectly identical and can only perform noisy measurements. The bit sequences after these first two steps hence are different, even at the two legitimate participants.

The extent of divergence, but also the key generation rate depend on the type of channel characteristic that is used for CRKG. Early approaches employed the Received Signal Strength Indicator (RSSI). Due to its limited information content, their rates were limited to about $2-22\,bits$ per second [31]. Using Channel Impulse Responses (CIR) instead increases this rate. However, it also entails additional divergence of the detected bit sequences, as the effect of measurement errors grows [24], [13], [23].

Information Reconciliation (IR) then allows Alice and Bob to agree on identical and reject diverging components after quantization. Current state of the art suggests they exchange descriptions of their local results to identify reciprocal parts [3]. It inherently entails public communication between the legitimate parties, which is the source of the main energy consumption and delays within CRKG [11], and, even worse, leaks information about the keying material to the adversary.

We propose interaction-free information reconciliation, to eliminate these drawbacks. To overcome the necessity of exchanging information, we implement a machine learning approach that allows the legitimate parties to identify their reciprocal components locally. It reduces the impact of the observations at the eavesdropper at the same time.

We apply our information reconciliation to real world ultra-wideband (UWB) CIR measurements. Like related work we measure the performance by the fraction of bits that still diverge at the legitimate parties after IR (referred to as "Bit Disagreement Rate", BDR). Achieving a BDR of 0.003, our approach non-interactively outperforms CIR state of the art *with* interaction. Our results also indicate that an eavesdropper, even equipped with the same trained network, can still not reconcile into the same bit sequence, as her channel is inherently different.

In summary, we make the following *contributions*:

- we propose a novel non-interactive *Information Reconciliation* approach based on CNNs and Siamese networks
- we analyze the core parameters of this network and thereby define an effective instantiation
- we demonstrate effectiveness and security of our approach with comprehensive real world and synthetic UWB CIRs

The remaining paper is structured as follows: Sec. II will describe the system model and define the core problem. In Sec. III the state-of-the-art is outlined. Sec. IV presents the solution design of our new approach. In Sec. V we describe our concrete instantiation and show its performance. Sec. VI

This work has been supported by the German Research Foundation (DFG) through the Centre for Tactile Internet with Human-in-the Loop, EXC 2050.

978-1-7281-4490-0/20/$31.00 ©2020 IEEE

concludes and gives an outlook.

II. Background and Problem Statement

Although not restricted to it, our motivating use case is CRKG using UWB CIR, as described in [26]. In general, we base our work on the assumptions of the widely applied UWB multipath propagation model as defined in [8].

The **system model** is depicted in Fig. 1: Alice and Bob execute alternating transmissions and thereby obtain their respective estimations $\hat{h}_{AB}$ and $\hat{h}_{BA}$ of the shared channel. An eavesdropper Eve overhears both transmissions and thus obtains her own channel estimates $\hat{h}_{AE}$ and $\hat{h}_{BE}$. Following the assumption of reciprocity, the estimates $\hat{h}_{AB}$ and $\hat{h}_{BA}$ are highly correlated. Due to uniform scattering, the correlation with Eve's observation rapidly decreases the further she moves away from the exact location of Alice or Bob [8]. In practice, distances greater than $\lambda/2$ yield effectively uncorrelated observations [32].

Subsequent to this *Channel Probing* the following processing is necessary to derive a common key (see, e.g., [3]): By *Quantization* the obtained analog signals are transformed into bit strings. Subsequently, *Information Reconciliation* (IR) reduces divergence of results at Alice and Bob, originating from non-reciprocal interferences and hardware imperfections. This is attempted by exchanging certain information about the keying material without revealing the key itself. Finally, *Privacy Amplification* attempts to account for any information leakage towards potential attackers. The resulting bit string is considered a shared secret.

The secret key rate C_s of this key generation has an upper bound defined by [3]:

$$C_s \leq min(I(A;B), I(A;B|E)) \tag{1}$$

Here, $I(.,.)$ represents the mutual information between two nodes. This means, the secret key rate is restricted by the **information leakage** towards an eavesdropper E. Hence, to have a secure *and* efficient key generation, this information leakage has to be reduced as much as possible, ideally to zero.

The CRKG protocols have only two steps that leak information to an eavesdropper: at the shared randomness source itself and during information reconciliation through exchanging data about the preliminary keying material. Secrecy of channel characteristics to arbitrary third parties is the fundamental assumption of Physical Layer Security in general. Following the respective channel models and the decorrelation argument derived from Jake's Uniform Scattering Model [8], we choose to accept this assumption.

