Wiretap Channel with Shared Key

Wei Kang and Nan Liu
School of Information Science and Engineering
Southeast University, Nanjing, Jiangsu, P. R. China, 210096
Email: {wkang, nanliu}@seu.edu.cn

Abstract—This paper studies the problem of secure communication over a wiretap channel where the transmitter and the legitimate receiver share a secret key, which is concealed from the eavesdropper. We find the secrecy capacity under this scenario. This result generalizes that of Yamamoto, which is applicable only to less noisy wiretap channels, to the general wiretap channel when no distortion is allowed at the legitimate receiver.

I. INTRODUCTION

In his historical paper [1], Shannon studied the problem of secure communication from an information theoretical perspective. His model consists of a transmitter (Alice), a receiver (Bob) and an eavesdropper (Eve) where Bob and Eve share the same observation. Secure communication is achieved via a secret key, which is shared between Alice and Bob and is concealed from Eve.

On the other hand, Wyner [2] considered the problem of secure communication in a degraded wiretap channel where Bob and Eve receive signals from Alice through a degraded broadcast channel, i.e., Eve’s observation is a degraded version of Bob’s observation. In this scenario, there is no secret key shared between Alice and Bob and secure communication is achieved by exploiting the quality difference between the channels of Bob and Eve. Later, Csiszar and Korner [3] extended Wyner’s result to the general (not necessarily degraded) wiretap channel.

Since secure communication can be achieved by either utilizing a secret key [1] or exploiting the channel difference between Bob and Eve [2], [3], there have been efforts of studying the problem of secure communication by considering the presence of a secret key in the problem of the wiretap channel, for example, [4], [5] studied the rate-distortion problem of the wiretap channel with shared key. Related problems include [6]–[8], which investigated the problem of secure communication in the wiretap channel with secure feedback from Bob to Alice (coded or uncoded). In the achievability schemes of [6]–[8], the secure feedback link is used to form a shared key between Alice and Bob, and secure communication between Alice and Bob is then achieved by utilizing the shared key generated from the feedback link and exploiting the channel difference between Bob and Eve. So far, in [4]–[8], tight results have only been obtained in special cases such as degraded or less noisy wiretap channels. For the general wiretap channel, the achievability and converse do not meet and the secrecy capacity has not been established. In this paper, we find the secrecy capacity for the general wiretap channel with shared key. Our result generalizes that of [4] from less noisy broadcast channel to the general broadcast channel when no distortion is allowed at the legitimate receiver.

II. PROBLEM STATEMENT AND RESULT

We consider a broadcast channel with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y} \times \mathcal{Z}$ and transition probability $V(\mathcal{y}, \mathcal{z}|\mathcal{x})$, which is a stochastic matrix. Assume a message $W$ is uniformly distributed in a discrete alphabet $\mathcal{W}$ and a key $K$ is uniformly distributed in a discrete alphabet $\mathcal{K}$. For any $\epsilon, \mu > 0$, a $(n, \epsilon, \mu)$-code includes a stochastic encoder and a decoder, where the stochastic encoder is defined by a conditional probability $Q(x^n|w, k)$ with $x^n \in \mathcal{X}^n$, $w \in \mathcal{W}$, $k \in \mathcal{K}$, the decoder is a mapping $\varphi : \mathcal{Y}^n \times \mathcal{Z}^n \to \mathcal{W}$, the average probability of decoding correctly is larger than $1 - \epsilon$, i.e.,

$$\frac{1}{|\mathcal{W}||\mathcal{K}|} \sum_{w \in \mathcal{W}} \sum_{k \in \mathcal{K}} \sum_{x^n \in \mathcal{X}^n} Q(x^n|w, k) \sum_{z^n \in \mathcal{Z}^n} V^n(A(w, k), z^n|x^n) > 1 - \epsilon$$ (1)

where the decoding set $A(w, k)$ is defined as $\{y^n \in \mathcal{Y}^n : \varphi(y^n, k) = w\}$, and the information leakage at $z^n$ is smaller than $\mu$, i.e.,

$$\frac{1}{n} I(W; Z^n) \leq \mu$$ (2)

if $Z^n$ is the random output sequence generated by the message $W$ through the channel. For a given $R_k > 0$, $R$ is an achievable secrecy rate for the above broadcast channel if for any $\epsilon, \mu, \eta > 0$ and sufficiently large $n$, there exists a $(n, \epsilon, \mu)$-code with $\frac{1}{n} \log |\mathcal{W}| > R - \eta$ and $\frac{1}{n} \log |\mathcal{K}| = R_k$. In other words, for a fixed secret key rate $R_k$, $R$, which is a function of $R_k$, is an achievable secrecy rate if there exist a sequence of codes where both the probability of error at the legitimate receiver and the information leakage at the eavesdropper goes to zero as the code length goes to infinity. The secrecy capacity is defined as the supremum of all achievable secrecy rates.
The main result of this paper is the following theorem.

