

Article

Not peer-reviewed version

---

# Outage Probability Analysis of Relay Communication Systems for Semantic Transmission

---

[Yangyang Cui](#)<sup>\*</sup> and [Xiaowei Wang](#)

Posted Date: 12 March 2025

doi: 10.20944/preprints202503.0871.v1

Keywords: Semantic Communication; Relay Node; Outage Probability; Channel Capacity



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

# Outage Probability Analysis of Relay Communication Systems for Semantic Transmission

Cui Yangyang \*  and Wang Xiaowei

School of Information Engineering, Shanghai Maritime University, 201306, Shanghai, China

\* Correspondence: 202230310028@stu.shmtu.edu.cn

**Abstract:** This paper focuses on relay semantic communication systems and deeply investigates their outage probability performance under different conditions. First, a three-point, two-link communication transmission model is constructed, which includes the source node, relay node, and destination node. The relay operates in half-duplex mode and is based on the DeepSC communication transmission model, with semantic and channel encoding/decoding implemented at both the source and relay nodes. Next, depending on the computational capabilities of the destination node, background knowledge, and channel conditions, the channel transmission strategies are classified into four scenarios: 1) Both the SR and RD links use semantic transmission, with the source and destination nodes having either the same or different background knowledge; 2) Both the SR and RD links use semantic transmission, but the signal is impaired, preventing the destination node from decoding; 3) The SR link uses semantic transmission, while the RD link performs bit transmission. For each scenario, the channel capacity formula and outage probability expressions are derived. In calculating the outage probability, the results are obtained by analyzing and deriving the distributions and conditions of relevant variables. Finally, Monte Carlo simulations are used to validate the theoretical analysis. The research results provide a theoretical basis for comprehensively understanding the performance of relay semantic communication systems and are of significant importance for advancing semantic communication technology.

**Keywords:** semantic communication; relay node; outage probability; channel capacity

## 1. Introduction

Traditional communication systems mainly focus on how to efficiently transmit symbol streams, while semantic communication emphasizes how to encode information so that the receiver can accurately understand the transmitted content. To achieve this effectively, researchers have utilized technologies like deep learning to design semantic encoders and decoders. For instance, neural networks are used to replace traditional modems, optimizing the expression of information during transmission. Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and Transformer models have all been applied to semantic communication, enhancing the system's intelligence. [1] proposed a deep learning-based semantic communication system called DeepSC, which is used for text transmission and outperforms traditional communication systems under low signal-to-noise ratio conditions.

In semantic communication, resource allocation (such as bandwidth, power, time, etc.) has become a critical issue. How to allocate limited communication resources based on the semantic importance of the information, avoid unnecessary data transmission, and ensure the integrity and accuracy of the information is one of the key directions of current research. [2] focuses on the resource allocation problem in text-based semantic communication, where the concept of Semantic Spectral Efficiency (S-SE) is introduced for the first time. By optimizing channel allocation and the number of semantic symbols transmitted, a resource allocation model is proposed to improve the efficiency of semantic communication.

[3] proposed a lightweight, distributed semantic communication system based on deep learning, called L-DeepSC, for low-complexity text transmission. In this system, data transmission from IoT devices to the cloud/edge is performed at the semantic level to improve transmission efficiency.

Semantic communication, as a new paradigm for next-generation networks, significantly reduces bandwidth requirements and improves communication efficiency by extracting the core semantic information from data and transmitting only the relevant content, even under low signal-to-noise ratio conditions. However, relying solely on semantic communication may face challenges in scenarios with high real-time requirements or strict data integrity demands. Therefore, combining the stability of traditional communication with the efficiency of semantic communication to form a hybrid communication architecture can not only meet the needs of different application scenarios but also provide new ideas and directions for the development of next-generation communication networks. [4] studied a heterogeneous communication framework that supports the coexistence of semantic communication and traditional bit-based communication. By proposing a Semi-NOMA (Semi-Non-Orthogonal Multiple Access) scheme, it enables flexible and efficient transmission of both semantic and bit streams. When compared with Orthogonal Multiple Access (OMA) and Non-Orthogonal Multiple Access (NOMA), it was shown that Semi-NOMA offers advantages in terms of semantic and bit rate, as well as power utilization efficiency.

Introducing relays into semantic communication systems is a highly promising innovation. The addition of relay nodes can effectively extend the communication coverage, compensating for the limitations of traditional direct communication in long-distance transmission or signal-blocked scenarios. A relay can receive semantic information from the source node and, leveraging its powerful signal processing and forwarding capabilities, analyze, optimize, and regenerate the semantic content. This ensures that the integrity and accuracy of the semantic information are maintained or even improved during transmission to the destination node. For example, in complex urban environments where tall buildings may cause signal blockage and attenuation, a relay can receive, process, and retransmit the semantic information at appropriate locations, overcoming obstacles and ensuring stable communication. Additionally, the relay can intelligently adjust its forwarding strategy based on network conditions and the destination node's requirements, compressing or expanding the semantic information as needed, thereby improving overall communication efficiency and quality. This lays a solid foundation for the widespread application of semantic communication in more complex scenarios and significantly drives the advancement of semantic communication technology towards greater maturity and reliability.

[5] proposed a semantic relay (SemRelay) framework to address the issues of limited resources in mobile devices and insufficient research in collaborative communication within semantic communication systems. By integrating deep learning techniques on edge devices, it achieves efficient text semantic communication and significantly improves spectrum efficiency and semantic transmission performance in multi-user scenarios through joint optimization of relay power and bandwidth allocation.

