

Article

# Enhancing Communication Reliability from the Semantic Level under Low SNR

Yueling Liu , Yichi Zhang <sup>\*</sup>, Peng Luo, Shengteng Jiang, Kuo Cao <sup>\*</sup>, Haitao Zhao  and Jibo Wei

College of Electronic Science, National University of Defense Technology, Changsha 410073, China; liuyueling16@nudt.edu.cn (Y.L.); pengluo.eric@outlook.com (P.L.); jiangshengteng@nudt.edu.cn (S.J.); haitaozhao@nudt.edu.cn (H.Z.); wjbhw@nudt.edu.cn (J.W.)

<sup>\*</sup> Correspondence: zhangyichi13@nudt.edu.cn (Y.Z.); caokuo18@nudt.edu.cn (K.C.)

**Abstract:** In the low signal-to-noise ratio region, a large number of bit errors occur, and it may exceed the channel error correction capability of the receiver. Traditional communication system may use the technology of automatic repeat-request to deal with this problem, which is time consuming and a waste of resources. To enhance the reliability of the communication system, we investigate reasoning and decoding at the semantic level instead of the grammar level. In particular, we propose a semantic communication model for text transmission, assisting the communication system to be more robust in terrible channel environments. Based on the traditional communication system, the language model BERT, part of speech tagging, and prior information concerning bit-flipping are introduced to enhance the semantic reasoning ability of the transceiver. Furthermore, this paper analyzes the effects of the sub-strategies on the performances of the improved communication model, such as the existence of a candidate set and language model. The numerical results show the effectiveness of our model in terms of improving the semantic accuracy measured by BLUE, the METEOR score, and the similarity score based on BERT between transmitted messages and recovered messages.

**Keywords:** bit-flipping; part of speech; prior information; semantic communication



**Citation:** Liu, Y.; Zhang, Y.; Luo, P.; Jiang, S.; Cao, K.; Zhao, H.; Wei, J. Enhancing Communication Reliability from the Semantic Level under Low SNR. *Electronics* **2022**, *11*, 1358. <https://doi.org/10.3390/electronics11091358>

Academic Editor: Athanasios Kanas

Received: 15 March 2022

Accepted: 22 April 2022

Published: 24 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the rapid development of communication technology, the explosive growth of data has consumed more and more limited available spectrum and power, resulting in a bottleneck for communication development. This is because the existing communication technology has achieved close to the Shannon capacity [1], which is the maximum rate of error-free transmission. Shannon's communication model targets accurate data transmission and the exact recovery of transmitted signals, which means pursuing accuracy at the bit level. To achieve this goal, the performance has been measured in the bit-error rate (BER) or the symbol-error rate (SER). To meet the growing demands of high-data-rate services and solve the limitation of the traditional architecture, it is necessary to upgrade the classic Shannon framework and introduce a new perspective to the communication system design. A true leap forward comes from incorporating semantics into communications [2].

The concept of semantic communication was first proposed by Weaver [3]. In [3], a three-level communication architecture was proposed: the first level is the grammar level, which mainly solves the problem of "how to transmit communication symbols accurately"; the second level is the semantic level, which mainly solves the problem of "how to convey the desired meaning precisely"; the third level is the pragmatic layer, which mainly solves the problem of "how the received meaning effectively affects behavior in the desired way". The scope of semantic communication was further broadened in [4], and three semantic communication sub-areas containing human-to-human, human-to-machine, and machine-to-machine communication were defined. As described in [5], semantic communication aims at the successful transmission of semantic information rather than the accurate reception of each single symbol or bit regardless of its meaning.

Thus, semantic communication provides a novel solution for dealing with limited available spectrum resources by reducing the sending of useless information, and it will play an important role in 6G communication [5].

Due to the vigorous development of artificial intelligence, semantic communication based on neural networks has gradually attracted researcher's attention, and various semantic communication theories and models were proposed. The conceptual framework of the semantic communication system and the definition of semantic information based on logical probabilities were proposed in [6,7]. Additionally, the work in [8] utilized situation-based logical principles to define semantic information. Furthermore, a theory of strongly semantic information was proposed in [9]. The authors in [2] represented semantic information as a factual statement in the propositional logic form and proposed a semantic communication framework incorporating the world model, reasoning process, and background knowledge. As an extension of the literature [2], the authors in [10] further explained the relationship of the model entropy, semantic entropy, and message entropy and proposed the concepts of semantic redundancy and semantic ambiguity. The work of [11] mathematically modeled the average semantic error between information based on semantic similarity. Based on the semantic communication framework, bidirectional long short-term memory model (LSTM) was applied to semantic encoding and decoding of text sources [12,13]. For the image and speech sources, the corresponding semantic feature extraction schemes were investigated in [14–17]. In [18], the authors presented a Transformer-based joint source–channel coding policy. To lower the cost of IoT devices, a lite distributed semantic communication system for IoT networks was developed to achieve better performance on text transmission in [19]. The work in [20] implemented the semantic transmission of speech signals. The authors in [21] studied a robust end-to-end semantic communication system for image sources. Moreover, the semantic communication system of multi-modal data was also investigated in [22]. Since the extraction and utilization of background knowledge are crucial for semantic communication, Y. Zhang et al. [23] proposed an agent-oriented semantic communication architecture, including context knowledge and a semantic encoding and decoding module. Besides, Y. Wang et al. [24] designed a semantic communication system on the basis of knowledge graphs.