**Theorem 1** The secrecy capacity of the general wiretap channel \( p(y, z|x) \) with shared key \( R_k \) is

\[
C_s = \max_{U \rightarrow V \rightarrow X \rightarrow Y \rightarrow Z} \min \left( (I(V; Y|U) - I(V; Z|U)) + R_k, I(V; Y) \right) \tag{3}
\]

where \([a]^+\) is defined as the maximum between 0 and \( a \).

**Remark:** If the wiretap channel is degraded, i.e., \( X \rightarrow Y \rightarrow Z \) form a Markov chain, the secrecy capacity found in Theorem 1 is consistent with the existing result of secrecy capacity for the degraded wiretap channel with shared key [4]. This can be seen as follows: we upper bound the secrecy capacity as

\[
C_s = \max_{U \rightarrow V \rightarrow X \rightarrow Y \rightarrow Z} \min(1, [I(V; Y|U) - I(V; Z|U)] + R_k, I(V; Y)) \leq \max_{U \rightarrow V \rightarrow X \rightarrow Y \rightarrow Z} \min (I(V; Y) - I(V; Z) + R_k, I(V; Y)) = \max_{U \rightarrow V \rightarrow X \rightarrow Y \rightarrow Z} \min (I(X; Y) + R_k - I(X; Z), I(X; Y)). \tag{4}
\]

This upper bound is achievable if we let \( U \) be independent of \((V, X, Y, Z)\) and \( V = X \), and therefore is the secrecy capacity for degraded wiretap channel with shared key [4].

To prove the above theorem, we establish equivalence between different rates in the following lemma.

**Lemma 1** The secrecy capacity \( C_s \) satisfies

\[
C_s = C_{s1} = C_{s2} \tag{5}
\]

where

\[
C_{s1} \triangleq \max_{U \rightarrow V \rightarrow X \rightarrow Y \rightarrow Z} \min (I(V; Y|U) - I(V; Z|U) + R_k, I(V; Y)), \tag{6}
\]

\[
C_{s2} \triangleq \max_{U \rightarrow V \rightarrow X \rightarrow Y \rightarrow Z} \min (I(V; Y|U) + I(V; Y|U) + R_k - I(V; Z|U), I(U; Y), I(U; Z)) \tag{7}
\]

The proof of Lemma 1 is in the Appendix. In the sequel, instead of \( C_s \), we prove \( C_{s2} \) is the secrecy capacity.

III. THE CONVERSE

We begin with the technique in [9, Page 314, Eqn (3.34)],

\[
H(Y^n) - H(Z^n) = \sum_{i=1}^{n} H(Y_{1,i}|Y^{i-1}Z_{i+1}^n) - H(Z_i|Y^{i-1}, Z_{i+1}^n), \tag{8}
\]

\[
H(Y^n|W, K) - H(Z^n|W, K) = \sum_{i=1}^{n} H(Y_{1,i}|Y^{i-1}, Z_{i+1}^n, W, K) - H(Z_i|Y^{i-1}, Z_{i+1}^n, W, K). \tag{9}
\]

Define auxiliary random variables

\[
U_i \triangleq (Y^{i-1}, Z_{i+1}^n). \tag{10}
\]

We also define a time sharing random variable \( Q \), which is independent of everything else and is uniform on the set \( \{1, 2, \ldots, n\} \). With the definition of \( U_i \) and \( Q \), we further define the following random variables

\[
U = (U_Q, Q), \quad V = (U, W, K), \quad X = X_Q, \quad Y = Y_Q, \quad Z = Z_Q. \tag{11}
\]

Note that the Markov Chain \( U \rightarrow V \rightarrow X \rightarrow (Y, Z) \) is satisfied. Using (8), (9) and the definition of the auxiliary random variables, there exist real numbers \( t_1 \) and \( t_2 \) such that [9, Page 315, Eqn (3.44)]

\[
\frac{1}{n} H(Y^n) = H(Y|U) + t_1, \quad \frac{1}{n} H(Z^n) = H(Z|U) + t_1, \tag{12}
\]

\[
\frac{1}{n} H(Y^n|W, K) = H(Y|V) + t_2, \quad \frac{1}{n} H(Z^n|W, K) = H(Z|V) + t_2, \tag{13}
\]

where

\[
0 \leq t_1 \leq \min(I(U; Y), I(U; Z)), \quad 0 \leq t_2 \leq \min(I(V; Y), I(V; Z)). \tag{14}
\]