[6] proposed a SemRelay-assisted base station system for transmitting text to mobile devices to solve the problem of resource-constrained mobile devices being unable to deploy deep learning-based semantic encoders and decoders. By jointly optimizing relay positions and bandwidth allocation to maximize effective bit rate, the proposed algorithm, which uses a penalty-based approach, yields numerical results showing near-optimal performance, with SemRelay outperforming traditional decode-and-forward relays in terms of rate performance.

Building on this, [7] addressed the problem of deploying deep learning semantic encoders and decoders for multiple mobile users. [8] developed a novel intelligent relay-assisted semantic communication system, combining traditional deep learning methods with intelligent semantic relays. By restoring the meaning of sentences, the system minimizes semantic errors and addresses issues caused by channel variations, applicable in scenarios of deteriorating wireless channels or mismatched knowledge backgrounds between the transmitter and receiver.

[9] proposed a wireless relay channel semantic communication scheme based on autoencoders (AESC), which encodes and decodes sentences from the semantic dimension. The autoencoder module enhances the system's robustness against noise. Additionally, a novel semantic forwarding (SF) mode was designed for relay nodes to forward semantic information at the semantic level, especially in cases where the source and destination nodes do not share common knowledge.

[10] studied semantic communication in multi-hop relay networks to achieve reliable information transmission over long distances or under high path-loss attenuation conditions.

Semantic communication, by focusing on transmitting the core meaning of the source, has the potential to significantly reduce network traffic, effectively alleviating the problem of spectrum resource scarcity. Specifically, for different types of transmission sources such as text [1,11], images [12,13], speech [14], and video [15], researchers have proposed various semantic communication systems that significantly enhance the reliability of semantic transmission.

Full-duplex communication technology, which allows communication nodes to transmit and receive simultaneously, has been regarded as highly promising when combined with semantic communication. In this mode, researchers aim to optimize resource allocation, interference management, and other issues in full-duplex communication using deep learning techniques to improve the efficiency of semantic communication. [16] proposed a novel in-band full-duplex (IBFD) paradigm called Semantic Division Duplex (SDD), using semantic communication to address the reliable source reconstruction problem in traditional IBFD systems.

Semantic communication has shown immense application potential in fields such as intelligent transportation, the Internet of Things (IoT), and smart healthcare. For example, in autonomous driving, communication between vehicles no longer involves transmitting all sensor data but instead transmits semantic information relevant to current driving decisions, thereby improving system efficiency and reducing network load. [17] proposed a unified multi-user semantic communication system for multimodal data transmission, multi-user collaboration, and multi-task execution in edge-intelligent autonomous driving systems to address communication challenges in autonomous driving.

In this paper, we focus on relay semantic communication systems and deeply investigate their performance under different conditions, particularly the outage probability, which is a critical metric. The outage probability directly reflects the likelihood that the communication system fails to meet the predetermined communication quality requirements during transmission, making it essential for evaluating the system's reliability and stability.

We consider various scenarios, including those where semantic transmission occurs between the source node (S) and the relay node (R), as well as between the relay node and the destination node (D), with the destination node having either the same or different background knowledge (BK) as the source node. We also explore scenarios where the relay performs a function similar to traditional amplify-and-forward (AF) due to signal loss caused by long-distance transmission, and hybrid transmission scenarios where the source node and the relay node use semantic transmission, while the relay node and the destination node perform bit transmission. Through modeling and analyzing these different scenarios, we derive the corresponding communication capacity and outage probability expressions, providing a theoretical foundation for comprehensively understanding the performance of relay semantic communication systems.

The paper is organized as follows: Section 2 establishes the system model. Section 3 deals with the outage probability performance of the proposed relay selection strategy, providing both exact and high signal-to-noise ratio approximations. Section 4 presents the simulation and analysis results, while Section 5 summarizes the paper and provides our conclusions.

Notation:  $\mathbb{C}^{m \times n}$  and  $\mathbb{R}^{m \times n}$  represent sets of complex matrices and real matrices of size  $m \times n$ , respectively. Boldface variables represent matrices or vectors.  $\mathcal{C}$  denotes that a variable follows a circularly symmetric complex Gaussian distribution with mean  $\mu$  and covariance  $\Sigma$ .

## 2. System Model

We consider a three - point two - line communication transmission model, as shown in Figure 1, which includes the source node ( $S$ ), the relay node ( $R$ ), and the destination node ( $D$ ), with no direct link. The relay operates in half - duplex mode and uses DeepSC as the communication transmission model. Semantic and channel encoders and decoders are deployed at both the source node ( $S$ ) and the relay node ( $R$ ), where the semantics of the text can be effectively extracted through the Transformer model. It is assumed that the DeepSC transceivers are trained at a base station or cloud platform.

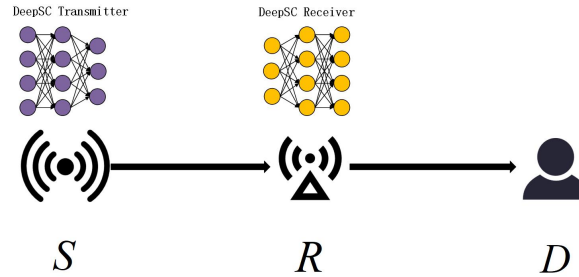


Figure 1. System Model.

