UNIVERSITY OF CALIFORNIA,
IRVINE

Physical Layer Security Key Generation for Fiber Optic and Inter-Vehicular Optical Communication Link

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Electrical Engineering

by

Imam Uz Zaman

Thesis Committee:
Professor Ozdal Boyraz, Chair
Professor Nader Bagherzadeh
Professor Mohammad Abdullah Al Faruque

2018

DEDICATION

To

my parents, my friends and my colleagues

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

It is a genuine pleasure to express my deep sense of thanks and gratefulness to my supervisor, Professor Ozdal Boyraz, whose expertise, guidance, feedback and support made is possible to complete my MS research project. I would also like to express my gratitude to Professor Abdullah Al Faruque for his prompt inspiration and advice. Furthermore, I would like to thank my thesis committee member, Professor Nader Bagherzadeh for his time to review this thesis document and provide his valuable opinion.

I am highly indebted and thoroughly grateful to my family for their endless support and encouragement throughout my study period. In addition, I am thankful to all my relatives and friends for believing in me. I am grateful to Optical Society of America and IEEE for publishing my work in their conferences and journals.

The text of this thesis is a reprint of the materials as it appears in the following papers:

"Polarization Mode Dispersion-Based Physical Layer Key Generation for Optical Fiber Link Security"[1] in Advanced Photonics Congress 2017, Optical Society of America.

"A physical layer security key generation technique for inter-vehicular visible light communication" in Advanced Photonics Congress 2017, Optical Society of America.

"Physical Layer Cryptographic Key Generation Utilizing PMD to Secure Optical Fiber Link", Journal of Lightwave technology, 2018.[2]

.

ABSTRACT OF THE THESIS

Physical Layer Security Key Generation for Fiber Optic and Inter-Vehicular Optical Communication Link
By

Imam Uz Zaman

Master of Science in Electrical Engineering

University of California, Irvine, 2018

Professor Ozdal Boyraz, Chair

Point to Point Optical Link (PPOL) and Inter Vehicular Visible Light Communications (IVVLC) as susceptible to many security threats including jamming, eavesdropping, interception and physical infrastructure attack. To address this problem, researchers have recently proposed to generate secret keys from the randomness of the physical environment such as wireless communication channel fading characteristics, phase fluctuation inside optical fiber and polarization mode dispersion in a long-haul network. In this thesis, I presented a novel symmetric physical layer based secret key generation schemes for both PPOL and IVVLC. Polarization Mode Dispersion (PMD) of long optical fiber links and intensity variations due to road surface roughness along with driving behavior are exploited as a random channel characteristic to generate cryptographic keys for PPOL and IVVLC, respectively. It has been demonstrated, modulation of a probe signal caused by PMD in a high speed PPOL (40Gb/s and 60Gb/s) is reciprocal with average Pearson correlation coefficient of 0.862, despite the presence of optical nonlinearities, dispersion, and noise in the system. 128-bit symmetric cryptographic key has been successfully generated using the proposed scheme. Moreover, PMD based encryption keys passed the National

Institute of Standards and Technology (NIST) tests. It is proved through simulations of a 50km link that, with optimal key generation settings, symmetric keys can be generated with high randomness (high P-values for NIST randomness tests) and with sufficient generation rates (>50%). Furthermore, this key generation algorithm successfully generated, with extremely low error, high entropy, secret keys with lengths up to 128-bits using a 1kbps probe signal with the proposed scheme. It is also found that the vehicular visible light channels have high entropy (e.g., 12-14 bits for 5000 samples) and that separate channels are highly uncorrelated to one another (e.g., Pearson correlation coefficient of 0.32). These results prove that physical communication channel can be exploited to generate symmetric cryptographic keys.

ORGANIZATION

In this thesis, a physical layer cryptographic key generation scheme for Point to Point Fiber Optic and Inter-Vehicular Optical Communication Link is presented. The thesis is organized in two chapters. Chapter 1 describes the security key generation scheme in a long-haul Point to Point Optical Link (PPOL). The system models, attack models, as well as key generation algorithms for a PPOL, are explained in Chapter 1, section 1.4. In addition, the simulation results and security key strength analysis for point to point optical link are quantified in Chapter 1, section 1.5. The symmetric security key generation scheme for Inter Vehicular Visible Light Communication (IVVLC) based on physical randomness is delineated in Chapter 2. The system model of IVVLC along with attack model, simulation results are explained in this chapter. Chapter 3 includes the summary and potential future research work on both PPOL and IVVLC security key generation.

Chapter 1

Physical Layer Security Key Generation for Point to Point Optical Link

1.1 Introduction

Over the past three decades, the accessibility and bandwidth demand of optical network has been increased tremendously. The Point-to-Point Optical Link (PPOL) has been employed in various applications ranging from Ethernet systems to telecommunications backbone infrastructure as well as military communication system. Optical Link incorporates optical fiber as a high speed transmission channel. However, like any other communication channel, Optical fiber is vulnerable to many security threats, involving jamming, eavesdropping, interceptions, and infrastructure attacks. In optical networks an adversary can eavesdrop on an optical system in a various way including physically tapping into the optical fiber [3], listening to residual crosstalk from an adjacent channel [4]. As the data rate of today's communication networks goes beyond 40Gb/s, the implementation of the real-time, low latency authentication and security of the data transmitted over optical fiber has become one of the most important areas of research. State-of-the-art data security (including fiber optic communication) is implemented by encrypting the data at the transmitter side and decrypting at the receiver side as shown in Fig 1.

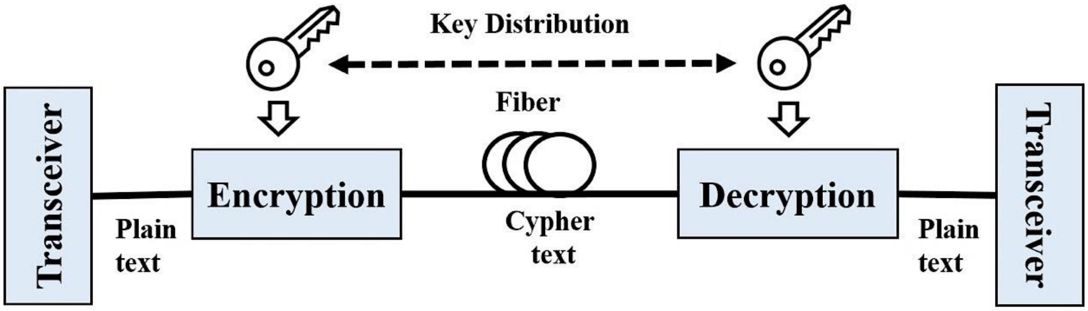


Figure 1. Schematic for data encryption and decryption system.

In general, cryptography requires one or more unique number known as keys. The cryptographic algorithm can be classified into two major categories [5]: Symmetric and Asymmetric. Symmetric algorithms (like Advanced Encryption Standard, RC4, Data Encryption Standard, etc.) use identical cryptographic keys for both encrypting the plaintext as well as decrypting cipher text [6]. Since symmetric algorithms do not require complex bit manipulations, they have low overhead and high performance [7]. However, the encryption key needs to be shared between two or more parties participating in the communication. In particular, the secret key is to be transmitted to the receiver side before the information is to be transmitted. It is impractical to ensure that no one will be able to tap communication channels during key exchange unless the channel is secured via cryptography and authentication. Hence, the only secure method of key exchange would be to personally transport the keys directly to the transmitters. This is a major drawback of symmetric key encryption. On the other hand, asymmetric algorithms (like RSA, Secure Sockets Layer, Digital Signature Algorithm etc.) do not involve a shared key for encryption and decryption, but public and private keys instead. However, higher computational power requirements, slow key generation process, more memory space requirements make asymmetric algorithms less suitable for time-critical and resource-limited applications. Recently, researchers have recently proposed to generate secret keys from the randomness of the physical environment as described in section *1.2*.

This thesis presents, a novel symmetric secret key generation scheme for Point-to-Point Optical Link (PPOL) communication by exploiting the underlying physical layer properties of the optical fiber under tight performance and memory constraints. In particular, it is shown that by exploiting the Polarization Mode Dispersion (PMD) phenomenon to generate symmetric random distortion for a bidirectional fiber transmission line.

The speed of the optical pulse inside an optical fiber is determined by the fiber refractive index. In the case of perfectly symmetrical (circular) optical fiber, the refractive indices are identical for both X and Y polarized light. However, fiber core asymmetry, internal-external stress, bends and twists of the fiber creates a difference in the refractive indices for X and Y polarized light. This difference is called birefringence. As a result, X and Y polarized lights propagates at a different speed. The delay between two polarization state is known as Differential Group Delay (DGD). Moreover, the average value of DGD is defined as PMD.

Evolution of PMD in the standard Single Mode optical Fiber (SMF-28) is totally stochastic, therefore the security strength of generated keys based on PMD is appreciably high. Secret key generation scheme by exploiting the underlying physical layer properties of the optical fiber is feasible under tight performance and memory constraints. In particular, this thesis proposes to exploit the Polarization Mode Dispersion (PMD) phenomenon to generate symmetric random distortion for a bidirectional fiber transmission line. The preliminary analysis presented in [1] shows that the encryption keys based on PMD have high reciprocity and high entropy.

The detailed system model to exploit the contingent nature of Polarization Mode Dispersion (PMD) using an optical switch and Randomly Spliced Polarization Maintaining Fibers (RSPMF) is described in this thesis. This thesis includes the characterization of the required length of RSPMF for different data rates and explained some of their effects on the overall key generation scheme. In addition to that, this thesis explains the security key generation scheme based on the proposed system model. The evaluated the effectiveness of the key generation scheme for 60Gb/s and 40Gb/s via mismatch rate analysis and state-of-the-art randomness tests created by the National Institute of Standards and Technology (NIST) are also delineated in this work. From the analysis, the optimal settings for our key generation scheme that maximizes the key randomness (98% test pass rate) and has a moderate key bit generation rate (60% on average) are found. Key randomness tests prove the randomness of the PMD modulated signal. Further, based on key generation parameters, it is observed that the generated keys are 40% different on average compared keys generated by a malicious non-invasive adversary.

1.2 Background

To address the limitation of the symmetric and asymmetric algorithm, various research groups and organizations have proposed to establish a hybrid solution for cryptographic algorithms [8]. Nevertheless, there are still some significant limitations to the current hybrid approach. It needs a key exchange session which utilizes an asymmetric algorithm whose lengthy computation time is generally not acceptable for safety-related applications which require a very short reaction time of 50 to 200 milliseconds [9]. The hybrid solution also requires an implementation of the asymmetry algorithm in the embedded devices, thus causing non-negligible memory space overhead. Moreover, the hybrid approaches cannot provide enough entropy due to high levels of predictability of the seed or user-given inputs and the deterministic nature of the key generation algorithm[10].

