Proof of theorems on multisource information theory and implementation in wireless relay network proceedings

Geng Zhang and Jie Shen†
China Electrical Power Research Institute,
E-mail: 172677249@qq.com, shenjie75@sohu.com

Ou Wang
Beijing University of Posts & Telecom
E-mail: ouwang@139.com

Lei lv
State Grid Meishan Electric Power Supply Company
13778888186@qq.com

In this paper, three theorems on multisource information theory are introduced, then analysis and proof of them are given. Source coding with SI and joint source coding problem can be used in wireless MAC channel, while source coding with one or two helpers can be used in wireless BC channel. Finally numerical results in a wireless relay network and discussion of them are given.

Keywords: Network coding; Broadcast channel; Relay channel

1. Introduction

Network information theory as a branch of information theory in the 1970’s had a vigorous development. Its name network information theory, since it considers the situation that multiple users transmit or receive signals at the same time, they form interrelated network system. Typical examples are: a source for multiple sinks - broadcasts; multiple sources for one sink-multiple accesses; a transport service between the source and the sink - relay channels, etc. [2].

For the above three channels, only the optimal solution of the multiple access channel has been found so far in wireless mobile communication technology. But with the formation of dual issues of the broadcast channel, after more than thirty years of efforts, has not been fully resolved by researchers[3,4]. A special case of Gaussian white noise multi-antenna in wireless channel is studied in [5]. With the application of relay network in electrical power Unified Communication, such that the flat, self-organization, ubiquitous trends, the network information theory in the wireless channel was much more studied in-
depth[6]. For power line communication system, organized as a mesh network, while meeting the trouble shouting problems, broadcast was adopted to inform the dispatching center.

The rest of the paper is organized as follows. In section 2, we discuss the multisource encoding problem with side information. In section 3, we focus on three-source multiple access channel (MAC) problem. In section 4, we analyse source encoding with two helpers problem. In section 5, we give an implementation of these problems in a wireless relay network, and discuss the numerical results. In section 6, we give the conclusion.

2. The Model of Multisource Network Information

Network information theory in the multi-user wireless communication problems, can be seen as a secondary information sharing. Here we discuss the source encoding problem with side information, which can be described in Figure 1:

![Source coding of side information](image)

The problem can be described as:

- There is a pair of joint distributed discrete memoryless sources, $X$ and $Y$. To be precise, for each $i$, $(X_i, Y_i) \sim P_{XY}$, and $(X_i, Y_i) \perp (X_j, Y_j)$, if $i \neq j$. For convenience of the following description, the two are connected by the channels shown in the figure with $W = P_{1X}$.
- $X$ is reconstructed by a single decoder. It is assumed here that the reconstructed signal is not lost, that is, when the encoding is long enough, then exist $\text{Pr}(\hat{X}^n \neq X^n) \to 0$.
- The decoder detects two messages: the sequence from the source $Y$ and the encoded bits sent by the encoder.
- Encoder used in block diagram? said that this is the need to design. The sequence $X^n$ is encoded with $R_X$ bits per symbol, so that it is possible to reconstruct the original information without loss in the
decoder by means of side information $Y^n$. The question is: What is the minimum $R_X$?

Mathematical description, for a block code $(n, R)$ there exists a pair of encoding/decoding maps, $f_n, g_n$, where

$$f_n: \mathcal{X}^n \rightarrow \{0,1\}^{nR}$$

$$g_n: \{0,1\}^{nR} \times 2^n \rightarrow \mathcal{X}^n$$

If the rate of the coded sequence $(n, R)$ can reach the rate $R$, the error probability is met:

$$P_e^{(n)} \triangleq \text{Pr}\left[X^n \neq g_n(f_n(X^n), Y^n)\right] \rightarrow 0$$

When $n \rightarrow \infty$, the probability taken as $(P_{XY})^n$.

For the problem of Figure 1, the minimum rate satisfies $R_{\text{min}} \leq H(X)$ and $R_{\text{max}} \geq H(X | Y)$. Theorem 1 shows that the minimum rate is $R_{\text{min}} \geq H(X | Y)$.

