sensors

MDPI

Article

Secrecy Capacity of a Class of Erasure Wiretap Channels in WBAN

Bin Wang [1], Jun Deng [1,*], Yanjing Sun [1,2], Wangmei Guo [3] and Guiguo Feng [3]

[1] School of Communication Engineering, Xi'an University of Science and Technology, Xi'an 710054, China; wangbin@mail.xidian.edu.cn (B.W.); yanjingsun_cn@163.com (Y.S.)
[2] School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China
[3] State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 10071, China; wangmeiguo@xidian.edu.cn (W.G.); fengguiguo@163.com (G.F.)
* Correspondence: dengj518@xust.edu.cn

Received: 31 October 2018; Accepted: 22 November 2018; Published: 26 November 2018

Abstract: In wireless body area networks (WBANs), the secrecy of personal health information is vulnerable to attacks due to the openness of wireless communication. In this paper, we study the security problem of WBANs, where there exists an attacker or eavesdropper who is able to observe data from part of sensors. The legitimate communication within the WBAN is modeled as a discrete memoryless channel (DMC) by establishing the secrecy capacity of a class of finite state Markov erasure wiretap channels. Meanwhile, the tapping of the eavesdropper is modeled as a finite-state Markov erasure channel (FSMEC). A pair of encoder and decoder are devised to make the eavesdropper have no knowledge of the source message, and enable the receiver to recover the source message with a small decoding error. It is proved that the secrecy capacity can be achieved by migrating the coding scheme for wiretap channel II with the noisy main channel. This method provides a new idea solving the secure problem of the internet of things (IoT).

Keywords: wiretap channel II; secrecy capacity; finite state Markov erasure wiretap channel; WBAN

1. Introduction

Due to the openness of wireless communication, the personal health information, which is exchanged on the wireless channel in WBAN, is readily fetched and attacked by hackers. To address this issue, there are usually two ways to enhance the security of wireless communications: one is the security guaranteed by information theory in Refs. [1–3], another is the security verified by the computational complexity in Refs. [4,5]. In this paper, we aim to study the secure transmission problem in WBAN on the basis of the information theory. Here, the secure transmission indicates the way to code the transmitted data so that the attackers cannot get the data. The concept of wiretap channel is introduced by Wyner in Ref. [6]. In his model, the source message was sent to the targeted user via a discrete memoryless channel (DMC). Meanwhile, an eavesdropper was able to tap the transmitted data via a second DMC. It was supposed that the eavesdropper knew the encoding scheme and decoding scheme. The object was to find a pair of encoder and decoder such that the eavesdropper's level of confusion on the source message was as high as possible, while the receiver could recover the transmitted data with a small decoding error. Wyner's wiretap channel model is called the discrete memoryless wiretap channel, since the main channel output was taken as the input of the wiretap channel in Ref. [7].

After Wyner's pioneering work, the models of wiretap channels have been studied from various aspects. Csiszar and Korner considered a more general wiretap channel model called the broadcast

channels with confidential messages (BCCs) in Ref. [8]. The wiretap channel was not necessarily a degraded version of the main channel. Moreover, they also considered the case where public data was supposed to be broadcasted through both main channel and wiretap channel. The degraded wiretap channels with discrete memoryless side information accessed by the encoder were considered in Refs. [9–11]. BCCs with causal side information were studied in Ref. [12]. Communication models with channel states known at the receiver were considered in Refs. [13,14]. Ozarow and Wyner considered another wiretap channel model called wiretap channel of type II [15]. The secrecy capacity was established there. In that model, the source data was encoded into N digital symbols and transmitted to the targeted user through a binary noiseless channel. Meanwhile, the eavesdropper was able to observe an arbitrary μ-subcollection of those symbols.

In the last few decades, a lot of capacity problems related to the wiretap channel II were studied. A special class of non-DMC wiretap channel was studied in Ref. [16]. The main channel was a DMC instead of noiseless, and the eavesdropper observed $\mu < N$ digital symbols through a uniform distribution. An extension of wiretap channel II was studied in Ref. [17], where the main channel was a DMC and the eavesdropper was able to observe μ digital bits through arbitrary strategies.

The model of finite-state Markov channel was first introduced by Gilbert [18] and Elliott [19]. They studied a kind of Markov channel model with two states, which is known as the Gilbert–Elliott channel. In their channel model, one state was related to a noiseless channel and the other state was related to a totally noisy channel. Wang in Ref. [20] extended the Gilbert–Elliott channel and considered the case with finite states.

This paper discusses finite-state Markov erasure wiretap channel (FSME-WTC) (see Figure 1). In this new model, the source data W is encoded into N digital symbols, denoted by X^N, and transmitted to the targeted user through a DMC. The eavesdropper is able to observe the transmitted symbols through a finite-state erasure Markov channel (FSMEC). Secrecy capacity of this new communication model is established, based on the coding scheme devised by the authors in Ref. [17].

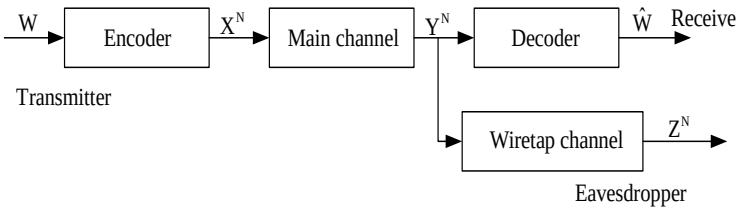

Figure 1. Communication model of degraded wiretap channels.