The most secure and solid alternative key generation scheme today is quantum optical-fiber cryptography, which provides impregnable security as assessed in [11], [12]. However, this quantum cryptography is an expensive and sophisticated solution suitable only for critical applications. Further, the secure quantum key distribution itself is a challenging task. Hence, a simple and efficient yet safe method for cryptographic key distribution is necessary. To address this problem, researchers have recently proposed to generate secret keys from the randomness of the physical environment [13]–[18].

The concept of the physical layer (PHY) based secret key generation is to exploit the randomly varying properties of the underlying physical layer. In the fiber-optic communication channel, the deployed optical fibers are considered as part of the physical layer (PHY). In [11] the author presented the use of quantum seals to test the integrity of the authenticity of a communication channel. The authors explained how a quantum physical layer senses tempering and how it communicates with the higher protocol layers to allow quantum seals to influence the security of data communication. Phase fluctuation in the optical fiber is exploited by using a large-scale Mach-Zehnder interferometer to generate and share keys in [14]. In [19], the authors showed that optical fiber communication encryption is possible based on four-wave mixing (FWM) in a very high non-linear bismuth-oxide fiber (Bi-NLF) and therefore requires specific fiber deployment. In [1], the authors showed that the stochastic nature of polarization mode dispersion in the optical fiber can be exploited to generate secret keys for cryptography.

1.3 Motivation

In most of the previous works, the methods related to physical-layer based fiber optic security requires a delicate and sophisticated system deployments which are complex and expensive. In addition to that, almost all of the previous works did not describe the key generation techniques established in their system modeling nor adequate analysis of secret key strength, key mismatch rates, and key entropy. In this work, I aim to implement a low cost, easily deployable symmetric cryptographic key generation technique based on uncompromisable physical randomness. Our goal is to solve the security challenges in resource-limited optical fiber links.

1.4 System Modelling

1.4.1 Randomly Spliced Polarization Maintaining Fiber Model

The proposed key generation scheme is based on the random variation of Polarization Mode Dispersion. PMD is related to Differential Group Delay (DGD) caused by birefringence in the optical fiber in a long haul network, $PMD = < \Delta\tau >$, where $< \Delta\tau >$ is the average value of DGD [20]. Birefringence varies along the fiber length and is totally nonstationary stochastic in nature. It arises from different internal and external stresses on the fiber including core asymmetry, non-uniform loading, bends, and twists. PMD is a random effect because it relies on the instantaneous weak birefringence state of the fiber link. Many experiments with the fiber of various lengths proved that PMD of a fiber link is proportional to the square root of the fiber length as in (1), where L is the length of the fiber.

$$PMD = PMD_{coefficient} \times \sqrt{L} \qquad (1)$$

Commercially available SMF-28 fibers have PMD coefficient of $0.04 ps/\sqrt{km} - 0.1 ps/\sqrt{km}$. Equation (1) reveals that, depending on the data rate, commercially available SMF-28 with PMD coefficient of $\approx 0.04 ps/\sqrt{km}$, the PMD effect manifests itself over an extremely large distance as shown in Table I, column 2[20]. As a consequence, the cryptographic keys generated from the modulated bit streams based on commercial SMF-28 possess low entropy. To mimic the effect of Polarization Mode Dispersion(PMD) of a long-haul fiber network in a smaller(≤50km) dispersion compensated Point to Point Optical Link (PPOL), two sections of randomly oriented Randomly Spliced Polarization Maintaining Fibers (RSPMF) are incorporated at both transceivers ends as in Fig. 4. Fig.2(a) shows the transmitted pre-defined bit sequence and Fig 2(b) shows the received signal at a 50km distance with RSPMF (blue) and without RSPMF (red).

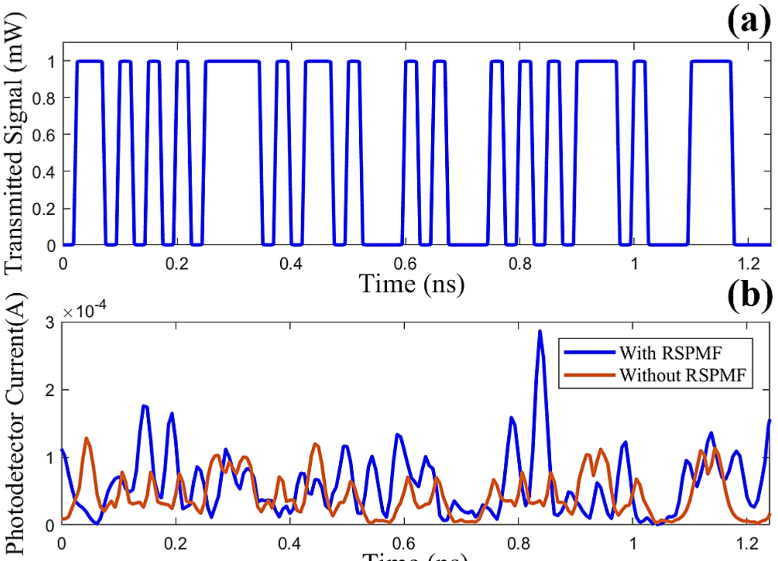


Figure 2. Transmitted random optical Signal at 40Gbps, (b) received signal with and without RSPMF of a 50km PPOL.

It is evident from Fig.2, the received signal in a 50km link without RSPMF has analogous amplitude variation as the pre-defined bit sequence. Therefore, the received signal cannot be exploited to generate cryptographic keys assuming that the attacker has the knowledge of the predefined bit sequence. The RSPMF is designed in such a way that the average value of DGD

(PMD) exceeds the maximum allowable PMD of the link, that is $\Delta\tau \geq 0.1T_B$ [20] where T_B is the bit period.

To demonstrate the concept $\Delta\tau \approx 0.25T_B$ is selected to achieve random amplitude modulation of the bit pattern. In the simulation, each PMF has beat length=1mm (from manufacturers specs) which implies $\Delta n=0.0016$. Given that DGD, $\Delta\tau \approx (\Delta n / c) L$, the total RSPMF length to achieve the desired PMD effect for 60Gb/s, 40Gb/s, 20 Gb/s and 10 Gb/s are summarized in Table I. For example, in a 60Gb/s system, $\Delta\tau$ exceeds 4.16ps ($\geq 0.25T_B$) when the RSPMF length is ≥8m. It can be seen from Fig 2 (b) that RSPMF enhances PMD, therefore, causes stochastic amplitude modulation of the bit pattern due to pulse splitting and random walk-offs between two orthogonal polarization states. In addition to the total length of the RSPMF, the number of spliced segments plays a crucial role in PMD modulation. Theoretically, one segment splits the input pulse into two. As a result, n number of segments can split the pulse $\leq 2^n$. Therefore, the higher the number of RSPMF segments the higher the random pulse modulation due to PMD. Figure 3 shows the random amplitude modulation of the input pulse pattern cause by RSPMF consists of 1,3,6 segments respectively in a 60Gb/s communication link.

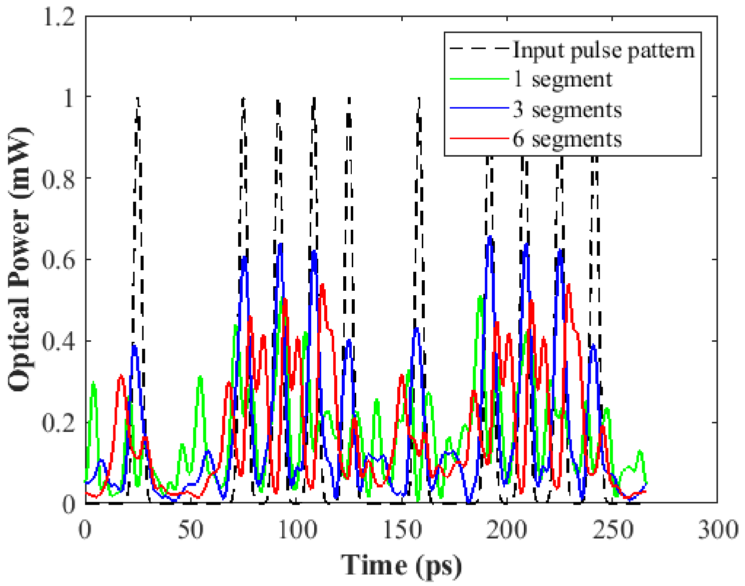


Figure 3. Effect of RSPMF segment number on the amplitude modulation of the optical pulse pattern due to PMD. The total length of RSPMF is 42m

Table I:

Summary of maximum allowable transmission distances for fibers with PMD coefficient of $0.04\,ps/\sqrt{km}$ and required RSPMF lengths

Data rate (Gb/s)	PMD limited link length for SMF-28 (km)	Required RSPMF length (m)
10	6.25 x 10^4	> 45
20	1.5 x 10^4	> 22
40	0.4 x 10^4	> 10
60	0.16 x 10^4	>8

The correlation among the modulated signals by RSPMF of different segments is given in Table II. It is evident that the modulation of the input pulse by RSPMF of different segment number (1,3,6) are highly uncorrelated from each other. In this thesis, to demonstrate the concept 6 segments of PMF with a randomly generated length between 8m and 16m are chosen to design each RSPMF of 42m. In total, the model incorporated two RSPMF (one at each transceiver side), in the simulation model.

Table II:
Correlation between the modulated bit pattern due to RSPMF

Segment Number	**1**	**3**	**6**
1	N/A		
3	0.431808	N/A	
6	0.225244	0.068404	N/A

1.4.2 Total Optical Point to Point Link Model

To prove the concept and to assess the feasibility of the model a point-to-point optical link simulation model is developed as shown in Figure 4. In this model, Alice and Bob are the two legitimate parties who want to communicate over a secured communication channel with symmetric cryptography. As mentioned earlier, symmetric encryption algorithms are faster and require less processing overhead compared to asymmetric algorithms. Using our presented method, Alice and Bob will be able to generate and exchange strong symmetric cryptographic keys to encode their plaintext without facing the key transportation challenges[20].