Thus, information leakage only occurs due to the exchange of information about the keying material during *IR*. Recent studies [15] have shown, that simple *IR* protocols leak up to 100% of information about the keying material and even recent approaches leak up to 30%.

The most robust way to guarantee zero information leakage is to not exchange any information about the preliminary keying material at all. Hence, in terms of information leakage, a completely interaction-free protocol, i.e. a protocol that refrains from message exchanges during IR, is optimal.

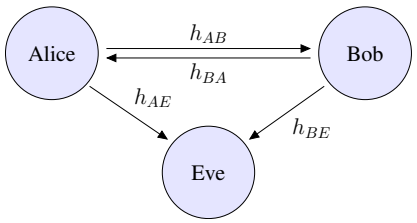


Fig. 1. Generic system model for CRKG: Alice and Bob measure the reciprocal channel and thereby obtain their estimates $\hat{h}_{AB}$ and $\hat{h}_{BA}$. Eve overhears this communication and estimates her own channels $\hat{h}_{AE}$ and $\hat{h}_{BE}$. Time-dependency $\hat{h}_{XY}(t)$ of the estimates is included in our system model, but the notion is omitted for brevity.

Besides information leakage, there are also **practical considerations** concerning the message exchange. Current CRKG implementations reach secret key generation rates of $2 - 22bit/s$ [12], [22]. As analyzed in [11], the messages transmission time during *IR* is one of the main latency drivers. Thus, reducing interactions and transmissions will inherently increase the achievable secret key rate. In accordance to the leakage property, a completely interaction-free solution would yield the best performance in terms of latency.

Finally, as CRKG is especially enticing for resource constrained devices [30], [11], the effective **energy usage** should be considered as well. Although CRKG itself requires considerably less energy than classical protocols, like ECDH, a major cause of energy consumption is exactly the message exchange throughout the IR protocol, causing up to $40,6\%$ of consumed energy [11].

We hence deem the development of a reliable, non-interactive Information Reconciliation scheme to be a worthy challenge: it reduces information leakage to adversaries and the communication overhead to the theoretically optimal minimum.

III. Related Work

Common approaches tackle IR by exchanging various information about the acquired randomness samples.

The simplest approaches are threshold-based. They include simple single threshold- [1] as well as guard-band based approaches [18]. Reconciliation is achieved by exchanging the set or a random subset of those values, which cross the threshold in certain manners. Transferring this set at least once in each direction, a common subset might be found, which is used for further processing.

Another group of approaches is based on Quantum Key Distribution: Extending the BBBSS protocol [2] and its improved version CASCADE [5], Jana et al. proposed an iterative reconciliation scheme [12]. They rely on dividing the keying material into blocks, about which the partners exchange parity information and try to adapt their bits until the parity matches.

Iterative approach like CASCADE typically last 30 rounds of bi-directional interaction.

Finally, the currently most prominent idea is to apply Error Correction Codes as reconciliation primitives as proposed by Zhang et al. [33]. Here, the key candidates are interpreted as code words of a linear block code and the respective decoding operation is executed. If decoded successfully, the observed error vector gets exchanged. By adding the received error vector to the own observation, the communication partner might shift its observation into the vicinity of the other observation and the subsequent decoding yields the same decoded word. There are proposals to use repetition codes [19], Hamming codes [6], BCH codes [25], [30], Reed-Solomon codes [33], Turbo codes [21] and LDPC codes [4], [17], which differ slightly in their achieved BDR and key-generation rates.

Several attempts to diminish the communication overhead by proposing uni-directional protocols have been made [7], [25], [20]. Nevertheless, the transmission of at least one message between the communication partners is required in all proposed schemes.

In summary, all existing approaches are based on interaction between the communication partners. They transmit messages containing information about the preliminary keying material – so all of them slow down the key exchange, cause energy expenditure, and leak information to the adversary. The idea of performing non-interactive, or blind, IR has not been pursued so far, to the best of our knowledge.

IV. Siamese Networks for Information Reconciliation

The **core idea** of our solution is to blindly extract the reciprocal channel characteristics that are unique to the positions of the legitimate partners. We train a machine learning model that distinguishes reciprocal components of legitimate measurements from those overheard by the adversary, for this purpose. The model can be trained in advance, once, and subsequently be used for blind information reconciliation. To make sure that Eve, even in possession of trained models, cannot approximate the sequences reconciled by Alice and Bob, we aim at extracting exactly those features, that represent the characteristics of the legitimate channel well.