Since the code satisfies the information leakage constraint of (2), we have

\[
n \mu \geq I(W; Z^n) = I(W, K; Z^n) - I(K; Z^n|W) = H(Z^n) - H(Z^n|W, K) - H(K|Z^n, W) \geq H(Z^n) - H(Z^n|W, K) - H(K) = H(Z^n) - H(Z^n|W, K) - nR_k = n(H(Z|U) + t_1 - H(Z|V) - t_2 - R_k) \tag{15}
\]

which implies

\[
t_1 - t_2 \leq R_k - I(Z; V|U) + \mu. \tag{16}
\]

We also have

\[
t_1 - t_2 \leq \min(I(U; Y), I(U; Z)). \tag{17}
\]

Thus, we have

\[
t_1 - t_2 \leq \min(R_k - I(Z; V|U) + \mu, I(U; Y), I(U; Z)). \tag{18}
\]

From Fano’s inequality, we have

\[
|W| = H(W) = H(W|K) \leq I(W; Y^n|K) + n\epsilon_n \leq I(W, K; Y^n) + n\epsilon_n = H(Y^n) - H(Y^n|W, K) + n\epsilon_n = n(H(Y|U) + t_1 - H(Y|V) - t_2 + \epsilon_n) = n(I(Y; V|U) + (t_1 - t_2) + \epsilon_n) \leq n(I(Y; V|U) + \min(R_k - I(Z; V|U) + \mu, I(U; Y), I(U; Z)) + \epsilon_n) \tag{19}
\]
which concludes the proof of the converse.

IV. THE ACHIEVABILITY

Consider a given distribution
\[ p(u, v, x, y, z) = p(u, v)p(x|v)W(y, z|x). \] (23)

**Case 1:** If \( R_k \leq I(V; Z|U) \), i.e.,
\[ R_k - I(V; Z|U) \leq 0 \leq \min(I(U; Y), I(U; Z)), \] (24)
we have
\[ C_{x2} = I(V; Y|U) - I(V; Z|U) + R_k. \] (25)

We will show that \( I(V; Y|U) - I(V; Z|U) + R_k \) is achievable as follows.

**Codebook generation:** Randomly generate a \( u^n \) sequence according to \( p(u) \). Then generate a three-dimensional codebook of size \((L, J, M)\). More specifically, for every \( 0 \leq l \leq L - 1, 0 \leq j \leq J - 1, 0 \leq m \leq M - 1 \), independently generate a sequence \( v^n_{ljm} \) conditioned on the \( u^n \) sequence according to \( p(v|u) \), where
\[
\frac{1}{n} \log L \rightarrow I(V; Y|U) - I(V; Z|U) + R_k, \\
\frac{1}{n} \log J \rightarrow R_k, \\
\frac{1}{n} \log M \rightarrow I(V; Z|U) - R_k. \] (26-28)

**Encoding:** If message \( W = l \) and shared key \( K = j \), uniformly select a codeword \( v^n_{lj0} \) to \( v^n_{lj(M-1)} \) and randomly generate a codeword \( x^n \) conditioned on \( v^n_{ljm} \) according to \( p(x|v) \). Transmit this codeword \( x^n \) into channel.

**Decoding:** Upon receiving \( y^n \) and knowing that the shared key \( K \) is equal to \( j \), the decoder determines a unique codeword that is jointly typical with \( y^n \) from all possible \( v^n_{ljm} \) sequences, i.e., \( L \times M \) possible codewords, where \( 0 \leq l \leq L - 1, 0 \leq m \leq M - 1 \). Since \( \frac{1}{n} \log L \times M \rightarrow I(V; Y|U) \), the decoder can decode with probability of error going to zero.

**Eqivocation:** We follow the steps in [3] in the calculation of the equivocation rate as follows.
\[
H(W|Z^n) = H(W, Z^n) - H(Z^n) = H(W, Z^n, V^n) - H(V^n|W, Z^n) - H(Z^n) \\
\geq H(V^n) + H(Z^n|V^n) - H(V^n|W, Z^n) - H(Z^n). \] (29)
The first term is equal to \( \log L \times J \times M \), i.e.,
\[
\frac{1}{n} \log H(V^n) \rightarrow I(V; Y|U) + R_k. \] (30)

For the second term, we have
\[
\frac{1}{n} \log H(Z^n|V^n) \rightarrow H(Z|V). \] (31)

For the third term, we have
\[
\frac{1}{n} \log H(V^n|W, Z^n) \rightarrow 0. \] (32)