The proposed system works in two modes i.e. communication mode and key generation mode. The state of the link can switch between these two modes very fast with current technologies. For instance, by using commercially available fast optomechanical MEMS switches, the switching time between these two modes can be less than 0.5ms. In the key generation mode, the communication link between Alice and Bob is A-B-C-D-E in Figure 4. The Differential Group Delay (DGD) between two Principal States of Polarization (PSP) will be higher and more stochastic due to the high PMD effect from the long SMF-28 fibers in between the transceivers and the RSPMF and SMF pigtails in this path. After the key establishment agreement between Alice and Bob, the system will go into communication mode and send signals via the P-B-C-D-Q path just as in a conventional point-to-point optical fiber link. In the key generation mode, the changes in the Differential Group Delay follow a Maxwell probability distribution as given in (2), where l, $\Delta\tau$, q^2 are, fiber length, the average DGD and the variance of the Maxwell distribution, respectively [21].

$$P(\Delta\tau, l) = \frac{2\;\Delta\tau^2}{\sqrt{2\Pi}q^3} exp[-\frac{\Delta\tau^2}{2q^2}] \qquad (2)$$

Due to the orthogonality of the input principle state of the polarization (PSP), any input polarization can be written as a vector sum of its components. Equation (3) states the output electric field vector in the time domain. In (5), r_+ and r_- are the complex projections, ε_{out+} and ε_{out-} are unit vectors of the output PSP and $\phi\pm$ are the constant phases picked by the polarization modes, $\Delta\tau = |\tau_+ - \tau_-|$ represents the DGD.

$$\overrightarrow{E_{out}(t)} = r_+ \overrightarrow{\varepsilon_{out+}} e^{j\phi+} E_{in}(t+\tau_+) + r_- \overrightarrow{\varepsilon_{out-}} e^{j\phi-} E_{in}(t+\tau_-) \tag{3}$$

In this research, the renowned discrete waveplate model is adopted and used Jones matrix calculations [19], [20] to simulate the model in MATLAB and VPI Transmission maker. In the simulation model, a long single-mode fiber is simulated as the concatenation of a large number of birefringent waveplates each having the same indices but different lengths and orientations. If the polarization dependent loss is not included, the frequency dependence of the Jones matrix, and temperature fluctuation, any waveplate can be represented by (4). Where $S(\theta)$ denotes the rotation of the fast axis of the wave plate by θ degree from the +x axis, L is the fiber length, n_{fast} and n_{slow} are the refractive indices for fast and slow modes, respectively.

$$M(\omega) = S(-\theta) e^{\frac{-j\omega L(n_{fast}+n_{slow})}{2c}} \times \begin{bmatrix} e^{\frac{-j\omega L(n_{fast}-n_{slow})}{2c}} & 0;0 & e^{\frac{j\omega L(n_{fast}-n_{slow})}{2c}} \end{bmatrix} \times S(\theta) \tag{4}$$

The simulation shows that there is high reciprocity of the modulated probe signal in Alice's and Bob's channel. Across ten probe signals (1024 samples), the average Pearson correlation coefficient between Bob's received samples and Alice's received samples was 0.862 when the data rate was 60Gb/s and 0.868 when the data rate was 40Gb/s (1.0 is the maximum).

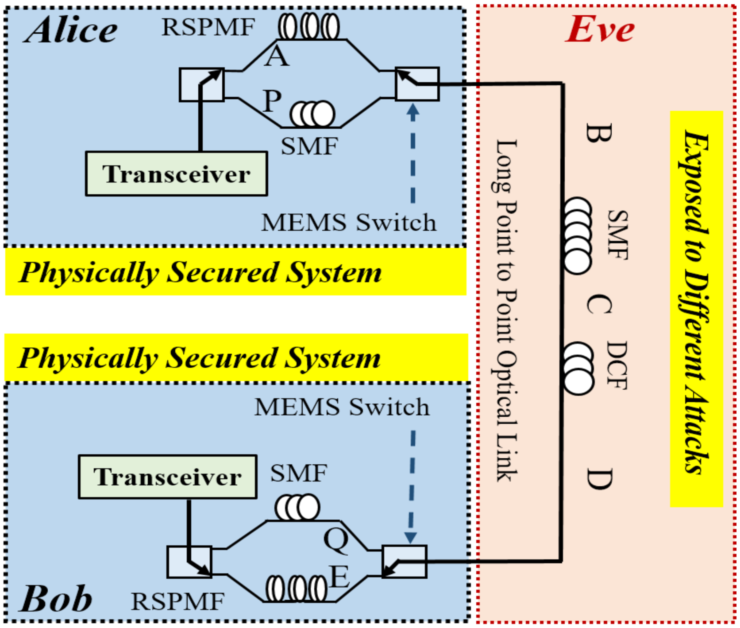

Alice
RSPMF
A
P
SMF
Transceiver
MEMS Switch
Physically Secured System
Eve
B
SMF
C
DCF
D
Long Point to Point Optical Link
Exposed to Different Attacks
Physically Secured System
MEMS Switch
Transceiver
SMF
Q
E
Bob
RSPMF

Figure 4. Proposed PMD based key generation scheme incorporating Randomly Spliced Polarization Maintaining Fibers (RSPMF). Alice and Bob are the two legitimate communication parties. The adversary (EVE) has access to the fiber network constitutes of SMF and Dispersion Compensated Fiber (DCF).

1.4.3 Attack Model

To define an attack model, this thesis assumes two communicating parties, Alice and Bob and a non-invasive eavesdropping adversary, Eve. The adversary, Eve, does not have direct physical access to the transceiver systems but may have access to a point along the communication link as shown in Fig. 4. As a result, Eve will not be able to observe the randomly spliced PMF, the probe signal, or input polarization state of the signal in the fiber that Bob observes from Alice and vice-versa. Moreover, in a later section, it will be showed that even if one assumes that Eve has information about these systems or features, Eve will not be able to generate the same keys as

Alice and Bob due to the stochastic nature of the PMD, which distributes over the entire length of the fiber.

1.4.4 Physical Layer Key Generation Scheme

The key generation scheme exploits the physical randomness from the PMD effect of the optical fiber link to generate symmetric secret keys for Alice and Bob. Alice and Bob initiate the key generation mechanism by sending predefined probe signal (bit sequence) to each other. These pre-defined signal pulses experience random polarization rotation, pulse splitting caused by pulse walk-offs between two orthogonal polarization states. The proposed scheme not only rely on $\Delta\tau$ but also rely on the polarization rotations of the pulses. That leads to stochastic amplitude modulations due to the stochastic nature of PMD. The randomness is observed as the photodetector current and then sampled, processed and quantized independently by Alice and Bob. Since they experience the same physical channel, the amplitude modulation measured at the two ends of the fiber would be highly correlated. Moreover, two parties perform a mismatch removal step to generate symmetric cryptographic keys to reduce the mismatch rate. In a practical application, one can anticipate that the key generation algorithm will run a few times (≈10) in a day. It is evident that the temperature, DGD of a deployed fiber varies significantly over a day due to hot spots [22]. A small change of polarization state and DGD at any hot spot of the link influences the overall experienced PMD effect and therefore, enable us to achieve random patterns more frequently than required.

1.4.5 Thresholds and Quantization

The first step in the key generation scheme involves Alice and Bob independently splitting their sets of received samples X_{Alice} and X_{Bob} into subsets/groups notated as $x_{Alice,i} \subseteq X_{Alice}$ and $x_{Bob,i} \subseteq X_{Bob}$ each of size G_{size} (the size of the last group may be less than or equal to G_{size}). Then, Alice and Bob take the average $\mu(x_{id,\ i})$ and the standard deviation $\sigma(x_{id,i})$ of each group and from them compute an upper threshold $Thr_{upper}(x_{id,i})$ and a lower threshold $Thr_{lower}(x_{id,i})$, where id ∈ {Alice, Bob}. The thresholds are defined as follows:

$$Thr_{upper}\left(x_{id,i}\right) = \mu\left(x_{id,i}\right) + \alpha * \sigma\left(x_{id,i}\right)$$

$$Thr_{lower}\left(x_{id,i}\right) = \mu\left(x_{id,i}\right) - \alpha * \sigma\left(x_{id,i}\right)$$

Where α is a programmable parameter that is assumed as constant. Nonetheless, it could be a unique value per threshold definition. This thresholding approach is the same as the one proposed by [23]. Per each subset $x_{id,(i,\ j)}$, Alice and Bob will generate thresholds and iterate through each sample $x_{id(i,\ j)}$. They will either quantize the sample into a binary value for their corresponding secret key and store its index for a later step or discard the sample and index. Alice and Bob will have a set with the quantized and stored key bits denoted as Kid and the stored indexes denoted as J_{id}. The pseudocode for the simple key bit quantization of a sample is defined as follows:

```
$if\ x_{id,(i,j)}\ \geq Thr_{upper}\left(x_{id,i}\right)\ then\ K_{id} + \{1\};\ J_{id} + \{j\};$
$else\ if\ x_{id,(i,j)}\ \leq Thr_{lower}\left(x_{id,i}\right)\ then\ K_{id} + \{0\};\ J_{id} + \{j\};$
$else\ do\ nothing$
```

1.4.6 Index Exchange and Index Mismatch Removal

After each group of samples is quantized into secret key bits, Alice and Bob will perform a secure mismatch removal step without revealing the secret key bits. This requires Alice and Bob to store the indices of the samples that were successfully quantized in their own sets, J_{alice} and J_{bob} respectively. Because of system noise and the choices for α and G_{size}, these sets may not match one another. Thus, Alice and Bob must exchange these sets (note that the actual values to avoid revealing key bits in this manner) to compare their own sets with the received sets and remove indices that do not exist in the received sets. Only the quantized secret key bits corresponding to the indices that have not been removed can be used in the final keys. The following is the pseudocode for the index removal and exchange steps from the perspective of one of the parties:

```
$Define\ id1, id2\ \in \{alice, bob\}\ s.t.\ id1\ \neq id2$
$for\ each\ j \in J_{id1}:$
$if\ j \notin J_{id2}\ then\ J_{id1} = J_{id1} - \{j\};$
$else\ do\ nothing$
```

1.4.7 Key Bit Generation Rate

Per each group of samples $x_{id,i}$, there are M_i secret key bits generated after the quantization and mismatch removal steps, where $M_i \leq G_{size}$. Thus, the key generation scheme will continue to generate secret key bits until it runs out of samples to quantize (N). The final number of generated secret key bits is notated as M, where $M \leq N$. Then, according to the selected cryptographic scheme's key length requirement L_K, the M secret key bits can be divided into M/L_K secret keys which will be stored and used for the current and future sessions. The high data and sampling rates (limited only by hardware) due to the optical fiber medium are highly suitable for this key generation technique because it makes up for the quantization method's tendency to have low key bit generation rates. In radio channel communication, the data rates are strictly limited and constrain the effectiveness of the technique.