The model of FSME-WTC can be applied to model the security problem of WBAN readily. Let us suppose that there are N sensors in WBAN. Then, we can treat the collection of symbols obtained from the sensors as a digital sequence of length N transimitted over an imaginary channel. The imaginary channel is not DMC because the symbols from the sensors are correlated. Markov chain is an important model to characterize the correlation of random variables since it will not bring too much complexity of the system. The wiretap channel is set as an erasure channel to model the situation where the attacker in WBAN is able to tap data from only part of the sensors. Thus, our model of FSME-WTC is to ensure that the attacker is not able to get any information from the WBAN when he/she can only observe data from at most $N\alpha$ sensors.

The importance of this model is obvious. As the technology of 5G advances towards the stage of commercial applications, wireless networks are becoming more and more significant in our daily lives [21,22]. Therefore, the security problem of wireless communication is critical from the aspects of both theory and engineering. Meanwhile, the finite state Markov channel is a common model to character the properties of wireless communication. Hence, the results of this paper are meaningful to many kinds of wireless networks with high confidentiality requirements, such as WBAN and IoT.

The remainder of this paper is organized as follows. The formal statements of Finite-state Markov Erasure Wiretap Channel and the capacity results are given in Section 2 (see also Figure 1). The secrecy capacity of this model is established in Theorem 1. Some concrete examples of this communication model are given in Section 3. The converse part of Theorem 1, relying on Fano's inequality and Proposition1, is proved in Section 4. The direct part of Theorem 1, based on Theorem 1 in [17], is proved in Section 5. Section 6 gives the proof of Proposition 1, and Section 7 finally concludes this paper.

2. Notations, Definitions and the Main Results

Throughout this paper, $\mathbb{N}$ is the set of positive integers. $[1:N] = \{1, 2, \ldots, N\}$ is the set of positive integers no greater than N for any $N \in \mathbb{N}$. For any index set $\mathcal{I} \subseteq [1:N]$ and random vector $Y^N = (Y_1, Y_2, \ldots, Y_N)$, denote by $Y_{\mathcal{I}}^N = (Y_1', Y_2', ..., Y_N')$ the "projection" of Y^N onto the index set $\mathcal{I}$ such that $Y_i = Y_i'$ for all $i \in \mathcal{I}$, and $Y_i =?$, otherwise.

Let $\mathcal{Y}$ be any finite alphabet not containing the "error" letter ? and $\mathcal{Y}_{\mathcal{I}}^N = \{(y_1, y_2, ..., y_N) : y_i \in \mathcal{Y}$ for $i \in \mathcal{I}$, and $i =?$ for $i \notin \mathcal{I}\}$. It follows that $Y_{\mathcal{I}}^N$ is distributed on $\mathcal{Y}_{\mathcal{I}}^N$ for any random vector Y^N over $\mathcal{Y}^N$.

Example 1. *Let $N = 5$, $\mathcal{I} = 1, 3$ and $\mathcal{X} = 0, 1$. Then,*

$$\mathcal{X}_{\mathcal{I}}^N = \{(0?0??), (0?1??), (1?0??), (1?1??)\}.$$

Let $X^N = (X_1, X_2, X_3, X_4, X_5)$ be an arbitrary random vector distributed on $\mathcal{X}^N$. Then, the random vector $X_{\mathcal{I}}^N = (X_1, ?, X_3, ?, ?)$ is distributed on $\mathcal{X}_{\mathcal{I}}^N$.

Definition 1. (Encoder) *Let the source message W be uniformly distributed on a certain message set $\mathcal{W}$. The (stochastic) encoder q_E is specified by a matrix of conditional probability $q_E(x^N|w)$ with $x^N \in \mathcal{X}^N$ and $w \in \mathcal{W}$. The value of $q_E(x^N|w)$ specifies the probability that we encode message w encoded into the sequence x^N.*

Definition 2. (Main channel) *The main channel is a DMC, whose input alphabet is $\mathcal{X}$ and output alphabet is $\mathcal{Y}$, where $? \notin \mathcal{X} \cup \mathcal{Y}$. The transition probability matrix of the main channel is denoted by $Q_{MC}(y|x)$ with $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. The input and output of the main channel are denoted by X^N and Y^N, respectively. For any $x^N \in \mathcal{X}^N$ and $y^N \in \mathcal{Y}^N$, it follows that*

$$Pr\{X^N = x^N, Y^N = y^N\} = Pr\{X^N = x^N\} Q_M C(y^N|x^N),$$

where

$$Q_{MC}(y^N|x^N) = \prod_{i=1}^{N} Q_{MC}(y_i|x_i).$$

Remark 1. *From the property of DMC, it holds that*

$$H(Y^N|X^N) = \sum_{i=1}^{N} H(Y_i|X_i).$$

Definition 3. (Wiretap channel) *Let $T_n, n \in N$ be the channel state of FSMEC at time n satisfying that $T_1 \to T_2 \to ... \to T_N \to ...$ forms a Markov chain. The transition of channel states is homogeneous, i.e., the conditional probability $Pr\{T_n = t_n|T_{n-1} = t_{n-1}\}$ is independent from the time index n. Moreover, the channel states are stationary, i.e., $T_1, T_2, ..., T_N, ...$ share a generic probability distribution p_T on a common*

finite set $\mathcal{T}$ of channel states. Moreover, let $Q_T(t'|t)$ be the probability that the state at the next time slot is changed to t' when the state is t currently. It follows that