In this model, the source node ( $S$ ) inputs a sentence  $s = [w_1, w_2, \dots, w_L]$ , where  $w_L$  represents the  $L$ -th word in the sentence. The DeepSC transmitter consists of two parts: the semantic encoder and the channel encoder, which are responsible for extracting semantic information from  $s$  and ensuring its successful transmission over the physical channel. The sentence is input into the DeepSC transmitter and mapped into a semantic symbol vector  $X_n = [x_1, x_2, \dots, x_{KL}]$ , where  $X_n \in \mathbb{R}^{KL \times 1}$ , and  $KL$  is the length of the semantic symbol vector after transforming the sentence. We note that the length of  $X_n$  changes with  $L$ , allowing more effective extraction of semantic information from sentences of varying lengths. In this model,  $K$  represents the average number of semantic symbols used for each word, and each semantic symbol can be directly transmitted over the communication medium.

The encoded symbol stream can be expressed as:

$$x = C_\alpha(S_\beta(s)) \quad (1)$$

where  $x \in \mathbb{C}^{M \times 1}$ ,  $S_\beta$  is the semantic encoder network with parameter set  $\beta$ , and  $C_\alpha(\cdot)$  is the channel encoder with parameter set  $\alpha$ .

At the relay node ( $R$ ), the received signal is:

$$y_R = h_{SR}x_S + n_{SR} \quad (2)$$

where  $y_R \in \mathbb{C}^{M \times 1}$ ,  $h$  represents the Rayleigh fading channel, which follows a  $\mathcal{CN}(0, 1)$  distribution, and  $n \sim \mathcal{CN}(0, \sigma_n^2)$ . The decoded signal can be expressed as:

$$\hat{s} = S_\chi^{-1}(C_\delta^{-1}(y)) \quad (3)$$

where  $\hat{s}$  is the recovered sentence,  $C_\delta^{-1}(\cdot)$  is the channel decoder with parameter set  $\delta$ , and  $S_\chi^{-1}(\cdot)$  is the semantic decoder network with parameter set  $\chi$ .

At the destination node ( $D$ ), depending on  $D$ 's computational power, either bit transmission or semantic transmission is employed. Specifically, if  $D$  has sufficient computational power to decode the semantic stream, semantic transmission is used over the  $RD$  link. If  $D$  lacks sufficient computational power, traditional bit transmission is used.

Both the  $SR$  link and  $RD$  link employ Rayleigh fading channel transmission models. In a Rayleigh fading channel,  $|h_{SR}|^2$  follows an exponential distribution with parameter  $\frac{1}{\sigma^2}$ , that is:

$$f_{|h|^2}(z) = \frac{1}{\sigma^2} e^{-\frac{z}{\sigma^2}}, z \geq 0. \quad (4)$$

For  $\gamma_{SR}$ , its expression is:

$$\gamma_{SR} = \frac{P_S |h_{SR}|^2}{WN_S}, \quad (5)$$

where  $P_S$  is the transmit power, and  $WN_S$  is the noise power.  $\gamma_{SR}$  is a linear transformation of  $|h_{SR}|^2$ , which is:

$$\gamma_{SR} = \frac{P_S}{WN_S} |h_{SR}|^2. \quad (6)$$

It can be derived that  $\gamma_{SR}$  follows an exponential distribution with parameter  $\lambda_{SR}$ , whose probability density function is:

$$f_{\gamma_{SR}}(y) = \lambda_{SR} e^{-\lambda_{SR} y}, y \geq 0, \quad (7)$$

where:

$$\lambda_{SR} = \frac{WN_S}{P_S \sigma^2}. \quad (8)$$

Similarly, we can derive the following for the  $RD$  link: In a Rayleigh fading channel,  $|h_{RD}|^2$  also follows an exponential distribution with parameter  $\frac{1}{\sigma^2}$ , that is:

$$f_{|h_{RD}|^2}(z) = \frac{1}{\sigma^2} e^{-\frac{z}{\sigma^2}}, z \geq 0. \quad (9)$$

For  $\gamma_{RD}$ , its expression is:

$$\gamma_{RD} = \frac{P_R |h_{RD}|^2}{WN_R}, \quad (10)$$

where  $P_R$  is the transmit power and  $WN_R$  is the noise power.  $\gamma_{RD}$  is a linear transformation of  $|h_{RD}|^2$ , which is:

$$\gamma_{RD} = \frac{P_R}{WN_R} |h_{RD}|^2. \quad (11)$$

It can be derived that  $\gamma_{RD}$  follows an exponential distribution with parameter  $\lambda_{RD}$ , whose probability density function is:

$$f_{\gamma_{RD}}(y) = \lambda_{RD} e^{-\lambda_{RD} y}, y \geq 0, \quad (12)$$

where:

$$\lambda_{RD} = \frac{WN_R}{P_R \sigma^2}. \quad (13)$$

Based on the computational power of the destination node ( $D$ ) and whether the channel conditions meet the requirements for semantic transmission, the channel transmission strategy can be divided into the following four cases:

- $SR$  and  $RD$  links both perform semantic transmission, and  $S$  and  $D$  have the same background knowledge ( $BK$ ): In this case, the destination node ( $D$ ) has sufficient storage and computational resources to decode the transmitted semantic information. Additionally, its semantic encoding logic is consistent with that of the source node ( $S$ ), meaning that  $S$  and  $D$  share the same background knowledge ( $BK$ ). The relay node ( $R$ ) functions similarly to a traditional decode - and - forward ( $DF$ ) relay, responsible for decoding and re - encoding data at the semantic level to ensure the integrity and efficient transmission of the semantic information.
- $SR$  and  $RD$  links both perform semantic transmission, but  $S$  and  $D$  have different background knowledge ( $BK$ ): In this case, the destination node ( $D$ ) also has sufficient computational resources to decode semantic information, but its background knowledge is not the same as the source node ( $S$ ). This mismatch in background knowledge may lead to deviations in the expression and understanding of the semantic information. Therefore, the relay node ( $R$ ) needs to perform additional tasks, such as converting and adapting the semantic information to ensure that the destination node can correctly interpret the received semantic information.