1.4.8 Generation Key Bit Mismatch Rate

After the entire key generation process, all the removed samples from the quantization step in Section C.1 as R_Q and the removed samples/bits from the index mismatch removal step as $R_{I.}$ The ratio of R_Q to R_I is known as the *Quantization Mismatch Rate* (*QMR*). Preferably, it is desirable to have a low *QMR* because a lower *QMR* implies a higher key bit generation rate. Unfortunately, the noise and choices for G_{size} and α can greatly affect the *QMR*. This research work would also like to evaluate the number of mismatching bits between Alice's generated key and Bob's generated key. These mismatching bits may be caused by extreme elements of noise that disturbed the reciprocity of the signals. Let's consider this as the *final mismatch rate* (*FMR*) which is equivalent to the Hamming distance (number of bits needed to be flipped for one key to be the same as another key) between Alice's and Bob's keys. Figure 5 provides a visual overview of the key generation scheme.

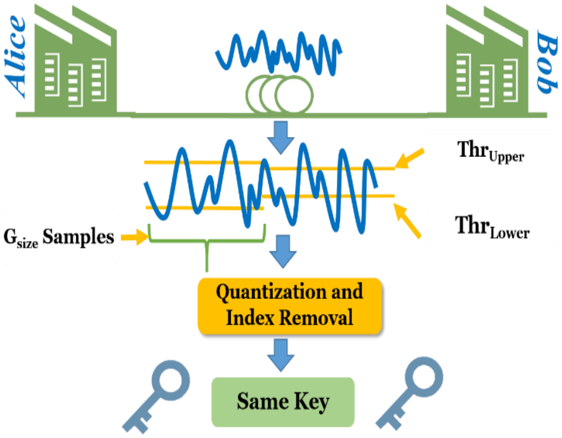


Figure 5. Key generation overview.

1.5 Simulation Results and System Verification

1.5.1 Simulation Tools

Signal propagation through the system is modeled by a combination of commercial simulation tools including VPI transmission Maker and Matlab. As a simulation tool, this work used the VPI transmission Maker to generate the probe signal's bit sequence, to create the modulated NRZ optical signal, to detect the signal at the detector and to include Relative Intensity Noise (RIN), shot noise, thermal noise, etc. The propagation of the optical signal in both directions based on the split-step method is modeled in MATLAB. The reciprocity check and the key generation algorithm are implemented in MATLAB as well. In our simulation, input laser power= 1mW, link length=50km, wavelength=1550nm, PMD coefficient of SMF= $0.04\, ps/\sqrt{km}$, PMF's beat length=1mm, SMF dispersion= $18\ ps/nm{\bullet}km$, DCF dispersion= $-100\ ps/nm{\bullet}km$, Non Linear Index= $2.6\times10^{-20}m^2\ /W$ RSPMF segments=6, total length of RSPMF=42m, length of each randomly oriented segment= 8m-15m. The simulations are performed according to different bit rates (40Gb/s, 60Gb/s) to check the reciprocity and entropy of the modulated signal due to high PMD. In this section, some results on our simulations using the system model in Section II are provided. The rate of change of average DGD is considerably slower than the data transmission rate (40Gb/s, 60Gb/s). Hence, the channel response is considered as constant during predefined probe signal propagation. An example of the received signal thresholding for Alice and Bob at 60 Gb/s is shown in Figure 6. A critical aspect of this section is to find settings that result in keys with strong randomness and moderately low mismatch rates.

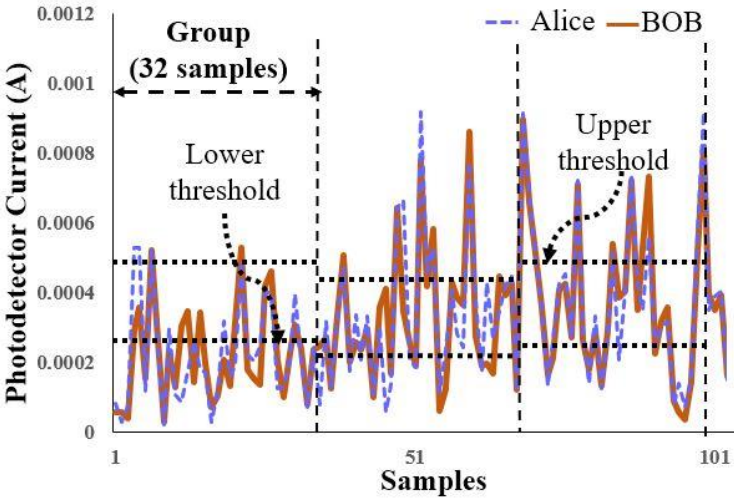


Figure 6. Sample thresholding for 60Gb/s PMD modulated data.

1.5.2 Quantization and Final Mismatch Rate (FMR) Evaluation

In the simulation, 10 probe signals (each one with the same 1028-bit pattern) are exchanged between Alice and Bob and ran the key generation algorithm with different settings. Figure 7(a) and Figure 7(b) show the quantization mismatch rates, *QMR*, according to the settings that are tested the key generation algorithm with for the 40Gb/s and 60Gb/s fiber links, respectively. As can be seen from Figure 7, α and G_{size} deeply affect the mismatch rate. As α decreases, the mismatches greatly increase and as G_{size} increases, the mismatches tend to increase as well. However, around $G_{size} \in [28, 32]$ the mismatch rates begin to plateau and decrease. Interestingly, one may observe that the quantization mismatch rates for $\alpha = 0.2$ and $\alpha = 0.3$ for the 40Gb/s rate simulation were quite similar (unlike in the 60 Gb/s rate simulation).

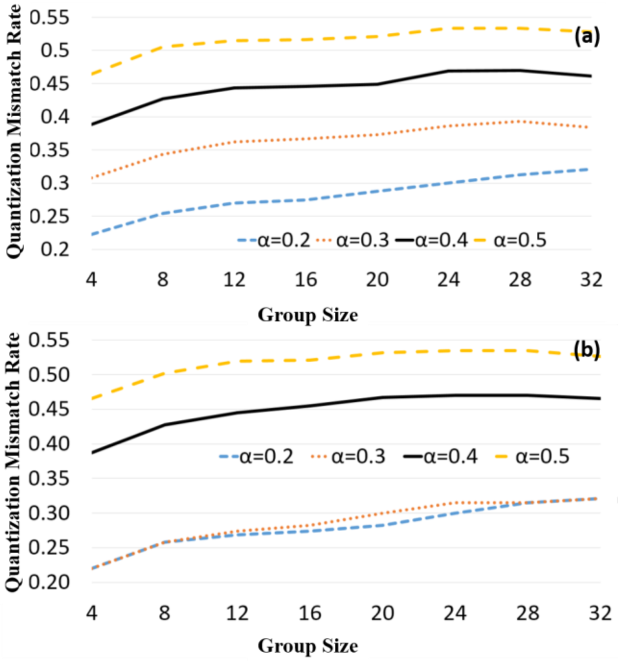


Figure 7. Alice-Bob quantization mismatch rates (QMR) from quantization and index exchange and removal steps for 40Gb/s (a) and 60Gb/s (b) rates when sampling offset = 1, LK = 128 bits, α ∈ [.2, .5] with a step size of 0.1, and Gsize ∈ [4, 32] with a step size of 4.

As shown in Fig. 8, it is also found that for both the 40Gb/s and 60Gb/s rates, the *FMR* values were ≤10% (worst case). On average (across all settings), the *FMR* values were about 5%, except when G_{size} = 5. Despite having *FMR* values less than 10%, it is observed that there exists a final key mismatch rate (# of mismatching keys / # of total keys) greater than 50% and less than 85% for each setting (besides $\alpha = 0.5$). However, these values can be considered negligible considering how many 128-bit keys that can be generated per sent probe signal (approximately less than or equal to five keys per probe signal).

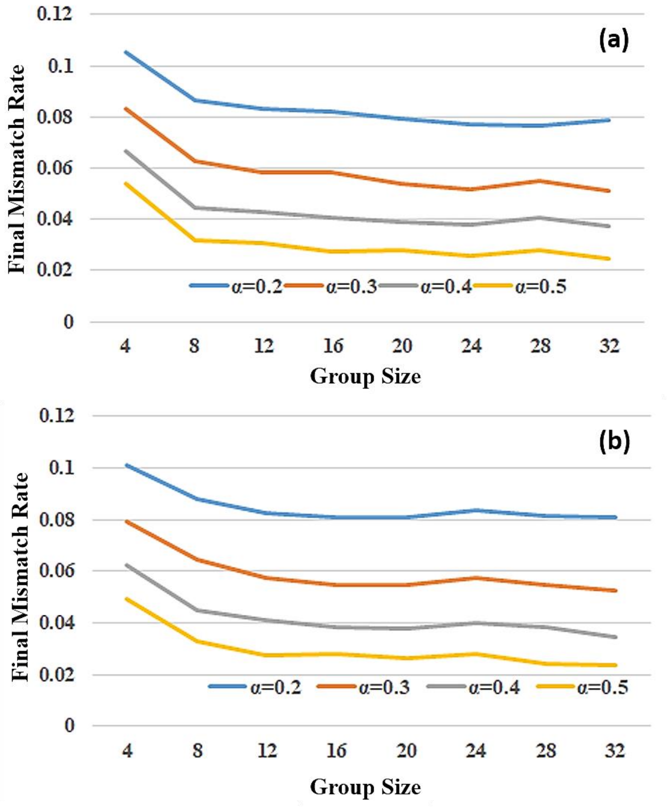


Figure 8. Final Mismatch Rate (FMR) between Alice and Bob for a 50km link (a) 40 Gb/s data rate (b) 60 Gb/s data rate.

1.5.3 Randomness Evaluation

1.5.3.1 NIST Randomness Test Setting

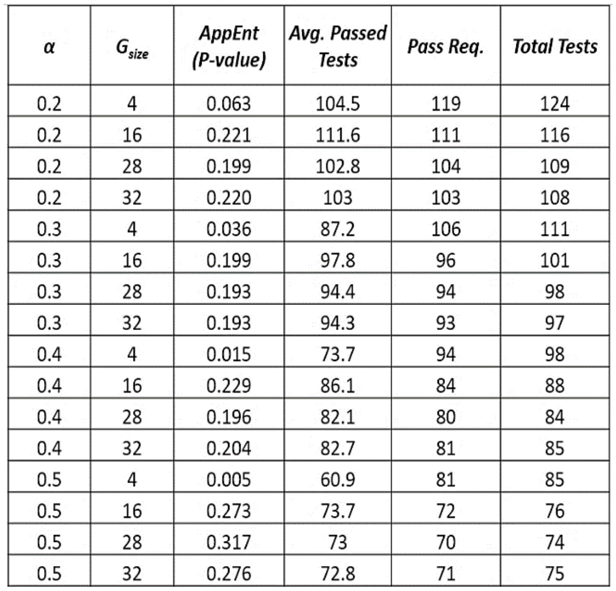
To evaluate the randomness (and equivalently, the security strength) of the generated secret keys, this research opted to use the randomness tests provided by NIST [24]. Each test is designed to evaluate a bit sequence's randomness via a specific pattern or an information theoretic metric. Since several NIST tests require certain lengths for their tests, it is not possible to use all the 16 tests because this key generation scheme is restricted by the key length L_K (some tests require 10^6 bits or more per bit sequence). Hence, this research work applies tests with more relaxed constraints. Therefore, the bit stream is set to size N as same as L_K and for certain tests, another value needs to be specified, block size M. The tests that are chosen to be used in this thesis are: Frequency; Block Frequency; Cumulative Sums Forward and Backward; Runs; Longest Run; Discrete Fourier Transform; Approximate Entropy (AppEnt); and Serial Forward and Backward.