$$Pr\{T^N = t^N\} = p_T(t_1) \cdot \prod_{i=2}^{N} Q_T(t_i|t_i - 1)$$

for $t^N \in \mathcal{T}^N$. The input of FSMEC is a digital sequence Y^N, which is actually the main channel output. Denote by Z^N the wiretap channel output. For each time slot n, the channel is either totally noisy, i.e., $Z_n =?$ or totally noiseless, i.e., $Z_n = Y_n$, which depends on the value of T_n. Thus, the channel output Z_n is totally determined by the channel input Y_n and the channel state T_n. Let $\mathcal{T}_1$ be the set of states under which the channel is noiseless. Then, it follows that $\mathcal{T}_0 = \mathcal{T} - \mathcal{T}_1$ contains the states where the channel is totally noisy. Denote by $Q_{WC}(z|y,t)$ the probability that the channel outputs z when the channel input is y and the channel state is t. It follows that

$$Q_{WC}(z|y,t) = \begin{cases} \delta(z,y) t \in \mathcal{T}_1, \\ \delta(z,?), t \in \mathcal{T}_0, \end{cases}$$

where

$$\delta(a,b) = \begin{cases} 1, a = b, \\ 0, a \neq b. \end{cases}$$

For any $y^N \in \mathcal{Y}^N$, $z^N \in \mathcal{Z}^N$ and $t^N \in \mathcal{T}^N$, it is readily obtained that

$$Pr\{Y^N = y^N, Z^N = z^N | T^N = t^N\} = Pr\{Y^N = y^N\} \prod_{i=1}^{N} Q_{WC}(z_i|y_i,t_i).$$

Remark 2. *Throughout this paper, it is supposed that T^N is independent from W, X^N and Y^N.*

Proposition 1. *$X^n \to Z^n \to T^n$ forms a Markov chain for every $1 \leq n \leq N$.*

Proof. The proof of Proposition 1 is given in Section 6. Proposition 1 would be used to establish the converse part of Theorem 1 (see Section 4). □

Definition 4. (Decoder) *The decoder is specified by a mapping $f_D : Y^N \to W$. To be particular, the estimation of the source message is $\hat{W} = Y^N$, where Y^N is the main channel output. The average decoding error probability is denoted by $P_e = Pr\{W \neq \hat{W}\}$.*

Definition 5. (Achievability) *A positive real number R is said to be achievable, if, for any real number $\varepsilon > 0$, one can find an integer N_0 such that, for any $N > N_0$, there exists a pair of encoder and decoder of length of length N satisfying that*

$$\frac{1}{N} log|\mathcal{W}| > R - \varepsilon, \frac{1}{N} I(W; Y^N) < \varepsilon \text{ and } P_e < \varepsilon. \tag{1}$$

Definition 6. (Secrecy capacity) *A real number C_s is said to be the secrecy capacity of the communication model if it is achievable for every $0 \leq R \leq C_s$ and unachievable for every $R > C_s$.*

Theorem 1. *Let B_n be the function of T_n defined in Definition 3 such that $B_n = 1$ if $T_n \in \mathcal{T}_1$, and $B_n = 0$, otherwise. If it follows that*

$$\lim_{N \to \infty} Pr\{|\frac{1}{N} \sum_{n=1}^{N} B_n - \alpha| < \iota\} = 1 \tag{2}$$

for any $\iota > 0$, the secrecy capacity of the communication model in Figure 1 is $(1-\alpha)C_M$, where C_M is the capacity of the main channel, i.e.,

$$C_M = \max_{P_X} I(X;Y). \tag{3}$$

Proof. The proof of Theorem 1 is divided into the following two parts. The first part, given in Section 4, proves that every achievable real number R must satisfy $R \leq (1-\alpha)C_M$, which is the converse half of the theorem. The second part, given in Section 5, proves that every real number R satisfying $0 \leq R \leq (1-\alpha)C_M$ is achievable, which is the direct half. □

Theorem 1 claims that, if the Markov chain $\{T_n\}$ satisfies Label (2), then the secrecy capacity of the wiretap channel model depicted in Figure 1 is $(1-\alpha)C_M$. In the rest of this section, we will introduce a class of Markov chains satisfying (2) in Theorem 2, and provide the secrecy capacity of the related wiretap channel model in Corollary 1.

A stationary Markov chain is call ergodic if, for each pair of states $t, t' \in \mathcal{T}$, it is possible to go from state t to t' in expected finite steps. One can prove that, if a Markov of chain is ergodic, the stationary probability distribution of the state is unique.

Theorem 2. (Law of Large Number for Markov Chain) *If the Markov chain $\{T_n\}$ is ergodic, let π be the unique stationary distribution of the state. Then, it follows that*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} I(T_n = t) = \pi(t)$$

for each channel state t, where $I(T_n = t)$ is 1 or 0, indicating whether $T_n = t$ is true or not.

With the theorem above, we immediately obtain that

Corollary 1. *If the Markov chain $\{T_n\}$ is ergodic with the unique stationary distribution π over $\mathcal{T}$, then the secrecy capacity of the wiretap channel model depicted in Figure 1 is given by*

$$C_s = (1 - \pi(\mathcal{T}_1))C_M,$$

where C_M is the capacity of the main channel, and

$$\pi(\mathcal{T}_1) = \sum_{t\in\mathcal{T}_1} \pi(t).$$

3. Examples

This section gives two simple examples of FSMEC defined in Definition 3. Example 2 is for discrete memoryless erasure channel (DMEC) and Example 3 is for a simple two-state FSMEC.