We leverage two ML concepts for this purpose. For extraction, the task is to take unidimensional, sequential data and to output a sequence of bits. The input data at Alice and Bob will be subject to transformations, most importantly due to gain differences and temporal shifting [27]. According to literature, such feature extraction tasks are well performed using convoluational neural networks (CNNs) [14], [34]. Their real-valued output at the last layer can be quantized using a simple threshold function.

We also want to maximize the advantage of the extraction at legitimate parties versus extraction at an adversary in an arbitrary position (different to the exact location of Alice or Bob). We hence want to train the network to project correlated input sequences nearer, and decorrelated input sequences further apart in the output space. For this purpose, we train our CNN in a Siamese Network setup. In combination with a discriminating loss function, like *contrastive* [9] or *triplet* loss [10], this architecture is explicitly designed to enable the learning of discriminative features within a single network. Here, the base network learns to extract and use the unique features that represent the CIR reciprocity, and to disregard all others.

Together, the architecture of suitable CNNs trained in a Siamese Network with contrastive/triplet loss, provide the properties needed for interaction-free IR.

Our **concrete realisation** combines these concepts as follows: The Siamese network itself is instantiated by creating two of the above CNNs as Siamese twins with shared weights (they effectively are two views of the same network). To train the network with *contrastive loss*, the data is prepared by defining pairs of data, which are flagged as collected from reciprocal measurements or not (i.e. pairs of observations from either Alice and Bob or from Alice and Eve or Bob and Eve). Each of the Siamese CNN instances then processes one CIR of this pair. By feeding the output of both CNNs combined with the similarity flag to the *contrastive loss*, the shared base network learns to discriminate between the different input pairs. To train with *triplet loss*, no similarity flag but triplets of anchor (Alice), positive (Bob) and negative (Eve) samples and three Siamese instances are used to learn the discriminating features. After one-time training, the base CNN is deployed at the terminals and used to generate reconciled outputs without seeing the partners input. The core setup is depicted in Fig. 2.

An attacker with access to the CNN, e.g. a malicious insider, is still not capable of reconciling the same sequence as the overheard messages are different: the obtained h_{AE} is not reciprocal to h_{AB} (h_{BE} and h_{BA} accordingly). Hence, the output of the attackers CNN is different to those of the legitimate terminals[1].

A **deployment** then is done in the three following steps:

1) Instantiate the Siamese Network with an appropriate base CNN
2) Train the network once with *contrastive* or *triplet* loss
3) Deploy the base CNN at the terminal and use them for IR on live CIRs

As the CNN are trained to distinguish reciprocal from non-reciprocal features in the CIR locally, this information reconciliation does not require any interaction.

V. Evaluation

While it is clear that the designed information reconciliation approach is non-interactive, it remains to be evaluated, to which extent it yields identical, or at least similar bit sequences at

[1]In fact, the CIRs h_{AE} and h_{EA} would reconcile into the same output. But this is a valid key exchange scenario between reciprocal terminals and not an attack.

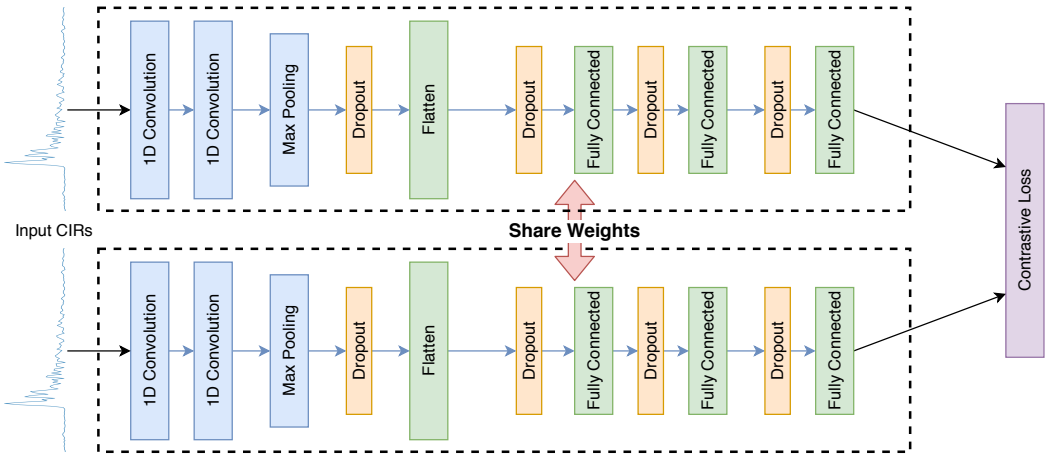


Fig. 2. The core setup of the proposed Siamese Network — the CNNs within the dotted lines share their weights and are essentially one and the same network. This CNN will become the resulting network deployed at CRKG nodes in practice.
For the *triplet loss*, 3 instances of the base CNN would be created and their respective outputs combined in the triplet loss function.

the legitimate parties, and how different the sequences are that adversaries can extract.