In this paper, we focus on information recovery from the semantic level instead of the bit level. More specifically, we consider the semantic transmission of text and investigate using semantics to improve the reliability of communication systems. Due to the language model being widely used for analyzing human language and predicting words from given texts, the language model is introduced into our novel context-based semantic communication system to enhance semantic accuracy. Compared with models that adopted a joint semantic and channel encoding scheme for text transmission, our proposed model designs semantic encoding and channel encoding separately. This brings three main advantages. First, our proposed model is more compatible because it allows the use of existing traditional communication techniques. Our proposed scheme does not change the structure of the traditional communication model and utilizes semantics to serve as the supplement to the traditional communication system. Second, our model is more suitable to be applied in non-differentiable channels, which is more robust to different scenarios. Third, our strategies are more flexible as they can be used in non-jointly designed communication systems. The contributions of this paper are summarized as follows:

- A part-of-speech-based encoding strategy is proposed. The encoding process tries to add some check information about the semantic features to assist the decoding process at the receiver.
- A context-based decoding strategy for recovering messages is proposed, in which the language model is employed to extract the contextual correlation between items, thereby enhancing the semantic reasoning ability of the receiver. Moreover, the prior information concerning the codewords is utilized to improve the semantic accuracy of the recovered messages.

- Semantic metrics of text such as BLUE, the METEOR score, and the similarity score based on BERT are employed to measure the semantic error. Based on the simulation results, the proposed model outperforms the traditional communication system in the low signal-to-noise ratio (SNR) regime.

The rest of this paper is organized as follows. Section 2 establishes the system model and explains the performance metrics. Section 3 analyzes the proposed semantic encoding and decoding strategies concretely. Section 4 presents the simulation results. Finally, the conclusions are drawn in Section 5.

## 2. System Model and Problem Formulation

Considering the transceiver as an intelligent agent, our work attempts to enhance the transceiver with semantic reasoning ability based on the traditional communication system and to improve communication reliability under low SNR.

### 2.1. System Model

We consider a semantic communication system, which is composed of a semantic encoder, channel encoder, channel, channel decoder, and semantic decoder, as shown in Figure 1. At the transmitter, parts of speech such as noun, verb, and adjective are used to categorize words in grammar to add semantic features during encoding. At the receiver, the language model is employed to extract the contextual correlation. Besides, the prior information concerning bit flipping is utilized in the semantic decoding module to help recover sentences. Let  $V$  be the set of all words in the corpus, and we define  $s = [w_1, w_2, \dots, w_m]$  as the sentence with  $m$  items to be transmitted, where  $w_i \in V$  is the  $i$ th word in the sentence. Let  $S_{k_p}(\cdot)$  be the semantic encoder integrating knowledge of part of speech  $k_p$ .  $C(\cdot)$  denotes the channel encoder.

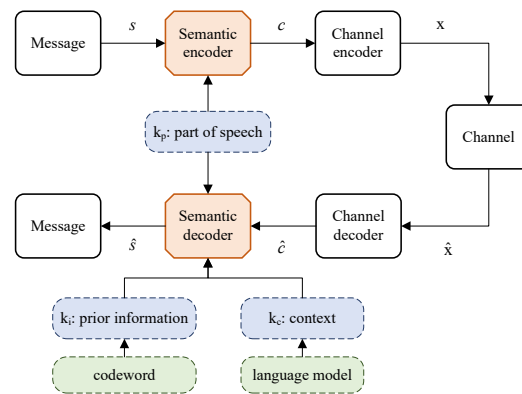


Figure 1. The semantic communication system.

From Figure 1, the transmitter first converts the sentence  $s$  into a sequence of binary bits  $c$  with the help of the semantic encoder  $S_{k_p}(\cdot)$ . Next, binary bits  $c = S_{k_p}(s)$  are fed into the channel encoder to cope with the influence of channel noise and distortion. Thus, the whole encoding process can be represented as follows:

$$x = c(S_{k_p}(s)), \tag{1}$$

where  $x$  is the transmitted signal at the transmitter. Let  $y$  be the sequence of observations at the receiver, which can be formulated as

$$y = hx + n, \tag{2}$$

where  $h$  denotes the channel coefficient and  $n \sim \mathcal{CN}(0, \sigma_n^2)$  represents the additive Gaussian noise.