**Theorem 1** the minimum achievable rate is $R_{\text{min}} = H(X | Y)$. The proof omitted here.

**Proof:** First, for a typical sequence $X^n$, assume that the encoder $f_n$ will map $X^n$ to bits $nR$, and $R < H(X)$, let

$$\mathcal{A}_i = f(i) \cap \tilde{T}_{p_X}, \quad i \in \{1,2,\ldots,2^{nR}\} \quad (1)$$

It is a typical sequence that $X$ is encoded into a set $i$. As the sequence $X^n$ in the set $\tilde{T}_{p_X}$ exponentially increases, $\mathcal{A}_i$ also contains exponential multiple sequences. So that only the determination of indicator $i$, then to determine the sequence $X$. As the fuzzy set $\mathcal{A}_i$, coding design is the fuzzy set design.

We want to observe the sequence $Y^n$, the decoder can be obtained with a high probability $X$. Which $X^n$ and $Y^n$ have a strong correlation, note the conditional distribution $V = P_{XY}$, and $X^n \in \tilde{T}_v(Y^n)$.

This set is an order of magnitude smaller than the set $\tilde{T}_{p_X}$

$$\tilde{T}_v(Y^n) \approx e^{nH(X|Y)}$$

Where $Y^n$ subjects to distribution $P_Y$. if $\mathcal{A}_i$ is small enough, such that $\mathcal{A}_i \cap \tilde{T}_v(Y^n)$ has high probability of containing only one sequence. The use of side information is to reduce the fuzzy set $\mathcal{A}_i$ until it contains only one
sequence. This creates a paradox, on the one hand, $\mathcal{A}_j$ to be small, and on the other hand $\mathcal{A}_j$ to cover as little as possible all the typical sequences $X$.

Using the most extreme approach, random select $\mathcal{A}_i$. For a fixed-rate $R$ sequence $x^n \in \mathcal{T}_n$, independent uniform selection $i \in \{1, 2, \ldots, 2^{nR}\}$, $x^n$ will be included in the collection $\mathcal{A}_i$, the equivalent of a random item into in one of the $2^{nR}$ drawers. After all the sequences are randomly assigned, the set of decoder sequences is fixed and invisible.

- Encoder: for the source $X^n = x^n$, it will search a $i$, such that $i \in \mathcal{A}_i$.
  The index $i$ is passed in the $nR$ bits.
- Decoder: when get $i$ and side information $Y^n = y^n$, decoder search the set: $\mathcal{A}_i \cap \mathcal{T}_n(y^n)$. If there is only one $x^n$, the source sequence is considered $x^n$; otherwise, the set is empty or contains more than one element, resulting in an error.

The bit error rate can be analyzed by a method similar to the Cover’s channel coding theorem. The probability $x^n$ in the set $\mathcal{T}_n(y^n)$ and $\mathcal{A}_i$ is very high, in two sets intersection probability is also high. If the collection $\mathcal{A}_i \cap \mathcal{T}_n(y^n)$ also contains other source sequences, an error occurs. This corresponds to a sequence in $\mathcal{T}_n(y^n)$ which is allocated to $\mathcal{A}_i$, according to the union bound, this probability does not exceed:

$$\left| \mathcal{T}_n(y^n) \right| \frac{1}{2^{nR}} \geq 2^{n(H(X|Y) - R)} \quad (2)$$

As long as $R > H(X|Y)$, with the $n$ increase, the error rate approaches 0, that is, the rate $R$ can be reached. ■

3. MAC channel with side information

Theorem 1 can be extended to more than two sources. For example, there are three sources of joint distribution $X, Y, Z$, using the rate of source as $R_Y = H(Y|Z)$ presents $Y$, and the code with rate $R_X = H(X|Y,Z)$ represents $X$. That is, the rate triple $(R_X, R_Y, R_Z) = (H(X|Y,Z), H(Y|Z), H(Z))$ can completely represent the discrete memoryless channel (DMC) coding.