For the last term, we have
\[
\frac{1}{n} \log H(Z^n) \leq H(Z|U). \] (33)

Thus, we have
\[
\frac{1}{n} \log H(W|Z^n) \geq I(V; Y|U) - I(V; Z|U) + R_k. \] (34)

**Case 2:** If \( I(V; Z|U) \leq R_k \leq I(V; Z|U) + \min(I(U; Y), I(U; Z)) \), i.e.,
\[ 0 \leq R_k - I(V; Z|U) \leq \min(I(U; Y), I(U; Z)), \] (35)
we have
\[ C_{x2} = I(V; Y|U) - I(V; Z|U) + R_k. \] (36)

We will show that \( I(V; Y|U) - I(V; Z|U) + R_k \) is achievable as follows.

**Codebook generation:** Randomly generate \( M \) many \( u^n \) sequences, i.e., \( u^n_0, \ldots, u^n_{M-1} \). Then generate a two-dimension codebook of size \((L, J)\). More specifically, for every \( 0 \leq l \leq L - 1, 0 \leq j \leq J - 1 \), randomly select a \( u^n \) sequences from \( u^n_0, \ldots, u^n_{M-1} \), say \( u^n_m \), with uniform distribution, and independently generate a \( v^n \) sequence, called \( v^n_{lj} \), conditioning on \( u^n_m \) according to \( p(v|u) \), where
\[
\frac{1}{n} \log M \rightarrow R_k - I(V; Z|U), \\
\frac{1}{n} \log L \rightarrow I(V; Y|U) - I(V; Z|U) + R_k, \\
\frac{1}{n} \log J \rightarrow R_k. \] (37-39)

**Encoding:** If message \( W = l \) and shared key \( K = j \), select codeword \( v^n_{lj} \) and randomly generate a codeword \( x^n \) conditioned on \( v^n_{lj} \) according to \( p(x|v) \). Transmit this codeword \( x^n \) into the channel.

**Decoding:** Upon receiving \( y^n \), the decoder first determines a unique \( u^n \) sequence, which is joint typical with \( y^n \). Since \( M \) satisfies
\[
\frac{1}{n} \log M \rightarrow R_k - I(V; Z|U) \leq \min(I(U; Y), I(U; Z)), \] (40)
the decoder can determine the correct \( u^n \) sequence with probability of error going to zero. Then, the decoder determines a unique \( v^n \) sequence, which is joint typical with \( (u^n, y^n) \). There are around \( L/M \) many \( v^n \) sequences joint typical with a certain \( u^n \) sequences given \( K = j \). Since
\[
\frac{1}{n} \log \frac{L}{M} \rightarrow I(V; Y|U) - I(V; Z|U) + R_k - (R_k - I(V; Z|U)) = I(V; Y|U), \] (41)
the decoder can decode with probability of error going to zero.

**Eqivocation:** We follow the steps in [3] in the calculation of the equivocation rate. Let \( S \) be the random variable which
selects the $u^n$ sequence through the selection of the $v^n$ sequence in the encoding process. Then we have

$$H(W|Z^n) \geq H(W|Z^n, S)$$

$$= H(W, Z^n|S) - H(Z^n|S)$$

$$= H(W, Z^n, V^n|S) - H(V^n|W, Z^n, S) - H(Z^n|S)$$

$$= H(W, V^n|S) + H(Z^n|V^n, W, S) - H(V^n|W, Z^n, S) - H(Z^n|S)$$

$$\geq H(V^n|S) + H(Z^n|V^n)$$

$$- H(V^n|W, Z^n, S) - H(Z^n|S). \quad (42)$$

The first term is equal to $\log \frac{L \times J}{M}$, i.e.,

$$\frac{1}{n} \log H(V^n|S) \rightarrow I(V; Y|U) + R_k. \quad (43)$$

For the second term, we have

$$\frac{1}{n} \log H(Z^n|V^n) \rightarrow H(Z|V). \quad (44)$$

For the third term, we have

$$\frac{1}{n} \log H(V^n|W, Z^n, S) \rightarrow 0. \quad (45)$$

For the last term, we have

$$\frac{1}{n} \log H(Z^n|S) \leq H(Z|U). \quad (46)$$

Thus, we have

$$\frac{1}{n} \log H(W|Z^n) \geq I(V; Y|U) - I(V; Z|U) + R_k. \quad (47)$$

**Case 3:** If $R_k \geq I(V; Z|U) + \min(I(U; Y), I(U; Z))$, then

$$I(V; Y|U) + \min(R_k - I(V; Z|U), I(Y; U))$$

$$= I(V; Y|U) + \min(R_k - I(V; Z|U), I(U; Y)). \quad (48)$$

The achievability scheme is same as the one in case 2 when $R_k = \min(I(U; Y), I(U; Z))$.