Example 2. *Suppose that the set of channel states $T = 0, 1$ with $T_1 = 1$ and $T_0 = 0$. Meanwhile, let*

$$p_T(0) = Q_t(0|0) = Q_t(0|1) = 1 - \alpha \tag{4}$$

and

$$p_T(1) = Q_t(1|0) = Q_t(1|1) = \alpha.$$

The state transition diagram of the channel states in this example is depicted in Figure 2. It is obvious that the FSMEC is in fact specialized into a DMEC with the transition probability

$$Q_{WC}(z|y) = \begin{cases} \alpha, & z = y, \\ 1-\alpha, & z =?, \\ 0, & otherwise. \end{cases}$$

From Theorem 2 in Ref. [6], the secrecy capacity of the communication model in Figure 1, with DMEC as the wiretap channel, is

$$\begin{aligned} &\max_{P_X} I(X:Y|Z) \\ &= \max_{P_X} I(X;Y) - I(X;Z) \\ &\overset{(a)}{=} \max_{P_X} I(X;Y|T) - I(X;Z|T) \\ &= \max_{P_X} \{[I(X;Y|T=0) - I(X;Z|T=0)]P_T(0) + [I(X;Y|T=1) - I(X;Z|T=1)]P_T(1)\} \\ &\overset{(b)}{=} \max_{P_X} I(X;Y|T=0)P_T(0) \\ &\overset{(c)}{=} \max_{P_X} I(X;Y)P_T(0) \\ &\overset{(d)}{=} (1-\alpha)C_M, \end{aligned}$$

where X and Y are the input and output of the main channel, respectively, and Z is the output of the wiretap channel under the channel state T; (a) follows from the facts that $X \to Z \to T$ forms a Markov chain (cf. Proposition 1) and T is independent from X and Y; (b) follows from the fact that $Y = Z$ when $T = 1$, and Z is determined when $T = 0$; (c) follows from the assumption that T is independent from X and Y; and (d) follows from (3) and (4).

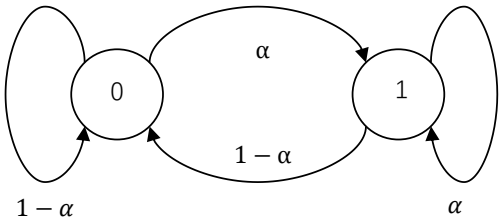

Figure 2. State transition diagram of discrete memoryless erasure channels.

Clearly, Formula (2) holds with $B_n = T_n$. Thus, in this case, the result of Theorem 1 in this paper coincides with that of Theorem 2 in Ref. [6].

Example 3. *Let $\mathcal{T} = 0, 1, \mathcal{T}_1 = 1, \mathcal{T}_0 = 0$,*

$$p_T(0) = p_T(1) = \frac{1}{2},$$

$$Q_t(0|0) = Q_t(1|1) = p,$$

$$Q_t(1|0) = Q_t(0|1) = 1 - p,$$

and $B_n = T_n$. We arrive at a simple two-state Markov erasure channel whose transition diagram is depicted in Figure 3. Furthermore, observe that

$$\begin{aligned} D\left[\sum_{n=1}^{N} B_n\right] &= D\left[\sum_{n=1}^{N} T_n\right] \\ &= E\left[\left(\sum_{n=1}^{N} T_n\right)^2\right] - \left(E\left[\sum_{n=1}^{N} T_n\right]\right)^2 \\ &= \left(\sum_{n=1}^{N}\sum_{m=1}^{N} E[T_n T_m]\right) - \frac{N^2}{4} \\ &= \sum_{n=1}^{N} \frac{(N-n)}{2}(2p-1)^n, \end{aligned}$$

where the last equality follows because $E[T_n T_m] = \frac{1}{2}$ *when* $m = n$*, and*

$$\begin{aligned} E[T_n T_m] &= Pr\{T_m = 1, T_n = 1\} \\ &= Pr\{T_m = 1, T_{n-1} = 1, T_n = 1\} + Pr\{T_m = 1, T_{n-1} = 0, T_n = 1\} \\ &= Pr\{T_m = 1, T_{n-1} = 1\} Q_T(1|1) + Pr\{T_m = 1, T_{n-1} = 0\} Q_T(1|0) \\ &= E[T_m T_{n-1}]p + (\frac{1}{2} - E[T_m T_{n-1}])(1-p) \\ &= \mathbf{(2p-1)E[T_m T_{n-1}] + \frac{1-p}{2}} \\ &= ... \\ &= (2p-1)^{n-m} E[T_m T_m] + \frac{1-p}{2}\sum_{i=1}^{n-m-1}(2p-1)^i \\ &= \frac{1+(2p-1)^{n-m}}{4} \end{aligned}$$

when $m < n$*. It is obvious that*

$$\lim_{N\to\infty} \frac{1}{N^2} D\left[\sum_{n=1}^{n} B_n\right] = 0$$

for $0 < p < 1$*. Formula (2) is then established immediately from the Markov Large Number Law. Applying Theorem 1, the secrecy capacity of the communication model in this case is* $\frac{1}{2}C(p)$*. Figure 4 shows the relationship between the secrecy capacity and the crossover probability p in this example.*

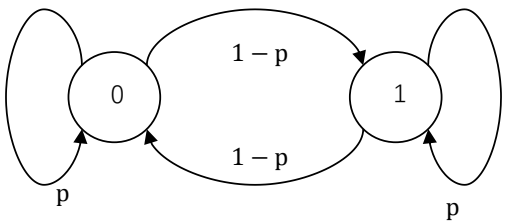

Figure 3. State transition diagram of a two-state Markov chain.