The former Splenian problem is extended to three-source multiple access channel (MAC) problem, as shown in figure 2.
At time $i$, there are two input $X_i \in X$, $Y_i \in Y$, and an output $Z_i \in Z$. The probability of distribution is

$$
\Pr(Z^n = z^n \mid X^n = x^n, Y^n = y^n) = \prod_{i=1}^{n} W(z_i \mid x_i, y_i)
$$

Where $W : X \times Y \rightarrow Z$ is transmit probability matrix. Let two users have two sets of input information bits to be transmitted: $M_X \in \{1, 2, \ldots, 2^n \}$ and $M_Y \in \{1, 2, \ldots, 2^n \}$. They are mapped to:

$$
f_X : \{1, 2, \ldots, 2^n \} \rightarrow X^n
g_Y : \{1, 2, \ldots, 2^n \} \rightarrow Y^n
$$

The mapping of the decoder is:

$$
\phi : Z^n \rightarrow \{1, 2, \ldots, 2^n \} \times \{1, 2, \ldots, 2^n \}
$$

The decoded result is $(\hat{M}_X, \hat{M}_Y) = \phi(Z^n)$. If a message is decoded incorrectly, an error occurs. If the rate of the message pair is reachable, the error probability approaches zero if and only if $n$ approaches infinity. The set of these rate pairs is called the capacity domain of the MAC channel. The capacity domain is convex pentagon, and the capacity pair satisfies inequality of theorem 2.

**Theorem 2:** The capacity domain of the three sources satisfies the following inequalities [7]

$$
(R_X, R_Y) : \begin{align*}
R_X &\leq I(X; Z \mid Y) \\
R_Y &\leq I(Y; Z \mid X) \\
R_X + R_Y &\leq I(X, Y; Z)
\end{align*}
\tag{3}
$$

**Proof:** Let the user $X$ select the independent and identically distributed random matrix $P_X$ for the code book, the user $Y$'s codebook matrix distribution...
is $P_Y$. When $Z^n$ received, the decoder looks for code word pairs from the codebook, the typical joint distribution of which is denoted by $P_{XYZ}$. The following four situations have an error:

1. Sending codeword pairs is not a typical distribution;
2. The two codewords are wrong, the received codewords are typical distribution;
3. The user $X$ code word is correct but the user $Y$ code word is wrong, and forms a typical distribution with the received code word;
4. Contrary to case 3.

Similar to the two source conditions, the source code is $g_x: \mathcal{X}^n \rightarrow \{0,1\}^{nH(X|Y,Z)}$
$g_y: \mathcal{Y}^n \rightarrow \{0,1\}^{nH(Y|Z)}$

According to the side information $Z^n$, the above equation can be compressed or reconstructed $X^n, Y^n$ with high probability. The $i, j$ can be selected from the codebook of $X$ and $Y$

$g^{-1}_x(i) \in \mathcal{X}^n$ and $g^{-1}_y(j) \in \mathcal{Y}^n$

The error probability of the MAC channel depends on the error rate of the joint distribution of the source code word, and the probability of both errors is low. Obviously, the $i$ and $j$ of the selected $|g^{-1}_x(i)|$ and $|g^{-1}_y(j)|$ should be as large as possible. The following two formulas show the potential of these two sets:

$$\max_i |g^{-1}_x(i)| \geq 2^{nH(X)} / 2^{nH(X|Y,Z)} = 2^{nI(X;Z|Y)}$$
$$\max_j |g^{-1}_y(j)| \geq 2^{nH(Y)} / 2^{nH(Y|Z)} = 2^{nI(Y;Z)} \quad \text{(4)}$$

Using the similarity method of theorem 1 to prove the two sources, we can get the conclusions, here omitted.

4. Source encoding with two helpers

The following analysis has two helpers of the situation, how to effectively use side information, see Figure 3. Unlike Figure 2, the encoder observes the sequence $Y$ and only needs $X$ to be reconstructed. If a part of the sequence $Y$ does not depend on $X$, then this information need not be considered. The following theorem illustrates the reachability in this case.
Theorem 3: The rate pair \((R_X, R_Y)\) is reachable if and only if there are some conditions that satisfy the Markov condition \(U \rightarrow Y \rightarrow X\),
\[
R_X \geq H(X \mid U) \quad R_Y \geq I(U; Y) \tag{5}
\]

Proof: Obviously, \(R_Y = I(U; Y)\) is the loss of the source codeword, and these losses are discarded irrelevant information. The discarded information \(Y\) can be reconstructed according to the information \(U\), the former is the reconstructed side information \(X^n\), but \(X\) only to achieve rate \(H(X \mid U)\). There is a balance between the two, if \(U\) has larger amount of information, then the \(I(U; Y)\) greater, but the \(H(X \mid U)\) is smaller; and vice versa. ■

Theorems 1 and 2 are general conclusions about the reachability of multiple access channel, and theorem 3 gives the lower bound of the achievable rate of the broadcast channel. These three theorems can be used as a theoretical basis for the analysis of network information theory in the process of networking of electrical power wireless communication systems NB IoT.