In the following, we first describe the data used to evaluate our approach, we test our instantiations with respect to the most relevant parameters after, and, finally, we will present the achieved results.

A. Training and Evaluation Data

The data for training and evaluation of the Siamese Network were obtained in two different ways: on the one hand we conducted extensive real world measurements to show the practical applicability of the approach, and on the other hand, we used synthetically generated CIRs to demonstrate that the approach generalizes well.

The **real world measurements** were obtain in a measurement campaign as described in [26]. Here, reciprocal UWB CIR measurements in a typical indoor environment were conducted, with an eavesdropping attacker observing all communication. The center frequency is $4\,GHz$ with $500\,MHz$ bandwidth and the sampling rate $1\,ns$, allowing a spatial resolution of $30\,cm$. Measurements were taken every $370\,ms$, with reciprocal experiments being synchronized to periods within $2\,ms$, well within the channel's coherence time. Overall, 12663 CIR pairs and adversarial observations were taken.

The measurements were set up in 7 different scenarios: 4 static (called $SA, \ldots, SD$) and 3 dynamic (IA, IB, ME). In SA-SC Alice, Bob, and Eve form an equilateral triangle with different rotations; in SD Eve resides right next to Bob. IA, IB incorporate additional movements: Eve moves right along the Line-of-Sight between Alice and Bob (IA), or perpendicular to it (IB). In ME the terminals move randomly in the room.

The data of each scenario was split into 70% training and 30% evaluation sets. All training data was then concatenated, permuted and used for training.

The **synthetic data** was created by adapting the deterministic Kunisch-Pamp channel model for UWB CIRs. Given a predefined environment and setting for transceiver position and properties, this model allows to deterministically generate noisy impulse responses with correlations similar to real world measurements. We used the obtained real world measurements to perform a Bayesian optimisation of the model parameters (cmp. [28]). Thereby, the resulting parameter set resembles the environment of the real world measurements. Given the model and this specific parameter set, we can create arbitrary CIRs for this environment.

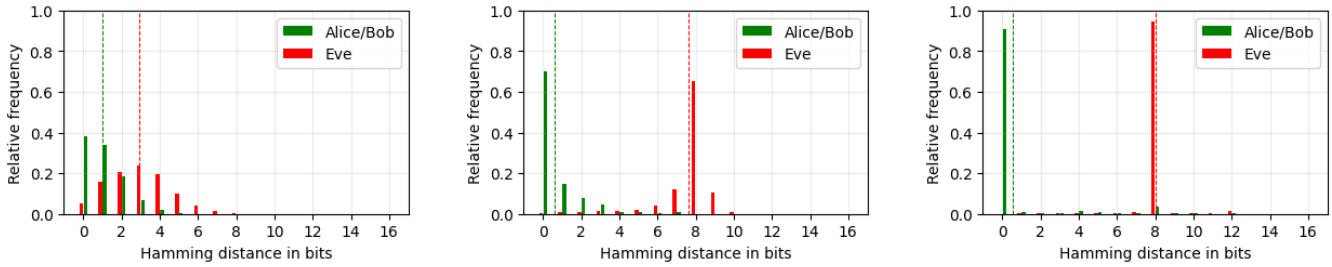
B. Instantiation of the Siamese Network

The architecture of the model has three core properties that are relevant for our evaluation: the base network with its respective architecture, the activation function of the final layer and finally the loss function used for learning.

For the base network, we implemented two different approaches: first, a rather simple network to test feasibility of the approach in general, and second, a more generalized network to show applicability.

Since the first network, **CNN1**, is intended solely to demonstrate the approach's feasibility, we omitted all measures for generalisations. The network is compromised of a single *1D convolutional layer* (1DConv) with 128 kernels of width 3, followed by 3 *fully connected layers* (FC) with 1024, 512 and 256 nodes, respectively. The final layer, which output we call embedding, is again *fully connected*, with the targeted embedding size chosen to be 16. All layers are activated as *rectified linear units* (ReLU). It is expected, that this network overfits the training data and is not generally applicable.

To demonstrate effective general applicability, we devise a second, more generalizing network **CNN2**. It consists of 2 *1DConvs* with respectively 32 and 64 kernels of width 3, followed by an 1D MaxPooling layer of size 2. After flattening, there are 2 FC layers with 256 nodes and the final FC layer sized to the embedding. We include a dropout with rate 0.5 before each FC layer (including the embedding layer). Again, all layers are ReLU activated.