- *SR* and *RD* links both perform semantic transmission, and *S* and *D* have the same background knowledge (*BK*), but signal degradation prevents *D* from decoding: In this case, although the destination node (*D*) has sufficient computational resources and shares the same background knowledge as the source node (*S*), the signal suffers significant degradation due to the long distance and poor channel quality, preventing *D* from correctly decoding the semantic information. As a result, the relay node (*R*) needs to operate like a traditional amplify - and - forward (*AF*) relay, amplifying and forwarding the signal. In this scenario, the system's channel capacity is equivalent to that of a two - hop *AF* relay.
- *SR* link performs semantic transmission, *RD* link performs bit transmission: In this case, the source node (*S*) and the relay node (*R*) use semantic transmission, but the relay node (*R*) and the destination node (*D*) only perform bit transmission. This model is applicable when the destination node (*D*) has limited computational resources. The relay node needs to decode the semantic information and re - encode it into a bit stream for transmission to the destination node. In this situation, the transmission performance of the *RD* link is determined by the channel capacity for bit transmission.

In summary, these four cases demonstrate the optimization strategies for transmission under different system conditions. By combining semantic transmission with traditional transmission methods, these strategies address various limitations, such as background knowledge matching, channel quality, and computational resources, providing a theoretical basis and practical guidance for efficient semantic transmission in relay communication systems.

### 3. Performance Analysis

To evaluate the performance of semantic communication in text transmission, this paper uses semantic similarity as the primary performance metric. Specifically, the semantic similarity  $\xi$  is defined as:

$$\xi = \frac{B(s)B(s)^T}{\|B(s)\| \|B(s)\|} \quad (14)$$

where  $B(\cdot)$  represents the bidirectional encoder representations from Transformers (BERT) model, which has made significant improvements in state - of - the - art sentence embedding methods. The pre - trained Sentence - BERT model is used. Unlike other semantic evaluation metrics like BLEU, the BERT similarity can more accurately measure the semantic distance between two sentences. Notably, the range of semantic similarity values is  $0 \leq \xi \leq 1$ , where  $\xi = 1$  indicates that the two sentences are identical, and  $\xi = 0$  means they have no similarity.

Based on semantic similarity, a new performance metric called semantic rate is introduced in literature [2] to measure the semantic information transmission rate achieved by DeepSC. Let  $I$  represent the semantic units (suts), i.e., the average amount of semantic information contained in a sentence  $s$ . Therefore, the semantic information for each semantic symbol can be expressed as  $\frac{I}{KL}$  (units: suts/symbol). Recall that the symbol rate is equal to the transmission bandwidth  $W$ . Thus, the effective semantic rate (units: suts/s) can be expressed as:

$$S = W \cdot \frac{I}{KL} \cdot \xi(K, \gamma) \quad (15)$$

where  $\epsilon$  is a function dependent on the specific semantic communication system and physical channel conditions. The value of  $\epsilon(K, \gamma)$  can be obtained by running the DeepSC tool, with the mapping relationship shown in Figure 2. Semantic similarity has been proven to be a function of received signal - to - noise ratio (SNR)  $\gamma$  and  $K$ . For any given  $K$ , the semantic similarity function  $\epsilon \in [0, 1]$  typically follows an "S" curve as SNR changes. Therefore, a generalized logistic regression method can be used to approximate it as a sigmoid function, expressed as:

$$\xi = a_1 + \frac{a_2}{1 + e^{-(c_1\gamma + c_2)}} \quad (16)$$

where  $a_1, a_2, c_1,$  and  $c_2$  are constant coefficients dependent on  $K$ .

Based on this, the channel capacity formulas for four scenarios can be derived: Based on this, the channel capacity formulas for the four cases can be derived:

### 3.0.1. Semantic Transmission in Both SR and RD Links, with S and D Having the Same Background Knowledge (BK)

At this point, the semantic rate (suts/s) of the SR link can be represented as:

$$R_{SR}^{\text{sem}} = \frac{1}{2} \frac{W}{KL} \xi(K, \gamma_{SR}) \quad (17)$$

where  $W$  is the channel bandwidth,  $I$  is the average amount of semantic information per sentence, measured in semantic units (suts),  $K$  is the average number of semantic symbols per word in the original sentence,  $L$  is the number of words in the sentence, and  $\epsilon(K, \gamma_{SR})$  is the semantic similarity.

The bit rate for transmission is:

$$R_{SR}^{\text{bit}} = \mu \frac{R_{SR}^{\text{sem}}}{I/L} = \frac{1}{2} \mu \frac{W}{K} \xi(K, \gamma_{SR}) \quad (18)$$

where  $\mu$  is the conversion factor for encoding the source information into bits, indicating the average number of bits per word.