The block size (M) was set to: 128, 1, and 3 for the Frequency, Approximate Entropy, and Serial tests, respectively, where the rest of the tests did not have an adjustable block size. For each pair of α and Gsize that was chosen in the previous tests, each NIST test is run over all of Alice's generated 128-bit keys generated from the 10 exchanged probe signals between Alice and Bob. To pass a test, a 128-bit key (bit stream) must not have a predictive pattern and must have a P-value greater than the default 0.01 significance level (this also means that only 1% of the tests can fail). To derive the P-value, a standard normal distribution and a *chi-square* reference distribution is used.

1.5.3.2 Evaluation of Security Strength of Alice-Bob Generated Symmetric Keys

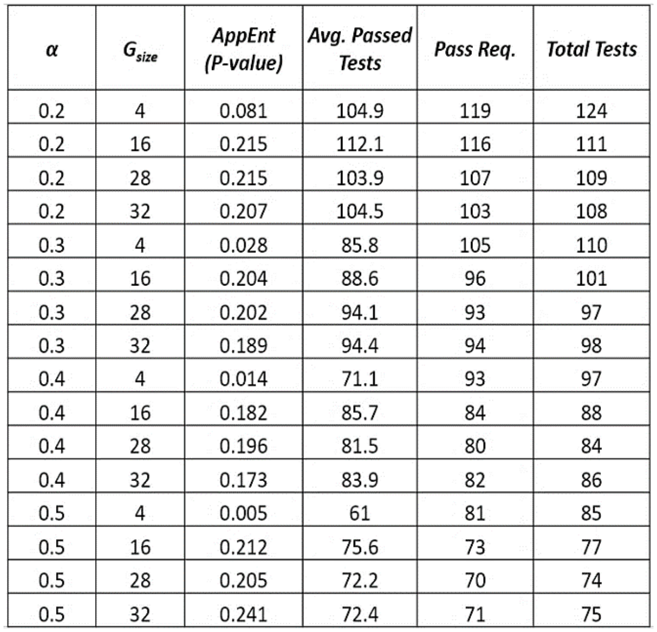
In Table III and Table IV, per each key generation setting Columns 1 and 2, provide: the averaged P-values for the Approximate Entropy test in Column 3; the average number of passed tests for each setting in Column 4; the minimum number of tests required to pass (per test type) in Column 5; and finally, the total number of tests ran are in Column 6. Table III corresponds to the tests when the data rate is 40Gb/s and Table IV corresponds to the tests when the data rate is 60Gb/s. As seen in the tables, the results vary according to the choices of the data rate, $G_{size,}$ and α. In general, from observations alone, it can be understood that the lower G_{size} and α are, the less the randomness. However, one benefit of having lower values for these settings is that there is a higher key bit generation rate. Thus, there is a possibility that there is a Pareto optimal solution for the key generation settings, the key randomness, and the key bit generation rate. Interestingly, the best solutions (in terms of randomness and decent key generation rates) occur when G_{size} = 28 or G_{size} = 32 for all values of α where the minimum number of tests were passed and there were moderately high entropy values (the average P-value was greater than 0.2 and individual P-values reached up to 0.7-0.9).

Table III:
The results of the NIST randomness tests of a 50km fiber link for generated keys when the data rate is 40Gb/s



α	*G_{size}*	*AppEnt (P-value)*	*Avg. Passed Tests*	*Pass Req.*	*Total Tests*
0.2	4	0.063	104.5	119	124
0.2	16	0.221	111.6	111	116
0.2	28	0.199	102.8	104	109
0.2	32	0.220	103	103	108
0.3	4	0.036	87.2	106	111
0.3	16	0.199	97.8	96	101
0.3	28	0.193	94.4	94	98
0.3	32	0.193	94.3	93	97
0.4	4	0.015	73.7	94	98
0.4	16	0.229	86.1	84	88
0.4	28	0.196	82.1	80	84
0.4	32	0.204	82.7	81	85
0.5	4	0.005	60.9	81	85
0.5	16	0.273	73.7	72	76
0.5	28	0.317	73	70	74
0.5	32	0.276	72.8	71	75

Table IV:
The results of the NIST randomness tests of a 50km fiber link for generated keys when the data rate is 60Gb/s.



α	*G_{size}*	*AppEnt (P-value)*	*Avg. Passed Tests*	*Pass Req.*	*Total Tests*
0.2	4	0.081	104.9	119	124
0.2	16	0.215	112.1	116	111
0.2	28	0.215	103.9	107	109
0.2	32	0.207	104.5	103	108
0.3	4	0.028	85.8	105	110
0.3	16	0.204	88.6	96	101
0.3	28	0.202	94.1	93	97
0.3	32	0.189	94.4	94	98
0.4	4	0.014	71.1	93	97
0.4	16	0.182	85.7	84	88
0.4	28	0.196	81.5	80	84
0.4	32	0.173	83.9	82	86
0.5	4	0.005	61	81	85
0.5	16	0.212	75.6	73	77
0.5	28	0.205	72.2	70	74
0.5	32	0.241	72.4	71	75

1.5.3.3 Evaluation of Security Strength in Presence of Noninvasive Attack by Eve

The security key generation scheme is based on the detected amplitude modulation of the probe signal that is caused by splitting, random walk-off and the mixing between two orthogonal polarization states which takes place statistically over the whole fiber length. It is seen from (5) that each pulse (E_{in}) of the probe signal splits into two in the presence of high birefringence segment. After passing n segments of RSPMF one single pulse can be splitted successively up to 2^n. The predefined signal consists of 1024 bits (pulses). A total number of pulses can be $\leq 41x10^5$. Random walk-offs and mixing of this large number of pulses over a long distance(≈50km) is totally stochastic. As a result, the probe signal experiences a random intensity modulation. This modulation due to pulse splitting and mixing is highly sensitive to polarization changes due to hotspots[22], segment number of RSPMF, orientation among the segments, the total length of the RSPMF as well as input polarization state . For example, Figure 9 shows the variation of the modulation due to three different input linear polarization states (0^o,10^o and 20^o), keeping all other parameters of RSPMF and SMF same. It can be seen from Table V that the modulated signals are highly uncorrelated from each other. It is quite impossible for an attacker to know all these parameters and therefore, Eve will not be able to emulate the same bit pattern modulation even after possessing knowledge of the Jones matrix.

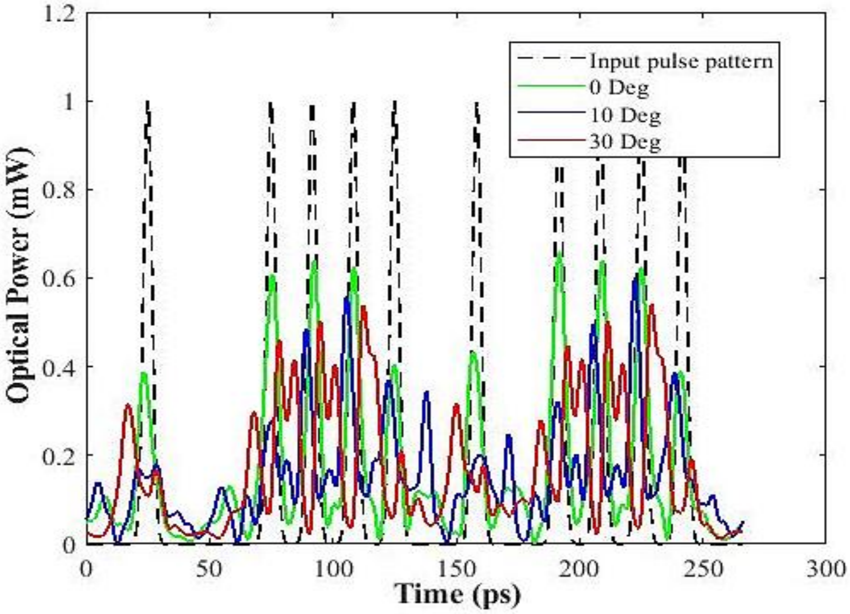


Figure 9. Random modulation of bit pattern due to different input polarization.

Table V:
Correlation among modulated signal for input polarization state.

Polarization angle	**0º**	**10º**	**20º**
0º	N/A		
10º	0.550008	N/A	
20º	0.068404	-0.05037	N/A

To further evaluate the security strength of the key generation scheme, the keys between Alice and Eve is generated, where Eve is imagined to be eavesdropping messages sent across the fiber link around 10 km from Alice. Eve is assumed to be a non-invasive attacker who is purely interested in deriving the same symmetric keys as Alice and Bob to decrypt and eavesdrop on the secured channel. A comparison study has been made on the newly generated keys between Alice and Eve with the same keys generated between Alice and Bob as in the previous section. The results are summarized in Figure 10, where the 128-bit keys generated between Alice-Bob and those between Alice-Eve are approximately 40-50% different from one another across all settings. This is a nontrivial mismatch rate range (keeping in mind that a single bit mismatch would cause encryption to fail) and indicates that there is almost no way that Eve can generate similar keys to Bob, unless they are virtually beside Bob (which is assumed to be impossible since Bob's transceiver is assumed to be physically protected).