(a) Contrastive Loss, D_W is Euclidean distance (b) Constrastive Loss, D_W as in (2) (c) Triplet Loss, D_W as in (2)

Fig. 3. Effectiveness of different **CNN1** parameter realizations. The dotted lines represent the BDR.

For both base networks, the activation of the embedding layer is especially important, due to its direct impact on the loss function. In the first iteration, this activation was set to ReLU as with the intermediate layers. This is reasonable when used in combination with the Euclidean distance or L2-Norm as distance function D_W.

Our goal, however, is to reconcile the input into a *binary* sequence, so we changed the activation to a *sigmoid* function, later. Restricting the final output to $[0, 1]$, the quantization is implicitly taken care of in this case: For a robust solution, we simply chose a threshold of $\theta = 0.5$ for quantization of the final embedding.

This quantization design invalidates the L2 norm as chosen distance metric. It also does not reflect the intention of the IR process and its final data: The main goal is to have equal *binary* sequences for the legitimate partners and different ones for the attacker. We hence would prefer the Hamming distance between the binary sequences to measure the final loss.

Including this quantization and Hamming distance calculation would result in a non-differentiable function, so learning would not be possible. Therefore, we chose the continuous variant of the Hamming distance D_W:

$$D_W(x, y) = y(1 - x) + (1 - y)x \tag{2}$$

This metric can now be optimized to reach 0 for reciprocal/similar pairs (Alice and Bob) and 0.5 for non-reciprocal/dissimilar pairs (Eve).

Finally, the loss function used to learn the separation is of essential importance. Following the reasoning in Sec. IV, the two candidates are *contrastive loss* (CL) and *triplet loss* (TL).

The **contrastive loss** is applied as

$$L(W, Y, \vec{X_1}, \vec{X_2}) = (1 - Y)\frac{1}{2}D_W^2 + Y\frac{1}{2}(\max(m - D_W, 0))^2 \tag{3}$$

Here, $\vec{X_i}$ are the respective outputs of the Siamese instances, D_W is $D_W(\vec{X_1}, \vec{X_2})$ and Y is the similarity flag. The margin m is set to 0.5, as we aim for a Hamming distance of 0.5 for "non-similar" pairs, i.e. Eves observations.

The **triplet loss** is realised as

$$L(W, \vec{X_A}, \vec{X_P}, \vec{X_N}) = \max(D_W^P - D_W^N + m, 0) \tag{4}$$

D_W^P denotes the distance between the anchor $\vec{X_A}$ and the positive sample $\vec{X_P}$, D_W^N is the distance between the anchor and the negative sample $\vec{X_N}$. The margin m is again set to 0.5 with the same reasoning as above.

C. Results

For our evaluation we employ the widely used metric of BDR [11]. As we have a constant vector length, the BDR is equal to the average Hamming distance between the binary vectors. We will depict the histogram of achieved Hamming distances to convey their actual distribution, instead of the mere average. The length of our embedding vector, $16\,bit$, is also the maximum distance, i.e those of two inverted sequences. For an attacker, the worst case would be a distance of 0.5, in our case equal to a distance of $8\,bit$, because in this case the attacker can only guess which of his bits are correct. In the plots, the Y-axis always shows the relative frequency and the X-axis the Hamming distance in bits; the dashed line is the BDR.

1) Different Network Parameters and Losses: We show the evaluation results for different parameters of **CNN1** in Fig. 3. The interpretation of these results are twofold:

First, the plots demonstrate the general capability of the proposed network architecture to differentiate between CIRs of legitimate partners and those of an eavesdropper. Especially, Fig. 3(c) depicts the clear discriminative strength of the network: the learned embeddings of the legitimate partners Alice/Bob have very low Hamming distances (0.033 BDR), whereas the eavesdropper Eve has a Hamming distance close to 0.5 (0.469 BDR) for her observations.