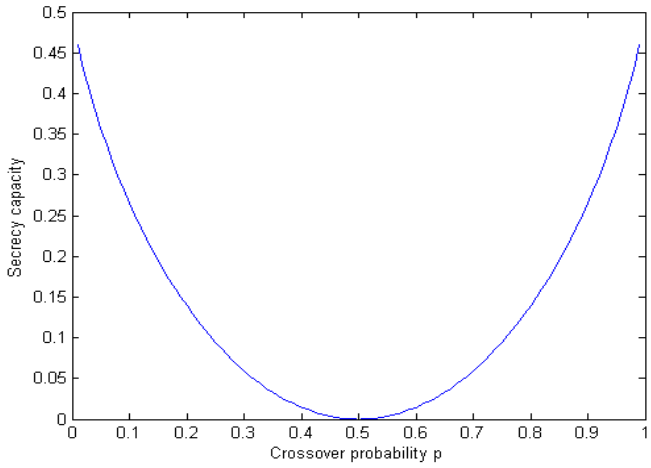

Figure 4. Secrecy capacity of the two-state Markov erasure wiretap channel in Example 3.

4. Converse Half of Theorem 1

This section proves that every achievable real number R must satisfy $R \leq (1-\alpha)C_M$. The proof is based on Fano's inequality (cf. Formula (76) in Ref. [6]) and Proposition 1.

For any give $\iota > 0$ and $\varepsilon > 0$, Formula (2) indicates that

$$Pr\{\frac{1}{N}\sum_{n=}^{N} B_n > \alpha - \iota\} > 1 - \varepsilon$$

or equivalently

$$Pr\{|\mathcal{I}(T^N)| > N(\alpha - \iota)\} > 1 - \varepsilon \tag{5}$$

when N is sufficiently large, where

$$\mathcal{I}(t^N) = \{n \in [1:N] : t_n \in \mathcal{T}_1\}.$$

Suppose that there exists a code of length N satisfying (1), i.e.,

$$\frac{1}{N}\log|\mathcal{W}| > R - \varepsilon, \frac{1}{N}I(W;Z^N) < \varepsilon \text{ and } P_e < \varepsilon.$$

Then, we have

$$NR < \log|\mathcal{W}| + N\varepsilon = H(W) + N\varepsilon = I(W;Y^N) + H(W|Y^N) + N\varepsilon < I(W;Y^N) + N\delta(P_e) + N\varepsilon,$$

where $\delta(P_e) \to 0$ as $P_e \to 0$, and the last inequality follows from the Fano's inequality. Since $I(W;Z^N) < N\varepsilon$, the formula above indicates that

$$NR < I(W;Y^N) - I(W;Z^N) + N\delta(P_e) + 2N\varepsilon. \tag{6}$$

The value of $I(W;Y^N) - I(W;Z^N)$ is upper bounded by

$$\begin{aligned} & I(W;Y^N) - I(W;Z^N) \\ & \overset{(a)}{=} I(W;Y^N|Z^N) \\ & \overset{(b)}{\leq} I(X^N;Y^N|Z^N) \\ & = I(X^N;Y^N) - I(X^N;Z^N) \\ & \overset{(c)}{=} I(X^N;Y^N|T^N) - I(X^N;Z^N|T^N), \end{aligned} \tag{7}$$

where (a) and (b) follow from the fact that $W \to X^N \to Y^N \to Z^N$ forms a Markov chain, and (c) follows from Proposition 1 and the fact that T^N is independent from X^N and Y^N.

For any $t^N \in \mathcal{T}^N$, denoting $Z^N(t^N) = Y^N_{\mathcal{I}(t^N)}$, Formula (7) is further deduced by

$$\begin{aligned} & I(W;Y^N) - I(W;Z^N) \\ & \leq I(X^N;Y^N|T^N) - I(X^N;Z^N|T^N) \\ & \overset{(a)}{=} I(X^N;Y^N|Z^N,T^N) \\ & = \sum_{t^N \in \mathcal{T}^N} \Big(I(X^N;Y^N|Z^N,T^N = t^N) \cdot Pr\{T^N = t^N\} \Big) \\ & = \sum_{t^N \in \mathcal{T}^N} \Big(I(X^N;Y^N|Z^N(t^N),T^N = t^N) \cdot Pr\{T^N = t^N\} \Big) \\ & \overset{(b)}{=} \sum_{t^N \in \mathcal{T}^N} I(X^N;Y^N|Z^N(t^N)) \cdot Pr\{T^N = t^N\}, \end{aligned} \tag{8}$$

where (a) follows because $X^N \to Y^N \to Z^N$ forms a Markov chain when given T^N, and (b) follows because X^N, Y^N and $Z^N(t^N) = Y^N_{\mathcal{I}(t^N)}$ are independent from T^N. For any fixed $t^N \in \mathcal{T}^N$, denote $\tilde{Z}^N = Z^N(t^N)$. On account of the chain rule, we have

$$H(Y^N) = \sum_{n=1}^{N} H(Y_n|Y^{n-1}), \tag{9}$$

$$H(\tilde{Z}^N) = \sum_{n=1}^{N} H(\tilde{Z}_n|\tilde{Z}^{n-1}), \tag{10}$$

and

$$\begin{aligned} H(\tilde{Z}^N|X^N) &= \sum_{n=1}^{N} H(\tilde{Z}_n|\tilde{Z}^{n-1},X^N) \\ &\leq \sum_{n=1}^{N} H(\tilde{Z}_n|X^n). \end{aligned} \tag{11}$$

Moreover, from the property of DMC, Remark 1 yields

$$H(Y^N|X^N) = \sum_{n=1}^{N} H(Y_n|X_n). \tag{12}$$

Combining Formulas (9)–(12), it follows that

$$I(W;Y^N) - I(W;Z^N(t^N)) \leq \sum_{n=1}^{N} (H(Y_n|Y^{n-1}) - H(\tilde{Z}_n|\tilde{Z}^{n-1}) - H(Y_n|X_n) + H(\tilde{Z}^n|X_n)). \tag{13}$$