The capacity of the RD link is:

$$C_{RD}^{\text{bit}} = \frac{1}{2} W \log_2 \left( 1 + \frac{P_{R\mathcal{G}} |h_{RD}|^2}{N_D} \right) \quad (19)$$

where the bit stream includes encoded semantic symbols, and  $r_c$  represents the average number of bits per semantic symbol using the traditional joint encoder. Thus, the effective semantic symbol rate on the RD link is:

$$R_{RD}^{\text{sem}} = \frac{C_{RD}^{\text{bit}} I}{r_c KL} \quad (20)$$

The effective semantic bit rate on the SR link is:

$$R_{RD}^{\text{bit}} = \mu \frac{R_{RD}^{\text{sem}}}{I/L} = \mu \frac{C_{RD}^{\text{bit}}}{r_c K} = \frac{1}{2} \mu \frac{W}{r_c K} \log_2 \left( 1 + \frac{P_{R\mathcal{G}} |h_{RD}|^2}{N_D} \right) \quad (21)$$

Assuming the predefined channel capacity is  $C$ , the signal - to - noise ratio (SNR) threshold is  $\gamma_{\text{th}}$ , and the outage probability is:

$$P_{\text{out}}(R) = P_r(R_{SD}^{\text{bit}} < C) = 1 - P_r(R_{SD}^{\text{bit}} > C) = 1 - P_r(\min\{R_{SR}^{\text{bit}}, R_{RD}^{\text{bit}}\} > C) \quad (22)$$

Based on  $R_{SR}^{\text{bit}} > C$ , we can derive:

$$\frac{1}{2} \mu \frac{W}{K} \xi(K, \gamma_{SR}) > C \Rightarrow \xi > \frac{2CK}{\mu W} \quad (23)$$

Substituting:

$$\xi(K, \gamma_{SR}) = a_1 + \frac{a_2}{1 + e^{-(c_1 \gamma_{SR} + c_2)}} \quad (24)$$

where  $\xi \in [0, 1]$ , we can obtain:

$$\gamma_{SR} > \frac{-\ln\left(\frac{\frac{2CK}{\mu W} - a_1}{a_2} - 1\right) - c_2}{c_1} \quad (25)$$

The SR link follows a Rayleigh channel model, so the probability  $P_r(R_{SR}^{\text{bit}} > C)$  is:

$$P_r(R_{SR}^{\text{bit}} > C) = 1 - F_{\gamma_{SR}} \left( \frac{-\ln \left( \frac{a_2}{\frac{2cK}{\mu W} - a_1} - 1 \right) - c_2}{c_1} \right) \quad (26)$$

Using the CDF for the Rayleigh distribution, this becomes:

$$P_r(R_{SR}^{\text{bit}} > C) = e^{-\lambda_{SR} \left( \frac{-\ln \left( \frac{a_2}{\frac{2cK}{\mu W} - a_1} - 1 \right) - c_2}{c_1} \right)} \quad (27)$$

Similarly, for  $R_{RD}^{\text{bit}} > C$ , we get:

$$\gamma_{RD} > 2 \frac{2r_c K C}{\mu W} - 1 \quad (28)$$

and the probability:

$$P_r(\gamma_{RD} > 2 \frac{2r_c K C}{\mu W} - 1) = e^{-\lambda_{RD} (2 \frac{2r_c K C}{\mu W} - 1)} \quad (29)$$

Hence, the outage probability is:

$$P_r(R_{SD}^{\text{bit}} < C) = 1 - e^{-\frac{\ln \left( \frac{a_2}{\frac{2cK}{\mu W} - 1} - 1 \right) + c_2}{c_1}} \times e^{-\lambda_{RD} (2 \frac{2r_c K C}{\mu W} - 1)} \quad (30)$$

### 3.0.2. Semantic Transmission in Both SR and RD Links, but S and D Have Different Background Knowledge (BK)

For the SR link, the effective semantic rate is:

$$R_{SR}^{\text{sem}} = \frac{1}{2} \frac{W}{KL} \epsilon(K, \gamma_{SR}) \quad (31)$$

which can be converted to the bit transmission rate:

$$R_{SR}^{\text{bit}} = \mu \frac{R_{SR}^{\text{sem}}}{I/L} = \frac{1}{2} \mu \frac{W}{K} \zeta(K, \gamma_{SR}) \quad (32)$$

For the RD link, the effective semantic rate is:

$$R_{RD}^{\text{sem}} = \frac{1}{2} \frac{W}{KL} \epsilon(K, \gamma_{RD}) \quad (33)$$

and the bit transmission rate is:

$$R_{RD}^{\text{bit}} = \mu \frac{R_{RD}^{\text{sem}}}{I/L} = \frac{1}{2} \mu \frac{W}{K} \zeta(K, \gamma_{RD}) \quad (34)$$

The overall communication capacity is determined by the minimum value of the two links:

$$R_{SD}^{\text{bit}} = \min\{R_{SR}^{\text{bit}}, R_{RD}^{\text{bit}}\} \quad (35)$$

Thus, the outage probability can be expressed as:

$$P_{\text{out}}(R) = P_r(R_{SD}^{\text{bit}} < C) = 1 - P_r(R_{SD}^{\text{bit}} > C) \\ = 1 - P_r(\min\{R_{SR}^{\text{bit}}, R_{RD}^{\text{bit}}\} > C) \quad (36)$$