Second, it shows the influence of the different described network parameters. Subfigure 3(a) shows contrastive loss with Euclidean distance as D_W. The general effectiveness of the network is visible, but not strongly expressed: the legitimate partners reconcile to the same bit sequence in only 38% of cases, and the distribution of the attacker, albeit clearly distinct, is still very close. The network in subfigure 3(b) employs the same loss, but with the continuous Hamming instead of the Euclidean distance. This clearly increases the separation between the legitimate and the adversarial observations: Alice/Bob have a BDR of 0.037, whereas Eve is restrained to a BDR of 0.475. Finally, subfigure 3(c) shows the increased performance of applying triplet loss: separation is significantly

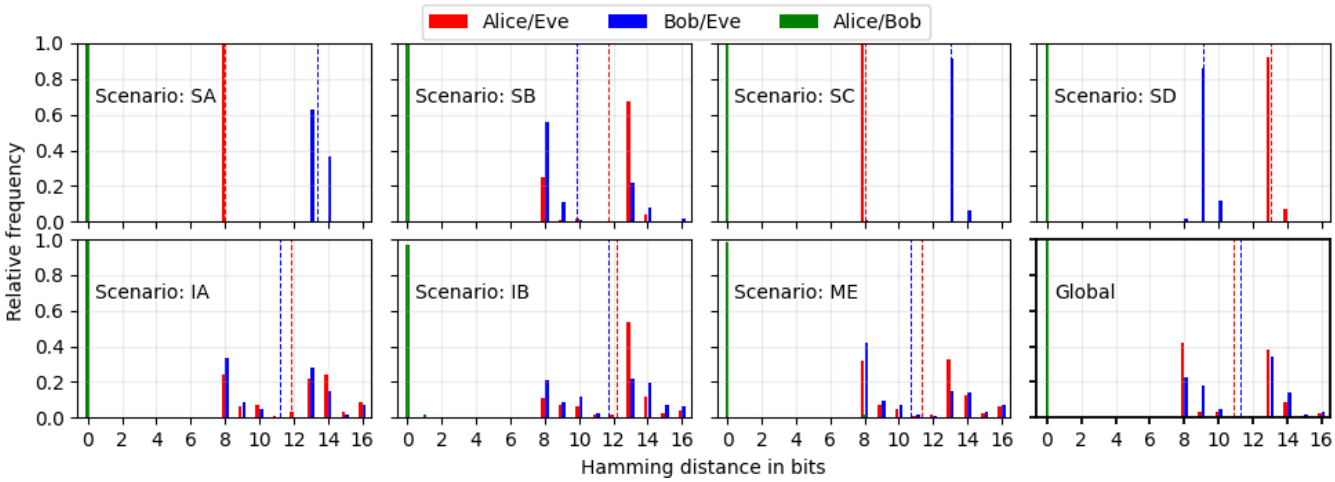


Fig. 4. The histograms of achieved Hamming distance after application of the trained model. The dotted lines represent the BDR.

better, with BDRs of 0.033 and 0.499, respectively, and a decreased standard deviation, e.g. from 0.087 to 0.049 for Eve.

It is worth noting that with **CNN1**, the attacker sometimes achieves a Hamming distance of 0, i.e. extracts the same binary sequence as Alice/Bob. This originates from the simple architecture and overfitting of **CNN1**. The more powerful and generalizing network **CNN2** completely removes this artifact as we will show in the next section.

2) General Real World Performance: Network **CNN2** was then used with continuous Hamming distance and triplet loss to evaluate the approach's applicability in real world scenarios.

The results of the evaluation with the real world measurements are depicted in Fig. 4. The lower right plot is a summary of all scenarios. The most notable outcome is that for all cases the reconciled sequences of the legitimate partners are equal in nearly all cases $(99,7\%)$, whereas the attacker reach at most a distance of 0.5. This means, that the legitimate partners robustly are reconciling into the same sequence, while the attacker is unable the gain any significant insight into this sequence, despite full knowledge of the used trained network.

Using the triplet loss with a margin $m = 0.5$, we expect the following: First, through the "pull" of the positive samples, the reciprocal observations will have a Hamming distance of 0. Second, since only negative samples with distance $< m$ contribute to the loss, there will be few attacker results lower than 0.5, i.e. $8\,bits$, because these are "pushed" to the upper half of the histogram. This should be particularly apparent in static scenarios, as these yield relatively stable observations. Finally, due to the generalizing measures, the results will be tolerant of movement and interference, which will be particularly visible in scenarios with such characteristics.

The evaluation results of our **static scenarios** accurately confirm these expectations: Alice and Bob achieve a BDR of 0, i.e. Hamming distance of 0 in all cases, with an overall BDR of 0.003. Only scenarios *SB* and *SC* deviate slightly, with a BDR of 0.005. The attacker achieved an average hamming distance of 0.692, while no result has a distance lower than $8\,bit$. As the static scenarios are also static for Eve, her reconciled sequences have rather stable distributions.

The results of the **dynamic scenarios** verify the general effectiveness: Despite the presence of unfavorable movements[2], the legitimate partners reach perfect reconciliation with frequencies of 1.0 (IA), 0.97 (IB) and 0.99 (ME). The attackers observations again are located in the histograms "upper" half, with a BDR of 0.717. Due to the additional interferences, the attackers binary sequences are much more scattered than in the static scenarios.