Considering that $\tilde{Z}^(n-1) \to Y^(n-1) \to Y_n \to \tilde{Z}_n$ forms a Markov chain, we have

$$I(Y^{n-1};Y_n) \geq I(\tilde{Z}^{n-1};\tilde{Z}_n)$$

or equivalently

$$H(Y_n) - H(\tilde{Z}_n) \geq H(Y_n|Y^{n-1}) - H(\tilde{Z}_n|\tilde{Z}^{n-1}).$$

Substituting the formula above into Formula (13), we have

$$\begin{aligned} & I(W;Y^N) - I(W:Z^N(t^N)) \\ &= \sum_{n=1}^{N} (H(Y_n) - H(\tilde{Z}_n) - H(Y_n|X_n) + H(\tilde{Z}_n|X_n)) \\ &= \sum_{n=1}^{N} (I(X_n;Y_n) - I(X_n;\tilde{Z}_n)). \end{aligned} \tag{14}$$

Noticing that

$$I(X_n;\tilde{Z}_n) = \begin{cases} 0, t_n \in \mathcal{T}_1, \\ I(X_n;Y_n), t \in \mathcal{T}_0. \end{cases}$$

Formula (14) is further deduced by

$$I(X^N;Y^N) - I(X^N;Z^N(t^N)) \leq \sum_{n=1}^{N} I(X_n;Y_n) - I(X_n;\tilde{Z}_n) = \sum_{n \notin \mathcal{I}(t^N)} I(X_n;Y_n) \leq (N - |\mathcal{I}(t^N)|)C_M.$$

Substituting the formula above with Formula (8) gives

$$\begin{aligned} & I(X^N;Y^N) - I(X^N;Z^N) \\ & \leq \sum_{t^N \in T^N} I(X^N;Y^N|Z^N(t^N)) \Pr\{T^N = t^N\} \\ & \leq \sum_{t^N \in T^N} (N - |\Im(t^N)|)C_M \Pr\{T^N = t^N\} \\ & \leq \Pr\{\Im(T^N) \geq N(\alpha - \iota)\}N(1 - \alpha + \iota)C_M + \Pr\{\Im(T^N) < N(\alpha - \iota)\}NC_M \\ & \leq N(1 - \alpha + \iota + 2\varepsilon)C_M, \end{aligned}$$

where the last inequality follows from (5). Combining (6) and the formula above yields

$$R < 1 - \alpha + \iota + 4\varepsilon + \delta(P_e).$$

$R \leq 1 - \alpha$ is finally established by letting ι, ε and P_e converge to 0. This completes the proof of converse half.

5. Direct Half of Theorem 1

This section proves that every real number R satisfying $0 < R \leq (1-\alpha)C_M$ is achievable, which is the direct half of Theorem 1. It suffices to prove the achievability of $(1-\alpha)C_M$. More precisely, for any given $\varepsilon > 0$, we need to prove the existence of the encoder–decoder pair (q_E, f_D) such that

$$\frac{1}{N}\log|\mathcal{W}| > R-\varepsilon, \frac{1}{N}I(W;Z^N) < \varepsilon \text{ and } P_e < \varepsilon.$$

The proof is based on the following theorem.

Theorem 3. *(Theorem 1 in Ref. [17]). Let a real number* $0 < \alpha' 1$ *be fixed and given. For any* $N \in \mathbb{N}$ *and* $\mu = N\alpha'$*, denote*

$$\Im_\mu = \Im_\mu(N) = \{\mathcal{I} \subseteq [1:N] : |\mathcal{I}| = \mu\}.$$

Then, for any real numbers $\varepsilon' > 0$ *and* $0 < R < (1-\alpha')C_M$*, one can construct a code of length N over the DMC defined in Definition 2 such that*

$$\frac{1}{N}\log|\mathcal{W}| > R-\varepsilon', \max_{\mathcal{I}\in\Im_\mu} I(W;Y_{\mathcal{I}}^N) < \varepsilon', P_e < \varepsilon'$$

when N is sufficiently large.

Proof. Let

$$\alpha' = \alpha + \iota$$

and

$$R = (1-\alpha-2\iota)C_M < (1-\alpha')C_M$$

for a small $\iota > 0$. Suppose that (q_E, f_D) is a code of length N satisfying

$$\frac{1}{N}\log|\mathcal{W}| > R-\varepsilon' > (1-\alpha-2\iota)C_M - \varepsilon'$$
$$\max_{\mathcal{I}\in\Im_\mu} I(W;Y_{\mathcal{I}}^N) < \varepsilon' \text{ and } P_e < \varepsilon'.$$

Applying the code (q_E, f_D) to the communication model in Figure 1, it is already satisfied that

$$\frac{1}{N}\log|\mathcal{W}| > (1-\alpha)C_M - \varepsilon \text{ and } P_e < \varepsilon,$$

when ε' and ι are sufficiently small. To establish $\frac{1}{N}I(W;Z^N) < \varepsilon$, let the value of N be sufficiently large such that