5. Theorems and Proofs

A broadcasting channel has one transmitter and multiple receivers, and a discrete memoryless version of it takes one transmitter \(X_i\) and feeds two receivers with \(Y_i\) as intended message \(M_1\), and \(Z_i\) as message \(M_2\) respectively, and an \(M_i\) as a common message that encoded by both receivers, according to conditional distribution \(W_{YZjX}\). In the degraded broadcasting channel, where
\[
W_{YZjX} = W_{YjX} \cdot \Phi_{ZjY}
\]
Hence there is a Markov relation $X \rightarrow Y \rightarrow Z$. Thus user $Z$ observe the noise version of $Y$, it is a common information. To optimal the coding strategy, introduce an auxiliary random variable $U$ such that the Markov relation is satisfied.

\[
R_y = I(X; Y | U) = I(U, X; Y) - I(U; Y) = I(X; Y) - I(U; Y) \tag{6}
\]

The optimal problem to choose $U$, such that $I(U; Z)$ is large, yet is $I(U; Y)$ constrained. So the simulation system can be presented as

![Simulation Architecture](image)

Applying these theories to design joint network and Gelfand-Pinsker coding for 2-receiver Gaussian broadcast channels with receiver side information. The coding method provides a unified coding structure for general side information configurations.

Using the proposed method and joint interference cancelation, we derive a unified inner bound to the capacity region of 2-receiver Gaussian broadcast channels under general side information configuration.

![Optimal Rate with Mutual Information](image)

2-receiver Gaussian broadcast channels under general side information configuration. And according to theory 3, the optimal rate of $Z, Y, X$ are $R_1 \leq C \left( \frac{P_1}{N_1} \right), R_2 \leq C \left( \frac{P_2}{N_2 + P_1} \right)$ and $R_3 \leq C \left( \frac{P - P_3 + N_2}{N_3 - P_1 \cdot N_2} \right)$ respectively, where $P_1 + P_2 \leq P$ [8].
The parameters for the simulation are set as follows: the frame length for the codeword is $5 \times 10^4$, QPSK modulation is used on the fading link which is changed for every codeword, the rate of the outer code is $1/2$, threshold for frame dropping scheme is 0.8. Fig.6 shows the overall BLER performance and the rate on the relay Unified Communication’s link.

The decoding process can only be implemented at codeword level rather than on the bit level. A LLR amplitude threshold is set, and when identifying unreliable bits, the joint entropy is measured to examine the feasibility of further lossless decoding. As the plot indicates, identifying unreliable bits actually increases the average joint entropy.

For one particular channel, the conditional probability density function for the overall received signal is

$$p(Y|Z) = \frac{1}{2\pi\sigma^2} \exp\left[\frac{-1}{2\sigma^2} \|Y - HZ\|^2\right]$$

(7)

And the overall LLR is

$$LLR = \sum_{Y,M_1} \exp\left(-\frac{\|X - Y\|^2}{N_0}\right)$$

$$\sum_{Z,M_2} \exp\left(-\frac{\|Y - Z\|^2}{N_0}\right)$$

(8)

increases the cardinality of a code symbols, since they become ternary rather than binary[10]. We discover that unreliable bits tend to spread evenly across an entire codeword, and hence their location cannot be efficiently encoded using run length coding. We also measure joint entropy at even lower SINR and discover that only when SINR is smaller than $-5$dB will the joint
entropy after the unreliable bit identification be smaller than the original joint entropy. And the error spread along with the decoding Markov chains.

6. Conclusion

Analysis and proof of three theorems on multisource information theory are given. The implementation of them in wireless relay networks has been introduced in wireless power communication. Numerical results show that the original information can be retrieved in accordance with theorems mentioned above.

References