Additionally to the measurements, we used the *Kunisch-Pamp* channel model for **synthetic attacks** to rule out that we missed an advantageous attacker position: We used the real-world measurements to optimize the parameter of the channel model to our measurement environment. Then we generated attacker observations for all positions in this room, in steps of at most $\lambda/2$, using this optimized model. This synthetic attack data was then again processed with the trained network. The results are shown in Fig. 5. The average Hamming distance for this synthetic attacks is 0.43. It is again visible, that in no positions the attacker reconciles into the same sequences as Alice/Bob. Nevertheless, the distribution of the histograms indicates that there are in fact more advantageous positions for the attacker. These might be positions in very close vicinity of the legitimate partners or positions where multipath clusters are shadowing each other.

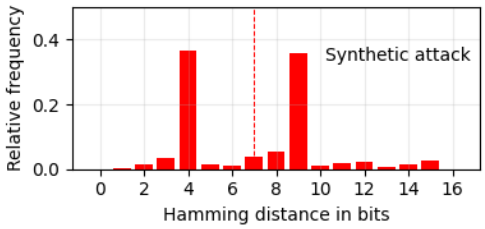


Fig. 5. The histograms of Eves Hamming distances for synthetic attack data.

[2]Movements per se are beneficial for CRKG, as they generate entropy to successive CIRs. Our experiment also includes detrimental movement, like Eve moving directly on the Line-of-Sight between Alice/Bob.

Overall, these results show that the trained network effectively implements a non-interactive Information Reconciliation. The synthetic attack even shows that attackers with knowledge of the network in especially favorable positions are not capable of obtaining the same sequence as the legitimate partners. Further, the average entropy of the single reconciled bits is $0.99\,bit$. Hence, the network has not learnt something static, but in fact extracts the reciprocal randomness. The overall success rate for the proposed IR is 0.992, which in turn can be represented as an BDR of 0.003. As invalid IR is a valid CRKG protocol outcome [3], successful reconciliation in 99.2% of all interaction cases is a very good result. Compared to state of the art CIR solutions, our results are very competitive: current solutions reach as low as 84% for successful IR [29]. But even high performing approaches like [16] reach at most a BDR of 0.004, which is still higher than the 0.003 achieved by our approach. This means, our solution is, in terms of effectiveness, at least as good as state of the art solutions. Additionally it is completely interaction free, i.e. it leaks no information at all and can be processed quicker.

VI. Summary and Outlook

In this work we tackled the question, how *Information Reconciliation* (IR) can be performed in a completely interaction-free manner to remove any information leakage, to inherently increase processing speed, and to reduce overhead.

We proposed a solution based on Siamese Networks towards this end, combining CNNs with contrastive and triplet loss. The CNNs are trained to extract the unique reciprocal properties of the input data. Such a pre-trained network can be readily used in practical applications to perform robust IR.

To show the effectiveness of our approach, we used extensive UWB CIR measurements to perform reconciliation in the context of Channel Reciprocity-based Key Generation (CRKG). We show that our approach can achieve Bit Disagreement Ratios as low as 0.003, while being completely interaction free. Additionally, we showed that an attacker with access to the trained network cannot reconcile the same sequences as the legitimate partners. On average the attacker achieves a BDR of 0.499%, i.e. she has no better chance than guessing each bits.

We conclude that our approach achieves successful IR, eliminates leakage of information about the keying material for CRKG to the adversary effectively, and reduces the overhead in terms of communication and time on top.

In current work, we mainly pursue two follow-up questions. First of all we investigate, to which extent alternative attacks, like training a specific adversarial model, could yield better chances for Eve. Some constellations in the synthetic setting yielded lower BDR for Eve, so we analyze reasons and design additional protection for these cases. Despite its performance, we are also working on improving our design both with respect to chosen parameters as well as types of networks.

References

[1] B. Azimi-Sadjadi *et al.*, "Robust key generation from signal envelopes in wireless networks," in *ACM CCS*, 2007.

[2] C. Bennett *et al.*, "Experimental quantum cryptography," *Journal of cryptology*, vol. 5, no. 1, pp. 3–28, 1992.

[3] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[4] M. Bloch, A. Thangaraj, and S. W. McLaughlin, "Efficient reconciliation of correlated continuous random variables using ldpc codes," *arXiv preprint cs/0509041*, 2005.

[5] G. Brassard and L. Salvail, "Secret-key reconciliation by public discussion," in *Eurocrypt*, 1993.