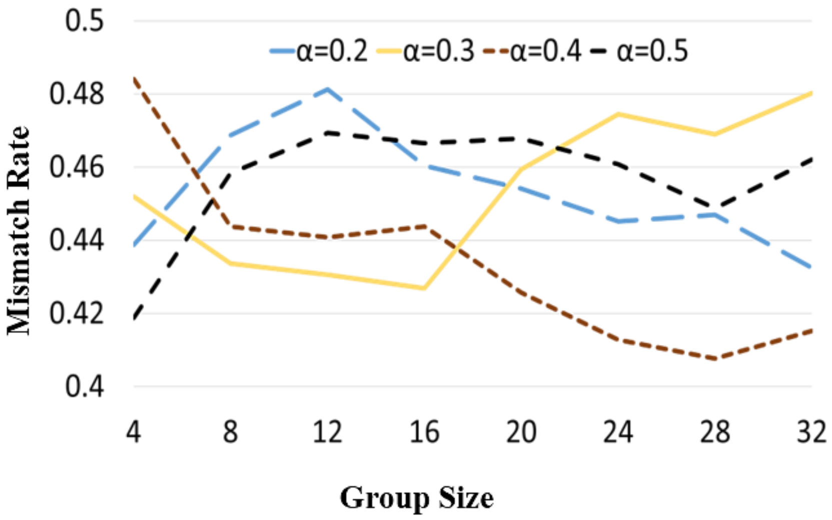


Figure 10. Averages of the final mismatch rates (FMR) between the keys that Alice and Bob generated (50km) and the keys that Eve generated (10km from Alice) at 60Gb/s.

1.6 Conclusion

In this chapter, a novel scheme for Point-to-Point Optical Link communication security is delineated to help resolve the high resource requirements and lack of a trustworthy source of high randomness of existing communication security solutions. The scheme includes a novel model and a physical layer symmetric cryptographic key generation technique. It focuses on exploiting the physical randomness manifested by the PMD effect. It is showed that this randomness makes it extremely hard for an adversary to generate the same cryptographic keys as the communicating parties. This key generation scheme successfully generated 128-bit keys with low final mismatch rates (≤10%) which could easily be truncated for 64-bit and 32-bit keys if necessary.

Chapter 2

Physical Layer Security Key Generation for Inter-Vehicular Visible Light Communication

2.1 Introduction

Visible light communication (VLC) is a rapidly growing wireless optical communication technology in which visible light is employed as the optical carrier exploiting the advantage of omnipresent commercial LEDs and photodiodes. Recently VLC has also been proposed as an effective alternative to radio-based (RF) wireless networks for short-range communication due to its many intrinsic advantages over RF such as high spectral availability, precise pointing due to high directional Line of Sight (LOS) propagation, and immunity to the multipath fading. All these qualities make VLC the best choice for Vehicle to Vehicle (V2V) communication, especially where vehicles need to be driven in a controlled close formation to increase traffic fluidity, road throughput and hence to decrease traffic jam. Data flow among vehicles is so vital in delivering vehicle information such as speed, brake, acceleration and any kind of warning for safety operation of vehicles. Like any other communication medium, VLC is susceptible to many security threats including jamming, eavesdropping, interception and physical infrastructure attack [25], and hence securing the communication with a reliable cryptographic design is desirable. However, the key management is the hardest problem in cryptography. The state-of-the-art cryptographic algorithm requires pre-shared keys, which is easily accessible to attackers if they have comprehensive knowledge of the system. In this chapter, a novel symmetric secret key generation scheme for vehicular VLC link is presented. In particular, the key generation scheme utilizes random modulation of a low data rate (1Kbps to 1Mbps) probe signal and therefore, the random intensity variation at the detector to generate symmetric keys. The random intensity variation is caused by the totally stochastic nature of road surface roughness and driving behavior of vehicle drivers. To increase the reliability a market-weighted headlamp beam model [26] and vehicle trajectory data by Next Generation SIMulation (NGSIM) program (by Federal Highway Administration) [27] are incorporated into our mathematical model. Extremely low error, high entropy, secret keys with lengths up to 128-bits using a 1kbps probe signal can be generated with the proposed scheme. It is also shown that vehicular visible light channels have high entropy (e.g., 12-14 bits for 5000 samples) and that separate channels are highly uncorrelated to one another (e.g., Pearson correlation coefficient of 0.32).

2.2 System Model

The proposed physical layer secret key generation method utilizes the randomness in the received signal due to road conditions and driving behavior. To prove the concept a model is developed that is based on the stochastic vehicle trajectory data provided by the NGSIM program and the road surface roughness, and the headlight modeling [28]. Since the Lambertian model is not an accurate model to simulate the intensity pattern of a vehicle's headlight and taillight, this thesis utilized a market-weighted headlamp beam model [29]. Using the luminous intensity (candela) table provided in this model one can calculate the corresponding illuminance value at any point of interest. In this thesis, the scope is limited to the Line of Sight (LOS) communication to generate symmetric secret keys. The illuminance (L) at the photodetector (PD) at the vertical angle (θ) and horizontal angle (ϕ) with respect to headlamp axis is determined by the following equation [30] where $r, dA, d\omega, \tau, I(\phi,\theta)$ are communication distance, photodetector (PD) area, solid angle, the angle between the photodetector normal and the incident direction, and luminous intensity respectively.

$$L = I(\phi,\theta) \times (d\omega / dA) = I(\phi,\theta) \times (\cos\tau / r^2)$$

Then the received Line of Sight (LOS) optical power (PRX-LOS) is calculated by $P_{RX-LOS} = (L \times A_r) / LER$ when $0 \le \tau \le \Omega$ otherwise $P_{RX-LOS} = 0$ [31] where A_r, **Ω**, and LER are the PD's total area, the half angle of PD's field of view (FOV) and the luminous efficacy of radiation, respectively. From the equation mentioned, the received optical power and hence photodetector current can be estimated effectively. Moreover, it is assumed that the taillight follows the same model as the headlight but with much lower intensity. The noise calculation algorithm includes shot noise due to background solar radiation and other artificial lights. Thermal noise associated with the receiver as mentioned in [31] are also included in the mathematical model. Relative velocity and hence relative lateral and longitudinal distances among vehicles result in random variation in intensity pattern. This randomness can readily be exploited to generate symmetric cryptographic keys.

To generate symmetric keys and to assess the feasibility of the key generation scheme a model of communication links between the vehicular transceivers *Alice(A)* and *Bob(B)* and another communication link between *Alice(A)* and the adversary *Eve (E)* (Figure 11) is developed. When Alice and Bob want to generate a symmetric key, they need to exchange a pre-defined probe signal (PRBS modulated bit pattern with a predefined length). To increase the reliability of the proposed method, the data of the vehicles such as speed, lateral coordinate (X), longitudinal coordinate (Y), time etc. are extracted from the vast amount of data provided by the NGSIM program [27]. Moreover, Using big data analysis the combination of three vehicle (Alice, Bob and Eve) are chosen in a way so that all three are omnipresent in the vicinity of each other in the real world with the intended point to point link establishment between Alice and Bob. This satisfies our communication link model (Figure 11). Then from the NGSIM data, the relative lateral (ΔX) and longitudinal (ΔY) distances between selected transceivers over a time duration are calculated. These relative distances are totally stochastic (Figure 12). This research work also includes stochastic road surface roughness (ΔH) in our mathematical model [28]. From all these information received intensity distribution at the photodetector is calculated. The NGSIM data has been interpolated to generate keys due to the lack of available data points. The simulation shows that there is a high reciprocity of the modulated probe signal (>.9 Pearson correlation coefficient between Alice and Bob’s signal) and there is also high randomness (approx.> 12 bits of Shanon entropy for each group of samples). It can be shown that the data received by both Alice and Bob is reciprocal and hence symmetric cryptographic keys can be generated successfully.

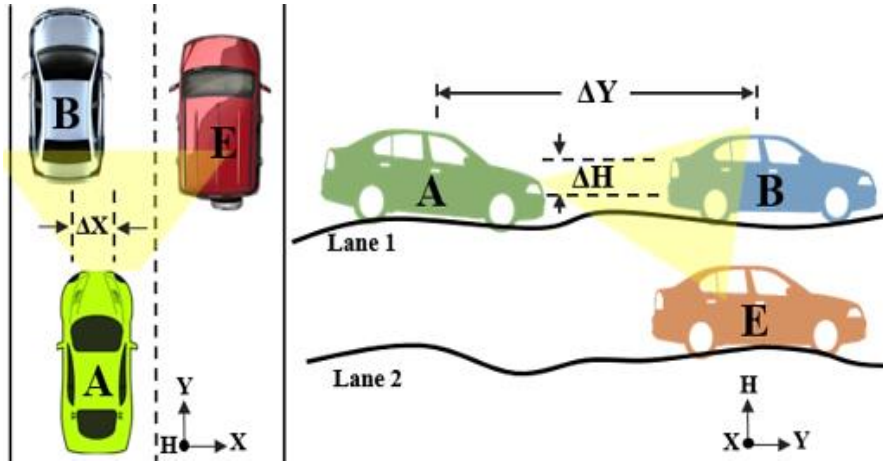


Figure 11. Vehicular Key Generation Model

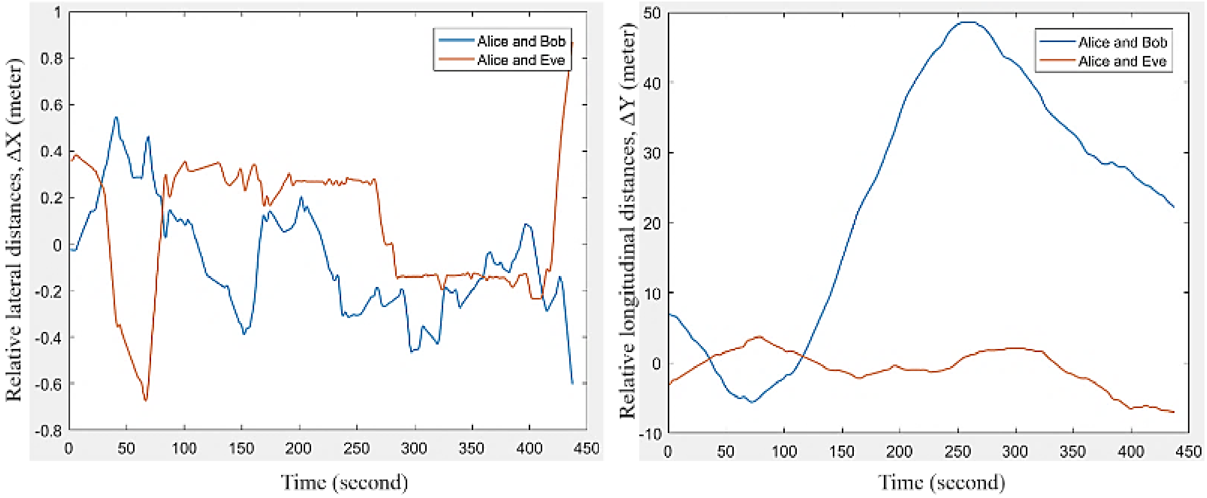


Figure 12. Relative lateral and longitudinal distances among Alice, Bob and Eve.