**V. Conclusion**

In this paper, we studied the problem of secure communication over a wiretap channel where the transmitter and the legitimate receiver share a secret key, which is concealed from the eavesdropper. We found the secrecy capacity under this scenario. This result generalizes that of Yamamoto, which is applicable only to less noisy wiretap channels, to the general wiretap channel when no distortion is allowed at the legitimate receiver.

**Appendix**

It is straightforward to see that

$$C_s \geq C_{s1} \geq C_{s2}. \quad (49)$$

We first show that $C_s \leq C_{s1}$. For a given distribution $p(u, v, x, y, z)$, if $I(V; Y|U) \geq I(V; Z|U)$, we have

$$\min([I(V; Y|U) - I(V; Z|U)]^+ + R_k, I(V; Y)) = \min(I(V; Y|U) - I(V; Z|U) + R_k, I(V; Y)). \quad (50)$$

Otherwise, we define a new set of distribution $p(u', v', x', y', z')$, where $p(v', x', y', z') = p(v, x, y, z)$ and $U' = V'$. In this case,

$$\min([I(V; Y|U) - I(V; Z|U)]^+ + R_k, I(V; Y))$$

$$= \min(I(V'; Y'|U') - I(V'; Z'|U') + R_k, I(V'; Y')). \quad (51)$$

Thus, for every distribution, the seemingly larger rate $\min([I(V; Y|U) - I(V; Z|U)]^+ + R_k, I(V; Y))$ can be achieved by $\min(I(V; Y|U) - I(V; Z|U) + R_k, I(V; Y))$ for some distribution. We conclude that $C_s \leq C_{s1}$, and therefore, $C_s = C_{s1}$.

We then show that $C_{s1} \leq C_{s2}$. We rewrite $C_{s1}$ as

$$C_{s1} \triangleq \max_{u \rightarrow v \rightarrow X \rightarrow (Y, Z)} \min(I(V; Y|U) - I(V; Z|U)) \quad (52)$$

$$\min(I(V; Y|U) - I(V; Z|U) + R_k, I(V; Y))$$

$$= \max_{u \rightarrow v \rightarrow X \rightarrow (Y, Z)} I(V; Y|U) +$$

$$+ \min(R_k - I(V; Z|U), I(U; Y)). \quad (53)$$

For a given distribution $p(u, v, x, y, z)$, if $I(V; U) \leq I(V; Z)$, we have

$$I(V; Y|U) + \min(R_k - I(V; Z|U), I(U; Y))$$

$$= I(V; Y|U) + \min(R_k - I(V; Z|U), I(U; Y), I(U; Z)). \quad (54)$$

Otherwise, we have

$$\min(I(V; Y|U) - I(V; Z|U))$$

$$\leq \min(I(V; Y|U) - I(V; Z|U) +$$

$$+ I(U; Y) - I(U; Z) + R_k, I(V; Y))$$

$$= \min(I(V; Y) - I(V; Z) + R_k, I(V; Y)). \quad (55)$$

We define a new set of distribution $p(u', v', x', y', z')$ such that $p(v', x', y', z') = p(v, x, y, z)$ and $U'$ is independent of $(V', X', Y', Z')$. Then, we have

$$\min(I(V; Y) - I(V; Z) + R_k, I(V; Y))$$

$$= \min(I(V'; Y'|U') - I(V'; Z'|U') + R_k, I(V'; Y'|U'))$$

$$= I(V'; Y'|U') +$$

$$+ \min(R_k - I(V'; Z'|U'), I(U'; Y'), I(U'; Z')). \quad (56)$$

where $I(U'; Y') = I(U'; Z') = 0$. Thus, for every distribution, the seemingly larger rate $\min(I(V; Y|U) - I(V; Z|U) + R_k, I(V; Y))$ can be achieved by $I(V; Y|U) + \min(R_k - I(V; Z|U), I(U; Y), I(U; Z))$ for some distribution. We conclude that $C_{s1} \leq C_{s2}$, and therefore, $C_{s1} = C_{s2}$.

**Acknowledgments**

This work was supported by the National Natural Science Foundation of China under Grants 60902008 and 60972027, the Research Fund of National Mobile Communications Research Laboratory, Southeast University (No. 2009B04 and 2008A02), International Science and Technology Cooperation Program under Grant 2008DFA12090, and Sino-Europe cooperation project S2010GR0902.
REFERENCES