[6] W. T. Buttler *et al.*, "Fast, efficient error reconciliation for quantum cryptography," *Physical Review A*, 2003.

[7] Z. Feng and L. Jingling, "Performance of an improved one-way error reconciliation protocol based on key redistribution," *China Comm.*, 2014.

[8] A. Goldsmith, *Wireless Communications*. Cambridge Univ. Press, 2005.

[9] R. Hadsell, S. Chopra, and Y. LeCun, "Dimensionality reduction by learning an invariant mapping," in *CVPR*, 2006.

[10] E. Hoffer and N. Ailon, "Deep metric learning using triplet network," 2014.

[11] C. S. F. Huth, "Physical-layer security architectures for the internet of things," Ph.D. dissertation, Ruhr University Bochum, Germany, 2018.

[12] S. Jana *et al.*, "On the effectiveness of secret key extraction from wireless signal strength in real environments," in *Proc. of the 15th International Conf. on Mobile computing and networking*, 2009, pp. 321–332.

[13] H. Koorapaty, A. Hassan, and S. Chennakeshu, "Secure information transmission for mobile radio," *Comm. Letters, IEEE*, 2000.

[14] H. Lee, P. Pham, Y. Largman, and A. Y. Ng, "Unsupervised feature learning for audio classification using convolutional deep belief networks," in *Advances in Neural Information Processing Systems*, 2009.

[15] G. Li, Z. Zhang, Y. Yu, and A. Hu, "A hybrid information reconciliation method for physical layer key generation," *Entropy*, vol. 21, 2019.

[16] H. Liu, Y. Wang, J. Yang, and Y. Chen, "Fast and practical secret key extraction by exploiting channel response," in *IEEE INFOCOM*, 2013.

[17] J. Martinez-Mateo, D. Elkouss, and V. Martin, "Interactive reconciliation with low-density parity-check codes," in *IEEE ISTC*, 2010.

[18] S. Mathur *et al.*, "Radio-telepathy: extracting a secret key from an unauthenticated wireless channel," in *MobiCom*, 2008.

[19] U. M. Maurer, "Secret key agreement by public discussion from common information," *Trans. Inf, Theory*, 1993.

[20] Y. Minsky, A. Trachtenberg, and R. Zippel, "Set reconciliation with nearly optimal communication complexity," *Trans. Inf. Theory*, 2003.

[21] K.-C. Nguyen, G. Van Assche, and N. J. Cerf, "Side-information coding with turbo codes and its application to quantum key distribution," *arXiv preprint cs/0406001*, 2004.

[22] N. Patwari *et al.*, "High-rate uncorrelated bit extraction for shared secret key generation from channel measurements," *IEEE Trans. Mob. Comp.*, 2009.

[23] A. Sayeed and A. Perrig, "Secure wireless communications: Secret keys through multipath," in *IEEE ICASSP*, 2008.

[24] E. H. Shehadeh *et al.*, "Intelligent mechanisms for key generation from multipath wireless channels," in *IEEE WTS*, 2011.

[25] W. Traisilanun, K. Sripimanwat, and O. Sangaroon, "Secret key reconciliation using bch code in quantum key distribution," in *IEEE ISCIT*.

[26] P. Walther *et al.*, "Improving quantization for channel reciprocity based key generation," in *IEEE LCN*, 2018.

[27] P. Walther, E. Franz, and T. Strufe, "Blind synchronization of channel impulse responses for channel reciprocity based key generation," in *IEEE LCN*, 2019.

[28] P. Walther, R. Knauer, and T. Strufe, "Passive Angriffe auf kanalbasierten Schlüsselaustausch," *SICHERHEIT*, 2020.

[29] M. Yuliana *et al.*, "A simple secret key generation by using a combination of pre-processing method with a multilevel quantization," *Entropy*, 2019.

[30] C. Zenger, "Physical-layer security for the internet of things." Ph.D. dissertation, Ruhr University Bochum, Germany, 2017.

[31] C. Zenger *et al.*, "Authenticated key establishment for low-resource devices exploiting correlated random channels," *ComNets*, 2016.

[32] J. Zhang, T. Q. Duong, A. Marshall, and R. Woods, "Key generation from wireless channels: A review," *IEEE Access*, vol. 4, 2016.

[33] J. Zhang, S. K. Kasera, and N. Patwari, "Mobility assisted secret key generation using wireless link signatures," in *IEEE INFOCOM*, 2010.

[34] B. Zhao, H. Lu, S. Chen, J. Liu, and D. Wu, "Convolutional neural networks for time series classification," *Journal of Systems Engineering and Electronics*, vol. 28, no. 1, pp. 162–169, 2017.