Using the probability properties, we get:

$$P_{\text{out}}(R) = 1 - P_r(R_{SR}^{\text{bit}} > C) \cdot P_r(R_{RD}^{\text{bit}} > C) \quad (37)$$

Substituting the expressions for  $R_{SR}^{\text{bit}}$  and  $R_{RD}^{\text{bit}}$ , we have:

$$P_{\text{out}}(R) = 1 - P_r \left( \gamma_{SR} > \frac{-\ln \left( \frac{a_2}{\frac{2CK}{\mu W} - 1} - 1 \right) - c_2}{c_1} \right) \cdot P_r \left( \gamma_{RD} > \frac{-\ln \left( \frac{a_2}{\frac{2CK}{\mu W} - 1} - 1 \right) - c_2}{c_1} \right) \quad (38)$$

### 3.0.3. Semantic Transmission in Both SR and RD Links, with S and D Having the Same Background Knowledge (BK), but Signal Degradation Causes D to Fail to Decode

In a two - hop relay communication system, the channel capacity is:

$$C_{SD}^{\text{bit}} = \frac{1}{2} W \log_2 \left( 1 + \frac{\gamma_{SR} \cdot \gamma_{RD}}{1 + \gamma_{SR} + \gamma_{RD}} \right) \quad (39)$$

The semantic channel capacity is derived as:

$$R_{SD}^{\text{sem}} = \frac{C_{SD}^{\text{bit}}}{r_c} \cdot \frac{I}{KL} \quad (40)$$

The bit rate for transmission is:

$$R_{SD}^{\text{bit}} = \mu \frac{R_{SD}^{\text{sem}}}{I/L} = \frac{1}{2} \frac{W\mu}{Kr_c} \log_2 \left( 1 + \frac{\gamma_{SR} \cdot \gamma_{RD}}{1 + \gamma_{SR} + \gamma_{RD}} \right) \quad (41)$$

Based on this, the outage probability is:

$$P_{\text{out}}(R) = P_r(R_{SD}^{\text{bit}} < C) = 1 - P_r(R_{SD}^{\text{bit}} > C) \quad (42)$$

Substituting the expression for  $R_{SD}^{\text{bit}}$ , we get:

$$P_{\text{out}}(R) = 1 - P_r \left( \frac{1}{2} \frac{W\mu}{Kr_c} \log_2 \left( 1 + \frac{\gamma_{SR} \cdot \gamma_{RD}}{1 + \gamma_{SR} + \gamma_{RD}} \right) > C \right) \quad (43)$$

Let  $M = 2^{\frac{2Kr_c C}{W\mu}} - 1$ , and simplify:

$$P_{\text{out}}(R) = 1 - P_r \left( \frac{\gamma_{SR} \cdot \gamma_{RD}}{1 + \gamma_{SR} + \gamma_{RD}} > M \right) \quad (44)$$

Under high SNR conditions, where  $\gamma_{SR} \gg 1$  and  $\gamma_{RD} \gg 1$ , we can approximate:

$$\frac{\gamma_{SR} \cdot \gamma_{RD}}{1 + \gamma_{SR} + \gamma_{RD}} \approx \frac{\gamma_{SR} \cdot \gamma_{RD}}{\gamma_{SR} + \gamma_{RD}} \quad (45)$$

and the final expression for the outage probability involves integrating the joint probability of two independent exponential random variables:

$$P_{\text{out}}(R) = 1 - \int_0^\infty \int_0^{M-u} \lambda_{SR} e^{-\frac{\lambda_{SR}}{u}} \cdot \lambda_{RD} e^{-\frac{\lambda_{RD}}{v}} \frac{1}{u^2} \frac{1}{v^2} du dv \quad (46)$$

### 3.0.4. Semantic Transmission on the SR Link and Bit Transmission on the RD Link

In this case, the effective semantic rate for the SR link is:

$$R_{SR}^{\text{sem}} = \frac{1}{2} \frac{W}{KL} \epsilon(K, \gamma_{SR}) \quad (47)$$

and the bit rate is:

$$R_{SR}^{\text{bit}} = \mu \frac{R_{SR}^{\text{sem}}}{I/L} = \frac{1}{2} \mu \frac{W}{K} \zeta(K, \gamma_{SR}) \quad (48)$$

For the RD link, the bit transmission rate is:

$$R_{RD}^{\text{bit}} = \frac{1}{2} W \log_2 \left( 1 + \frac{P_R |h_{RD}|^2}{WN_R} \right) \quad (49)$$

The overall communication capacity is:

$$R_{SD}^{\text{bit}} = \min\{R_{SR}^{\text{bit}}, R_{RD}^{\text{bit}}\} \quad (50)$$

The outage probability is:

$$P_{\text{out}}(R) = P_r(R_{SD}^{\text{bit}} < C) \quad (51)$$

Simplifying:

$$P_{\text{out}}(R) = 1 - P_r(R_{SD}^{\text{bit}} > C) = 1 - P_r(\min\{R_{SR}^{\text{bit}}, R_{RD}^{\text{bit}}\} > C) \quad (52)$$

The probabilities for the SR and RD links are derived as:

$$P_r(R_{SR}^{\text{bit}} > C) = e^{-\frac{\ln\left(\frac{a_2}{\frac{2CK}{\mu W} - 1} - 1\right) + c_2}{c_1}} \quad (53)$$

and:

$$P_r(R_{RD}^{\text{bit}} > C) = e^{-\lambda_{RD}(2^{\frac{2C}{W}} - 1)} \quad (54)$$