2.3 Attack Model

To address an attack model, this thesis assumes two communicating parties, Alice and Bob and a non-invasive eavesdropping adversary, Eve. Eve is present in the vicinity of the communication link between Alice and Bob as in Figure 11. Conventionally, the adversary Eve (E) does not have direct access to the transceiver systems but might have access to the communication link. Furthermore, in this thesis, it is shown that, even if Eve has all the information regarding the communication system or features, Eve will not be able to generate same keys as Alice and Bob due to the stochastic nature of vehicle trajectory and road surface roughness.

2.4 Pre-Shared Key Generation

In this key generation technique, each transceiver will take a set of samples from the pre-defined probe signal (PRBS) and quantize their signal strengths (in Amps) into symmetric key bits. Each transceiver will compute upper (Thrupper) and lower (Thrlower) thresholds of a group of samples, SampleGroup, with group size (g), based on their mean and variance by the following algorithm:

$$Thr_{upper} =< SampleGroup > +\alpha \times \sigma(SampleGroup)$$

$$Thr_{lower} =< SampleGroup > -\alpha \times \sigma(SampleGroup)$$

$$if\ \ SampleStrength\ \geq Thr_{upper} then\ Key\ Bit\ = 1$$

$$if\ \ \text{else if}\ SampleStrength < Thr_{lower} then\ Key\ Bit\ = 0$$

$$else\ do\ not\ quantize$$

Where <x> and σ(x) represent the mean and variance of x respectively. Both g and α parameters that can be derived or altered according to the variance of the system (the higher the variance, the higher α, and g should be).

The whole key generation scheme is demonstrated in the following figure:

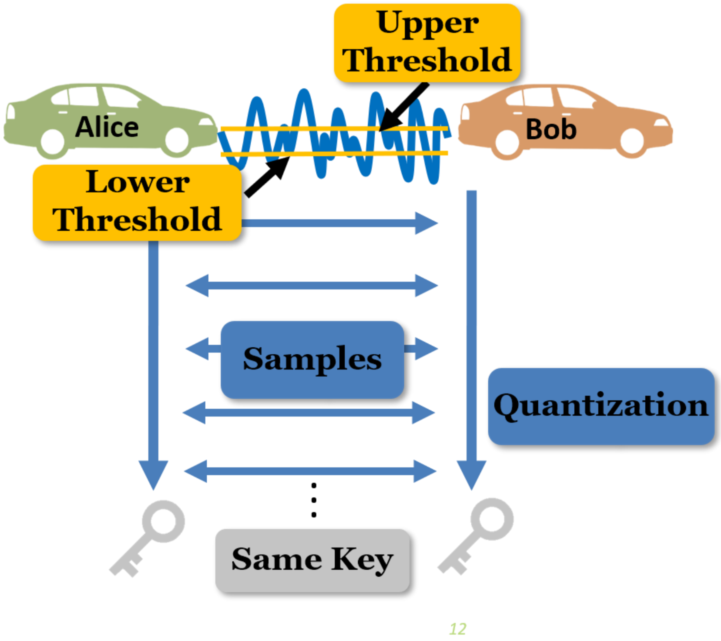


Figure 13. Key generation scheme for VVLC.

2.5 Simulation Results and System Verification

Signal propagation between the vehicles, the reciprocity check, entropy calculation, noise calculations, and the key generation algorithm are modeled in Matlab. The mathematical model utilized the US 101 (Hollywood Freeway) data from NGSIM to extract vehicles position related information and include it in our mathematical model. For key generation phase, in this work, the simulation is performed using 1Mbps probe signal. Figure 14 and Figure 15 show the sample randomly modulated electrical signal received by *Alice* and *Bob* and corresponding thresholding to generate keys respectively.

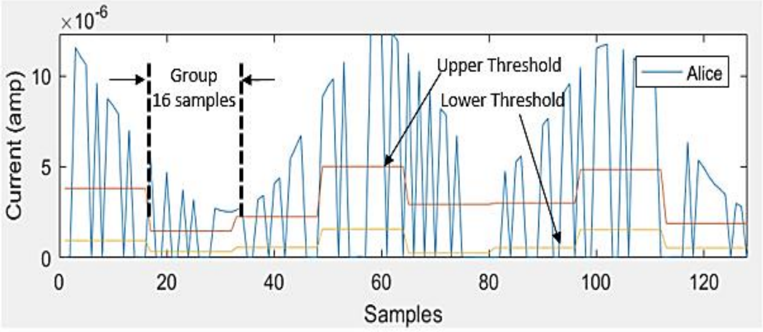


Figure 14. Sample signal and thresholding for Alice

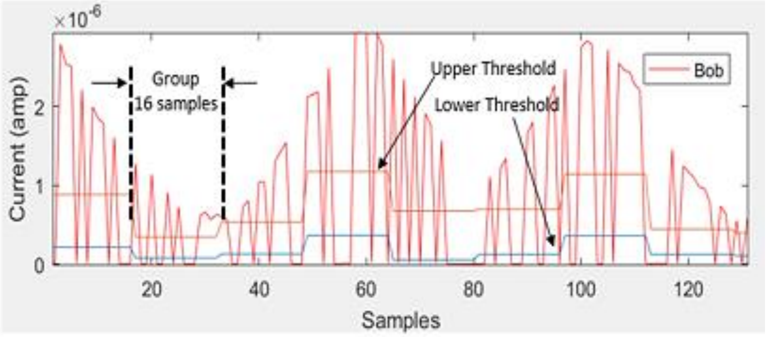
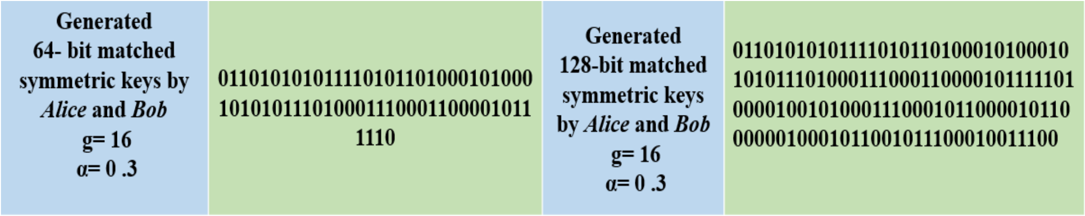


Figure 15. Sample signal and thresholding for Bob

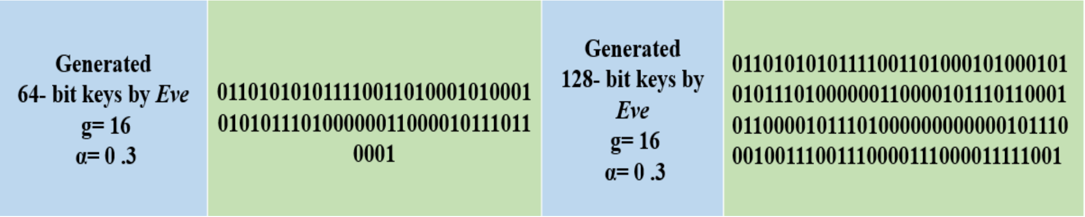
The generated keys of different lengths with parameters of group size, *g*, equal to 16 and α equal to 0.3 are shown in Figure 16 and Figure 17. The final bit mismatching rate is 0 (0 mismatching key bits/640 key bits) over 10 different 64-bit keys and 0 (0 mismatching key bits/1280 key bits) over 10 different 128-bit keys, demonstrating the applicability of this algorithm. To minimize the simulation time, 2^7-1 Pseudo Random Bit Sequence (PRBS) is used as the probe signal and 512 bits to emulate data propagation. When the key generation is done the vehicles can communicate just as conventional VLC link up to 1Gbps data rate.

For the *Alice* and *Eve* channel, using the same algorithm as used in *Alice-Bob* channel and with the similar PRBS (as it is assumed *Eve* has all the information of the link) the keys for the adversary *Eve* in Figure 17. A final key bit mismatching rate value of 0.0063 (4 mismatching key bits/640 key bits) for the *Alice-Eve* channel after creating 10 different 64-bit keys is calculated, and, interestingly, the same value of 0.0063 (8 mismatching key bits/ 1280 key bits) after creating 10 different 128-bit keys. As can be shown, it is clear that the resulting keys from the *Alice-Bob* channel are different than those generated from the *Alice-Eve* channel. Furthermore considering 5000 samples of received current values the entropy values (H, in bits) for each channel is computed, $H_{Alice\text{->}Eve}$ = 14.77, $H_{Eve\text{->}Alice}$ = 12.95, $H_{Alice\text{->}bob}$ = 13.36 and $H_{Bob\text{->}Alice}$ = 11.97, demonstrating the high channel variation. Moreover, the correlations for each channel is computed: $Corr_{Alice\text{-}Eve}$ = 0.78 and $Corr_{Alice\text{-}Bob}$ = 0.77 showing clear symmetricity, and $Corr_{Alice\text{-}Bob\text{ to Alice-Eve}}$ = 0.3163, demonstrating that the two main channels *Alice-Bob* and *Alice-Eve* are uncorrelated to each other. These results have been published in [17].



Generated 64- bit matched symmetric keys by *Alice* and *Bob* g= 16 α= 0 .3	**01101010101111010110100010100010101011101000111000110000101111110**	**Generated 128-bit matched symmetric keys by *Alice* and *Bob* g= 16 α= 0 .3**	**01101010101111010110100010100010 10101110100011100011000010111110 1 00001001010001110001011000010110 00000100010110010111000100111 00**

Figure 16. Sample key generation for Alice and Bob.



**Generated
64- bit keys by *Eve*
g= 16
α= 0 .3**
	01101010101111001101000101000101010111010000000110000101110110001	**Generated
		128- bit keys by
		Eve
		g= 16
		α= 0 .3**
			**0110101010111100110100010100010 1
			01011101000000011000010111011000 1
			01100001011101000000000000101110
			00100111001110000111000011111001**

Figure 17. Sample generated security keys by Eve

2.6 Conclusion

It is observed from the simulation result that driving behavior and road surface roughness can be exploited to generate high entropy symmetric cryptographic keys. 64 bit and 128-bit keys are generated to show the feasibility of the concept. Future work will include experimental validation of the system.

Chapter 3

Summary and Future Work

3.1 Point to Point Optical Link

In Chapter 1, a symmetric physical layer based secret key generation scheme is demonstrated for Point-to-Point Optical Link (PPOL) communication. The scheme exploits Polarization Mode Dispersion (PMD) as a random and inimitable channel characteristic. It is found that the randomness and security strength of generated cryptographic keys based on PMD is significantly high. In this thesis, it is presented that random modulation of a probe signal caused by PMD in a high-speed data communication network (40Gb/s and 60Gb/s) is reciprocal with average Pearson correlation coefficient of 0.862, despite the presence of optical nonlinearities, dispersion, and noise in the system. 128-bit symmetric cryptographic key has been successfully generated using the proposed scheme. Moreover, PMD based encryption keys passed the National Institute of Standards and Technology (NIST) tests. It is delineated through simulations with a 50km link that, with optimal key generation settings, symmetric keys can be generated with high randomness (high P-values for NIST randomness tests) and with sufficient generation rates (>50%). Furthermore, an attack model is defined of a non-invasive adversary intercepting at 10km into the link and found that the generated keys have high average key bit mismatch rates (>40%).