$$Pr|\mathcal{I}(T^N)| < N(\alpha+\iota) > 1-\varepsilon'. \tag{15}$$

The value of $I(W;Z^N)$ is upper bounded by

$$\begin{aligned}
&I(W;Z^N)\\
&\overset{(a)}{\leq} I(W;Z^N|T^N)\\
&= \sum_{t^N\in T^N} I(W;Z^N|T^N=t^N)\Pr\{T^N=t^N\}\\
&= \sum_{t^N\in T^N} I(W;Y^N_{\Im(t^N)}|T^N=t^N)\Pr\{T^N=t^N\}\\
&\overset{(b)}{=} \sum_{t^N\in T^N} I(W;Y^N_{\Im(t^N)})\Pr\{T^N=t^N\}\\
&= \sum_{t^N:|\Im(t^N)|<N(\alpha+\iota)} I(W;Y^N_{\Im(t^N)})\Pr\{T^N=t^N\}\\
&+ \sum_{t^N:|\Im(t^N)|\geq N(\alpha+\iota)} I(W;Y^N_{\Im(t^N)})\Pr\{T^N=t^N\}\\
&\overset{(c)}{\leq} \varepsilon' + NC_M\Pr\{|\Im(T^N)|\geq N(\alpha+\iota)\}\\
&\overset{\mathrm{d}}{\leq} \varepsilon'(1+NC_M),
\end{aligned}$$

where (a) follows because W is independent from Z^N; (b) follows because $Y^N_{\mathcal{I}(t^N)}$ is independent from T^N; (c) follows because $I(W;Y^N_{\mathcal{I}(t^N)})\leq H(W)\leq NC_M$ when $|\mathcal{I}(t^N)|>N(\alpha+\iota)$, and

$$I(W;Y^N_{\mathcal{I}}(t^N))\leq \max_{\mathcal{I}\in\Im_\mu} I(W;Y^N_{\mathcal{I}})<\varepsilon'$$

when $|\mathcal{I}(t^N)|<N(\alpha+\iota)$; and (d) follows from Formula (15). Consequently,

$$\frac{1}{N}I(W;Z^N)\leq\frac{\varepsilon'}{N}(1+NC_M)<\varepsilon'(1+C_M)<\varepsilon$$

when ε' is sufficiently small. The proof of the direct half is completed. □

6. Proof of Proposition 1

This section proves that $X^n\to Z^n\to T^n$ forms a Markov chain for every $n\in N$, which is Proposition 1. It suffices to prove that

$$\begin{aligned}
&Pr\{X^n=x^n,Z^n=z^n,T^n=t^n\}\Pr\{Z^n=z^n\}\\
&=\Pr\{X^n=x^n,Z^n=z^n\}\Pr\{Z^n=z^n,T^n=t^n\}
\end{aligned}\tag{16}$$

for any $x^n\in X^n, t^n\in\mathcal{T}^n$ and $z^n\in Z^n$. Suppose that x^n, t^n and z^n are given. Denote

$$\mathcal{I}(z^n)=\{1\leq i\leq n: z_i\neq ?\},$$

$$\mathcal{I}(t^n)=\{1\leq i\leq n: t_i\in\mathcal{T}_1\}.$$

If $\mathcal{I}(z^n)\neq\mathcal{I}(t^n)$, both sides of (16) equal 0. Formula (16) is established. If $\mathcal{I}(z^n)=\mathcal{I}(t^n)=\mathcal{I}$, terms in Formula (16) are deduced as follows. Firstly,

$$\begin{aligned}
&Pr\{X^n = x^n, Z^n = z^n, T^n = t^n\} \\
&= Pr\{X^n = x^n\} Pr\{T^n = t^n\} \cdot \\
&Pr\{Z^n = z^n | X^n = x^n, T^n = t^n\} \\
&= Pr\{X^n = x^n\} Pr\{T^n = t^n\} \cdot \\
&Pr\{Y_{\mathcal{G}}^n = z^n | X^n = x^n, T^n = t^n\} \\
&= Pr\{X^n = x^n\} Pr\{T^n = t^n\} \cdot \\
&Pr\{Y_{\mathcal{G}}^n = z^n | X^n = x^n\},
\end{aligned} \tag{17}$$

where the last equality follows because X^n and Y^n are independent from T^n. Moreover,

$$\begin{aligned}
&Pr\{X^n = x^n, Z^n = z^n\} \\
&= Pr\{X^n = x^n, Y_{\mathcal{G}}^n = z^n\} \\
&= Pr\{X^n = x^n\} Pr\{Y_{\mathcal{G}}^n = z^n | X^n = x^n\}.
\end{aligned} \tag{18}$$

Finally,

$$\begin{aligned}
&Pr\{Z^n = z^n T^n = t^n\} \\
&= Pr\{Y_{\mathcal{G}}^n = z^n, T^n = t^n\} \\
&= Pr\{Y_{\mathcal{G}}^n = z^n\} Pr\{T^n = t^n\},
\end{aligned} \tag{19}$$

where the last equality follows because Y^n is independent from T^n. Combining Formulas (17)–(19) results in Formula (16) also holding for x^n,z^n and t^n with $\mathcal{I}(z^n) = \mathcal{I}(t^n)$. The proof is completed.

7. Conclusions

Since the data in WBAN is highly related with the personal health, it is vital to protect this healthy information from attacks. In this paper, from the perspective of information theory, we studied the infrastructure of secure transmission system in WBAN, and solved the capacity problem of a class of finite-state Markov erasure wiretap channel for the IoT. The coding scheme used in this paper comes from the generalized wiretap channel II with the noisy main channel. The idea may be used to solve the capacity problems of other non-DMC wiretap channels. In a theoretical sense, the secure performance of our designed algorithm is not relevant with the computation capability of engaged computers and can guarantee the security of transmitted data in WBAN, by which the personal privacy could be significantly protected.

Author Contributions: Conceptualization, B.W.; Methodology, B.W.; Software, Y.S., W.G. and G.F.; Data Curation, W.G.; Writing—Original Draft Preparation, B.W.; Writing—Review & Editing, Y.S.; Supervision, J.D.; Funding Acquisition, J.D.