The final outage probability is:

$$P_{\text{out}}(R) = 1 - e^{-\frac{\ln\left(\frac{a_2}{\frac{2CK}{\mu W} - 1} - 1\right) + c_2}{c_1} - \lambda_{RD}(2^{\frac{2C}{W}} - 1)} \quad (55)$$

## 4. Simulation and Results

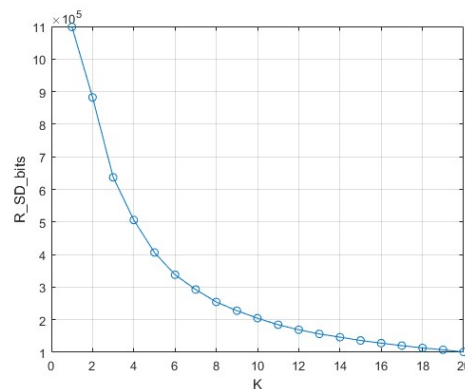
This section presents the theoretical analysis of outage probability for the four proposed semantic communication systems and verifies it through Monte Carlo simulations. The experimental parameters are shown in Table 1. In this simulation, the outage probability for four relay - based semantic communication models is tested under different values of  $K$  and transmit power of the source.

**Table 1.** Parameter values.

Parameter Name	Value
S - Transmission Power ( $P_S$ )	100W
Relay Transmission Power ( $P_R$ )	100W
Noise Power - Spectral Density ( $N_0$ )	$4 \times 10^{-24}$ W/Hz
Capacity Threshold (C)	$8 \times 10^8$ bit/s
Source - Coding Factor ( $\mu$ )	48
Average Number of Semantic Symbols per Word (K)	4symbols/word
Average Number of Bits ( $r_c$ )	8
System Bandwidth (W)	$1.8 \times 10^8$ Hz
Standard Deviation of Channel Gain ( $\sigma$ )	1

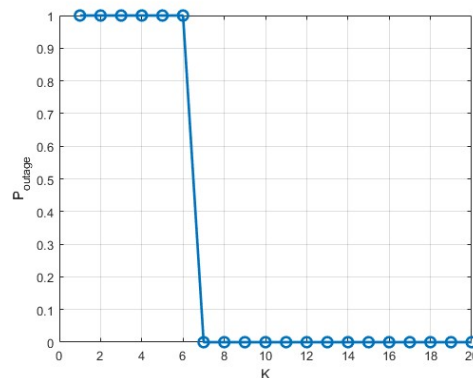
#### 4.1. First Case

In the BD channel, bit transmission is performed, so the SD link's channel capacity is actually determined by the SR link. Since the signal-to-noise ratio (SNR)  $\gamma_{SR}$  is much greater than 1, no matter how the transmit powers of S and R are changed,  $\epsilon(K, \gamma_{RD})$  is approximately equal to 1, meaning the outage probability can be approximated to 0. The SD link's channel capacity decreases as K increases. As shown in Figure 2.

**Figure 2.** The first case: the variation of SD channel capacity with respect to the K value.

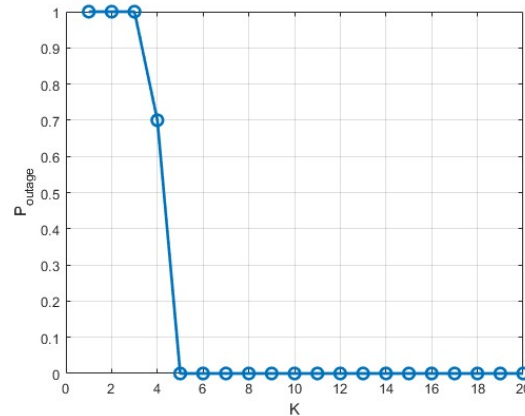
#### 4.2. Second Case

By adjusting the transmit power  $P$  of S and R, the semantic similarity  $\zeta$  between the SR and RD links can be modified, thereby altering the channel capacity of the SD and RD links. As the value of  $P$  increases, the SNR of the link increases. At this point, the SD link's channel capacity is determined by the link with the lower SNR. When K exceeds 6, the outage probability decreases and approaches 0. As shown in Figure 3.

**Figure 3.** The second case: the variation curve of outage probability with respect to the K value.

### 4.3. Third Case

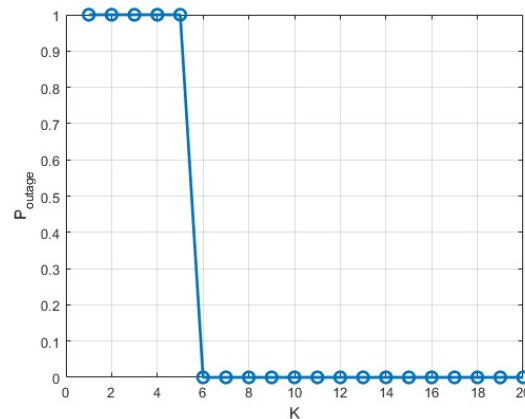
In this case, the outage probability decreases as  $K$  increases. Since the relay performs an amplify - and - forward (AF) function, this case is consistent with a traditional three - node communication model. The outage probability decreases as the SNR increases. As shown in Figure 4.