Further research will involve developing a more extensive attack model with higher position granularity along the fiber link. More efficient key generation algorithm will be developed and tested. Moreover, extensive experiments in a lab environment to verify the practicality of the model and key generation technique as shown in the following figure:

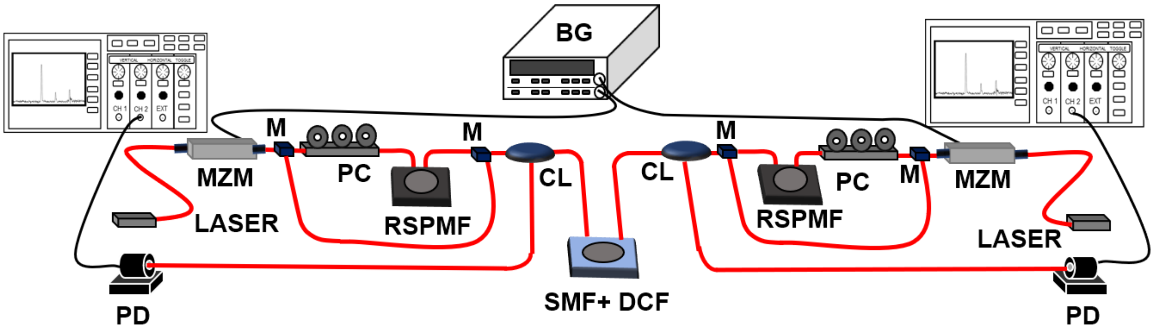


Figure 18. Experimental setup for PPOL key generation scheme. MZM: Mach Zehnder Modulator, PD: Photo Detector, CL: Circulator, BG: Bit sequence Generator, M: Mems switch.

In addition to abovementioned experimental validation, future research will include real-world key generation among different data centers or infrastructures to prove the key strength of the physical layer security generation scheme described in this thesis.

3.2 Inter-Vehicular Visible Light Communication

In chapter 2, A physical layer secret key generation scheme exploiting randomness of the road surface and the driving behavior is being proposed. The cryptographic key is generated based on real-world vehicle trajectory big-data to prove the concept. The research work proposed and simulated a novel symmetric key generation scheme which can be implemented in any existing vehicle to vehicle visible light communication. By analyzing and simulating numerous samples taken from NGSIM vehicle trajectory data, it is showed that natural driving behavior and road surface roughness can be exploited as a source of randomness to generate symmetric cryptographic security keys. As a proof of concept, 64 bit and 128-bit security keys are generated from the physical channel. Future work will include advanced attack model design, robust key generation algorithm development as well as real-world experiment to validate real-world application of this research work.

Bibliography

[1] I. U. Zaman, A. B. Lopez, M. A. A. Faruque, and O. Boyraz, “Polarization Mode Dispersion-Based Physical Layer Key Generation for Optical Fiber Link Security,” in *Advanced Photonics 2017 (IPR, NOMA, Sensors, Networks, SPPCom, PS) (2017), paper JTu4A.20*, 2017, p. JTu4A.20.

[2] I. U. Zaman, A. B. Lopez, M. A. A. Faruque, and O. Boyraz, “Physical Layer Cryptographic Key Generation by Exploiting PMD of an Optical Fiber Link,” *Journal of Lightwave Technology*, pp. 1–1, 2018.

[3] K. Shaneman and S. Gray, “Optical network security: technical analysis of fiber tapping mechanisms and methods for detection prevention,” in *IEEE MILCOM 2004. Military Communications Conference, 2004.*, 2004, vol. 2, pp. 711-716 Vol. 2.

[4] M. Furdek, N. Skorin-Kapov, M. Bosiljevac, and Z. Šipuš, “Analysis of crosstalk in optical couplers and associated vulnerabilities,” in *The 33rd International Convention MIPRO*, 2010, pp. 461–466.

[5] N. Patwari, J. Croft, S. Jana, and S. K. Kasera, “High-Rate Uncorrelated Bit Extraction for Shared Secret Key Generation from Channel Measurements,” *IEEE Transactions on Mobile Computing*, vol. 9, no. 1, pp. 17–30, Jan. 2010.

[6] H. Delfs and H. Knebl, *Introduction to Cryptography: Principles and Applications*. Springer Science & Business Media, 2007.

[7] N. R. Potlapally, S. Ravi, A. Raghunathan, and N. K. Jha, “A study of the energy consumption characteristics of cryptographic algorithms and security protocols,” *IEEE Transactions on Mobile Computing*, vol. 5, no. 2, pp. 128–143, Feb. 2006.

[8] H. Schweppe, Y. Roudier, B. Weyl, L. Apvrille, and D. Scheuermann, “Car2X Communication: Securing the Last Meter - A Cost-Effective Approach for Ensuring Trust in Car2X Applications Using In-Vehicle Symmetric Cryptography,” in *2011 IEEE Vehicular Technology Conference (VTC Fall)*, 2011, pp. 1–5.

[9] S. S. Gmbh and H. Str, “Automotive Security: Cryptography for Car2X Communication Dr. Torsten Schütze∗†,” 2011.

[10] C. W. O’donnell, G. E. Suh, and S. Devadas, “PUF-based random number generation,” in *In MIT CSAIL CSG Technical Memo 481 (http://csg.csail.mit.edu/pubs/memos/Memo-481/Memo-481.pdf*, 2004, p. 2004.

[11] T. S. Humble, “Quantum security for the physical layer,” *IEEE Communications Magazine*, vol. 51, no. 8, pp. 56–62, Aug. 2013.

[12] K. Lim, H. Ko, C. Suh, and J.-K. K. Rhee, “Security analysis of quantum key distribution on passive optical networks,” *Opt. Express, OE*, vol. 25, no. 10, pp. 11894–11909, May 2017.

[13] M. Rostami, J. B. Wendt, M. Potkonjak, and F. Koushanfar, “Quo Vadis, PUF?: Trends and Challenges of Emerging Physical-disorder Based Security,” in *Proceedings of the Conference on Design, Automation & Test in Europe*, 3001 Leuven, Belgium, Belgium, 2014, pp. 352:1–352:6.

[14] K. Kravtsov, Z. Wang, W. Trappe, and P. R. Prucnal, “Physical layer secret key generation for fiber-optical networks,” *Optics Express*, vol. 21, no. 20, p. 23756, Oct. 2013.

[15] J. Wan, A. B. Lopez, and M. A. Al Faruque, “Exploiting Wireless Channel Randomness to Generate Keys for Automotive Cyber-physical System Security,” in *Proceedings of the 7th International Conference on Cyber-Physical Systems*, Piscataway, NJ, USA, 2016, pp. 13:1–13:10.

[16] K. Guan, J. Cho, and P. J. Winzer, “Physical layer security in fiber-optic MIMO-SDM systems: An overview,” *Optics Communications*, vol. 408, no. Supplement C, pp. 31–41, Feb. 2018.

[17] I. U. Zaman, A. B. Lopez, M. A. A. Faruque, and O. Boyraz, “A Physical Layer Security Key Generation Technique for Inter-Vehicular Visible Light Communication,” in *Advanced Photonics 2017 (IPR, NOMA, Sensors, Networks, SPPCom, PS) (2017), paper SpTu1F.3*, 2017, p. SpTu1F.3.

[18] J. Wan, A. Lopez, and M. A. A. Faruque, “Physical Layer Key Generation: Securing Wireless Communication in Automotive Cyber-Physical Systems,” 2018.

[19] M. P. Fok and P. R. Prucnal, “Low-latency nonlinear fiber-based approach for data encryption and anti-jamming in optical network,” in *LEOS 2008 - 21st Annual Meeting of the IEEE Lasers and Electro-Optics Society*, 2008, pp. 743–744.

[20] S. Ten and M. Edwards, “An Introduction to Fundamentals of PMD in Fibers,” *WP5051*.

[21] F. Curti, B. Daino, G. D. Marchis, and F. Matera, “Statistical treatment of the evolution of the principal states of polarization in single-mode fibers,” *Journal of Lightwave Technology*, vol. 8, no. 8, pp. 1162–1166, Aug. 1990.

[22] M. Brodsky, N. J. Frigo, M. Boroditsky, and M. Tur, "Polarization Mode Dispersion of Installed Fibers," *J. Lightwave Technol., JLT*, vol. 24, no. 12, pp. 4584–4599, Dec. 2006.

[23] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, "Radio-telepathy: Extracting a Secret Key from an Unauthenticated Wireless Channel," in *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking*, New York, NY, USA, 2008, pp. 128–139.

[24] A. L. Bassham (NIST) *et al.*, "SP 800-22 Rev. 1a, A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications." [Online]. Available: https://csrc.nist.gov/publications/detail/sp/800-22/rev-1a/final. [Accessed: 21-Feb-2018].

[25] G. Blinowski, "Security issues in visible light communication systems," *IFAC-PapersOnLine*, vol. 48, no. 4, pp. 234–239, Jan. 2015.

[26] B. Schoettle, M. Sivak, and M. J. Flannagan, "High-Beam and Low-Beam Headlighting Patterns in the U.S. and Europe at the Turn of the Millennium," *ResearchGate*, Mar. 2002.

[27] V. Punzo, M. T. Borzacchiello, and B. Ciuffo, "On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 1243–1262, Dec. 2011.

[28] K. Bogsjö, K. Podgórski, and I. Rychlik, "Models for road surface roughness," *Vehicle System Dynamics*, vol. 50, no. 5, May 2012.

[29] B. Schoettle, "A market-weighted description of low-beam headlighting patterns in the U.S.: 2004," *ResearchGate*.

[30] P. Luo, Z. Ghassemlooy, H. L. Minh, E. Bentley, A. Burton, and X. Tang, "Fundamental analysis of a car to car visible light communication system," in *2014 9th International Symposium on Communication Systems, Networks Digital Sign (CSNDSP)*, 2014, pp. 1011–1016.

[31] P. Luo, Z. Ghassemlooy, H. L. Minh, E. Bentley, A. Burton, and X. Tang, "Performance analysis of a car-to-car visible light communication system," *Appl. Opt., AO*, vol. 54, no. 7, pp. 1696–1706, Mar. 2015.