Funding: This research was funded by the National Natural Science Foundation of China Nos. 51804304, 61571338 and U1709218, the Natural Science Basic Research Plan of Shaanxi Province No. 2018JM5052, the Key Research and Development Plan of Shaanxi Province No. 2017ZDCXL-GY-05-01, the National Key Research and Development Program of China Nos. 2016YFE0123000, YS2017YFGH000872, and 2018YFC0808301, the Xi'an Key Laboratory of Mobile Edge Computing and Security No. 201805052-ZD3CG36, the China Postdoctoral Science Foundation No. 2015M5826, the Scientific Research Program Funded of Shaanxi Provincial Education Department No. 2016JK1501, and the Shaanxi Provincial Postdoctoral Science Foundation of Shaanxi Provincial.

Acknowledgments: The authors would like to thank Ning Cai of Shanghai Tech University for helping to prove the work in this paper. The authors are grateful to the anonymous reviewers for their constructive comments on the paper.

Conflicts of Interest: The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Tolossa, Y.J.; Vuppala, S.; Kaddoum, G.; Abreu, G. On the uplink secrecy capacity analysis in D2D-enabled cellular network. *IEEE Syst. J.* **2017**, *12*, 2297–2307. [CrossRef]
2. Jameel, F.; Wyne, S.; Kaddoum, G.; Duong, T.Q. A comprehensive survey on cooperative relaying and jamming strategies for physical layer security. *IEEE Commun. Surv. Tutor.* **2018**. [CrossRef]
3. Kong, L.; Vuppala, S.; Kaddoum, G. Secrecy Analysis of Random MIMO Wireless Networks over $\alpha - \mu$ Fading Channels. *IEEE Trans. Veh. Technol.* **2018**. [CrossRef]
4. Zhang, P.N.; Ma, J. Channel Characteristic Aware Privacy Protection Mechanism in WBAN. *Sensors* **2018**, *18*, 2403. [CrossRef] [PubMed]
5. Anwar, M.; Abdyllah, A.H.; Butt, R.A.; Ashraf, M.W.; Qureshi, K.N.; Ullah, F. Securing Data Communication in Wireless Body Area Networks Using Digital Signatures. *Technol. J.* **2018**, *23*, 50–55.
6. Wyner, A.D. The Wire-Tap Channel. *Bell Syst. Technol. J.* **1975**, *54*, 1355–1387. [CrossRef]
7. Kramer, G. Topics in Multi-user Information Theory. *Found. Trends Commun. Inf. Theory* **2007**, *4*, 265–444. [CrossRef]
8. Csiszar, I.; Korner, J. Broadcast channels with confidential messages. *IEEE Trans. Inf. Theory* **1978**, *24*, 339–348. [CrossRef]
9. Chen, Y.; Han Vinck, A.J. Wiretap channel with side infor-mation. *IEEE Trans. Inf. Theory* **2008**, *54*, 395–402. [CrossRef]
10. Dai, B.; Luo, Y. Some new results on the wiretap channel with side information. *Entropy* **2012**, *14*, 1671–1702. [CrossRef]
11. Dai, B.; Han Vinck, A.J.; Hong, J.; Luo, Y.; Zhuang, Z. Degraded Broadcast Channel with Noncausal Side Information, Confidential Messages and Noiseless Feedback. In Proceedings of the 2012 IEEE International Symposium on Information Theory, Cambridge, MA, USA, 1–6 July 2012; pp. 438–442.
12. Dai, B.; Luo, Y.; Han Vinck, A.J. Capacity region of broadcast channels with private message and causual side information. In Proceedings of the 3rd International Conference on Image and Signal Processing (CISP 2010), Yantai, China, 16–18 October 2010; pp. 3770–3773.
13. Khisti, A.; Diggavi, S.N.; Womell, G.W. Secrete-key agreement with channel state information at the transmitter. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 672–681. [CrossRef]
14. Chia, Y.H.; El Gamal, A. Wiretap channel with causal state information. *IEEE Trans. Inf. Theory* **2012**, *58*, 2838–2849. [CrossRef]
15. Ozarow, L.H.; Wyner, A.D. Wire-tap channel II. *AT T Bell Lab. Technol. J.* **1984**, *63*, 2135–2157. [CrossRef]
16. He, D.; Luo, Y. A kind of non-DMC erasure wiretap chan-nel. In Proceedings of the 2012 IEEE 14th International Conference on Communication Technology, Chengdu, China, 9–11 November 2012; pp. 1082–1087.
17. He, D.; Luo, Y.; Cai, N. Strong Secrecy Capacity of the Wiretap Channel II with DMC Main Channel. In Proceedings of the 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–15 July 2016.
18. Gilbert, E.N. Capacity of a burst-noise channel. *Bell Syst. Technol. J.* **1960**, *39*, 1253–1265. [CrossRef]
19. Elliott, E.O. Estimates of error rates for codes on burst-noise channels. *Bell Syst. Technol. J.* **1960**, *42*, 1977–1997. [CrossRef]
20. Wang, H.S.; Moayery, N. Finite-state Markov channel—A useful model for radio communication channels. *IEEE Trans. Veh. Technol.* **1995**, *44*, 163–171. [CrossRef]
21. Lv, N.; Chen, C.; Qiu, T.; Sangaiah, A.K. Deep Learning and Superpixel Feature Extraction based on Sparse Autoencoder for Change Detection in SAR Images. *IEEE Trans. Ind. Inf.* **2018**. [CrossRef]
22. Chen, C.; Hu, J.; Qiu, T.; Atiquzzaman, M.; Ren, Z. CVCG: Cooperative V2V-aided Transmission Scheme Based on Coalitional Game for Popular Content Distribution in Vehicular Ad-hoc Networks. *IEEE Trans. Mob. Comput.* **2018**. [CrossRef]