**Figure 4.** The third case: the variation curve of outage probability with respect to the  $K$  value.

### 4.4. Fourth Case

In this case, the outage probability decreases as  $K$  increases. As shown in Figure 5.



**Figure 5.** The fourth case: the variation curve of outage probability with respect to the  $K$  value.

## 5. Conclusion

This paper provides an in - depth study of the outage probability performance in relay - based semantic communication systems and proposes four transmission strategies under different background knowledge, computational capabilities, and channel conditions. By establishing a three - node communication model, the channel capacity and outage probability expressions for each scenario were derived and verified through Monte Carlo simulations. The research findings indicate that:

- The effectiveness of semantic transmission is influenced by the degree of background knowledge matching. When the background knowledge (BK) of the source node (S) and the destination node (D) is consistent, the semantic transmission performance of the communication system is optimal.
- Relay nodes play a crucial role in semantic conversion and adaptation, particularly when background knowledge does not match. Relay nodes can optimize semantic information transmission, improving the accuracy of semantic decoding.
- Channel conditions have a significant impact on the success rate of semantic transmission. When channel quality is poor, the relay node needs to use an amplify - and - forward (AF) strategy to enhance the signal, thereby reducing the outage probability.

- Hybrid strategies of semantic and bit transmission are suitable for scenarios with limited computational resources. When the computational power of the destination node is insufficient, performing semantic decoding at the relay node and converting it to bit streams can enhance the overall transmission reliability.

The theoretical analysis and simulation results in this study provide new insights for optimizing relay strategies in semantic communication systems and are of significant importance for improving the stability and reliability of future intelligent communication systems. Future research can further consider full - duplex relay modes and joint optimization of resource allocation to further enhance the performance of semantic communication systems.

## References

1. Xie, H.; Qin, Z.; Li, G.Y.; et al. Deep learning enabled semantic communication systems. *IEEE Trans. Signal Process.* **2021**, *69*, 2663–2675.
2. Yan, L.; Qin, Z.; Zhang, R.; et al. Resource allocation for text semantic communications. *IEEE Wireless Commun. Lett.* **2022**, *11*, 1394–1398.
3. Xie, H.; Qin, Z. A lite distributed semantic communication system for Internet of Things. *IEEE J. Sel. Areas Commun.* **2020**, *39*, 142–153.
4. Mu, X.; Liu, Y.; Guo, L.; et al. Heterogeneous semantic and bit communications: A semi-NOMA scheme. *IEEE J. Sel. Areas Commun.* **2022**, *41*, 155–169.
5. Hu, Z.; You, C.; Liu, T.; et al. Semantic Communication Meets Edge Intelligence: Semantic-Relay-Aided Text Transmissions. *IEEE Internet Things J.* **2024**.
6. Liu, T.; You, C.; Hu, Z.; et al. Semantic-Relay-Aided Text Transmission: Placement Optimization and Bandwidth Allocation. In Proceedings of the 2023 IEEE Globecom Workshops (GC Wkshps), 2023, pp. 215–220.
7. Hu, Z.; Liu, T.; You, C.; et al. Multiuser resource allocation for semantic-relay-aided text transmissions. In Proceedings of the 2023 IEEE Globecom Workshops (GC Wkshps), 2023, pp. 1273–1278.
8. Ma, S.; Liang, W.; Zhang, B.; et al. An investigation on intelligent relay assisted semantic communication networks. In Proceedings of the 2023 IEEE Wireless Commun. Netw. Conf. (WCNC), 2023, pp. 1–6.
9. Luo, X.; Yin, B.; Chen, Z.; et al. Autoencoder-based semantic communication systems with relay channels. In Proceedings of the 2022 IEEE Int. Conf. Commun. Workshops (ICC Workshops), 2022, pp. 711–716.
10. Yin, Y.; Huang, L.; Li, Q.; et al. Semantic communication on multi-hop concatenated relay networks. In Proceedings of the 2023 IEEE/CIC Int. Conf. Commun. China (ICCC), 2023, pp. 1–6.
11. Sana, M.; Strinati, E.C. Learning semantics: An opportunity for effective 6G communications. In Proceedings of the IEEE 19th Annu. Consum. Commun. Netw. Conf. (CCNC), Las Vegas, NV, USA, January 2022, pp. 631–636.
12. Lee, C.-H.; Lin, J.-W.; Chen, P.-H.; Chang, Y.-C. Deep learning-constructed joint transmission-recognition for Internet of Things. *IEEE Access* **2019**, *7*, 76547–76561.
13. Bourtsoulatze, E.; Kurka, D.B.; Gündüz, D. Deep joint source-channel coding for wireless image transmission. *IEEE Trans. Cogn. Commun. Netw.* **2019**, *5*, 567–579.
14. Weng, Z.; Qin, Z. Semantic communication systems for speech transmission. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 2434–2444.
15. Tung, T.-Y.; Gündüz, D. DeepWiVe: Deep-learning-aided wireless video transmission. *arXiv preprint* **2021**, arXiv:2111.13034.
16. Niu, K.; Liang, Z.; Dong, C.; et al. Semantics-Division Duplexing: A Novel Full-Duplex Paradigm. *IEEE Wireless Commun.* **2024**, *31*, 307–314.
17. Feng, Y.; Shen, H.; Shan, Z.; et al. Semantic Communication for Edge Intelligence Enabled Autonomous Driving System. *arXiv preprint* **2024**, arXiv:2406.10606.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.