On the contrary, the received signal will be decoded by passing through the channel decoder and semantic decoder successively. Defining  $C^{-1}(\cdot)$  as the channel decoder, the sequence of observations  $y$  will be converted to the sequence of binary bits  $\hat{c} = C^{-1}(y)$ . Let  $S_{k_i, k_p, k_c}^{-1}(\cdot)$  be the semantic decoder integrating prior information  $k_i$ , part of speech  $k_p$ , and context  $k_c$ . Therefore, the recovered sentence can be represented as follows:

$$\hat{s} = S_{k_i, k_p, k_c}^{-1}(C^{-1}(y)). \quad (3)$$

## 2.2. Performance Metrics

Semantic communication cares less about the consistency between sending and receiving messages, but emphasizes whether the sender and receiver have the same understanding of the message. Therefore, traditional performance metrics such as the BER are no longer suitable for the semantic communication. Here, we measure the performances of the semantic communication using the bilingual evaluation understudy (BLEU) score [25] metric for the evaluation of translation with the explicit ordering (METEOR) score [26] and similarity score based on BERT [27]. BLEU is currently the most widely used automatic evaluation indicator. It evaluates the similarity between the result and the reference translation by using a sliding window. The formula is shown in the following

$$\text{BLEU} = BP \cdot \exp\left(\sum_{n=1}^N w_n \log p_n\right), \quad (4)$$

where  $BP$  denotes the penalty factor,  $w_n$  is the weight of the  $n$ -gram, and  $p_n$  is the  $n$ -gram precision score.

METEOR introduces external knowledge sources, such as WordNet, to achieve word alignment and expand the synonym set. Besides, it considers the parts of speech of words and evaluates the performance of sentences based on the harmonic average of precision and recall. The formula is given as follows:

$$F_{\text{mean}} = \frac{P_m R_m}{\alpha P_m + (1 - \alpha) R_m}, \quad (5)$$

$$\text{METEOR} = (1 - \text{Pen}) F_{\text{mean}}, \quad (6)$$

where  $P_m$  represents precision,  $R_m$  denotes recall,  $\alpha$  is the hyperparameter according to WordNet,  $F_{\text{mean}}$  is the harmonic mean combining the precision and recall, and  $\text{Pen}$  is the penalty coefficient.

The similarity score uses a word vector trained by the BERT model to obtain the semantic evaluation by calculating the cosine similarity. The calculation formula is given as follows:

$$\text{sim}_{\text{v}_{\text{BERT}}}(s_1, s_2) = \frac{\mathbf{v}_{\text{BERT}}(s_1) \mathbf{v}_{\text{BERT}}(s_2)^T}{\|\mathbf{v}_{\text{BERT}}(s_1)\| \|\mathbf{v}_{\text{BERT}}(s_2)\|}, \quad (7)$$

where  $\mathbf{v}_{\text{BERT}}(s_1)$  and  $\mathbf{v}_{\text{BERT}}(s_2)$  are word vectors of sentences  $s_1$  and  $s_2$ , respectively.

All metrics introduced above have values between 0 and 1, which indicate the semantic similarity between recovered text and transmitted text. The number 1 means the highest score, while 0 represents the scenario where two texts have no semantic similarity.

## 3. Encoding and Decoding Strategy

### 3.1. Encoding Strategy

This section introduces the details of the proposed encoding strategy. After preprocessing the corpus and breaking sentences into words, we utilize part of speech tagging (POS tagging) [28], a useful tool of the natural language toolkit, to classify words in grammar. For simplicity, we assign two bits as the check information to represent parts of speech and

group all words into four categories with the help of part-of-speech tagging. Denoting the set of parts of speech as

$$P = \{nouns, adjectives, verbs, others\}. \quad (8)$$

Let the set of the check information be

$$C_p = \{00, 01, 10, 11\}. \quad (9)$$

Define  $E(\cdot)$  as the mapping function for POS tagging. For example, 00, 01, 10, and 11 are set to represent nouns, adjectives, verbs, and others, respectively. For example,  $E(chair) = 00$  will be obtained using the mapping function.

We build the frequency distribution for all words in the corpus and sort them in descending order of their frequency. Let  $H(\cdot)$  be the function for Huffman coding to generate Huffman codewords for all words. The relationship between Huffman codewords and words is stored in the dictionary  $Dict_1(\cdot)$ . Considering both functions of  $H(\cdot)$  and  $E(\cdot)$  map words into bits, we concatenate the Huffman codewords and the check bits for each word. Therefore, the final output of the semantic encoder is described as follows:

$$c_i = H(w_i) \oplus E(w_i). \quad (10)$$

Similarly, the mapping between the final outputs and words is stored in the dictionary  $Dict_2(\cdot)$ . Consequently, the details of the proposed encoding process are summarized in Algorithm 1.

---

**Algorithm 1** Part-of-speech-based encoding method

---

**Input:** a text corpus;

**Output:** codeword dictionary;

- 1: Build frequency distribution for all words in the corpus and sort them in descending order of their frequency.
  - 2: Build function  $H(\cdot)$  for Huffman encoding to generate Huffman codewords for all words.
  - 3: Initialize the  $P, C_p$  and build mapping function for POS tagging  $E(\cdot)$ .
  - 4: **for**  $w_i$  in corpus **do**
  - 5:    $h_i = H(w_i)$
  - 6:    $e_i = E(w_i)$
  - 7:    $c_i = h_i \oplus e_i$
  - 8:    $Dict_1(h_i) = w_i$
  - 9:    $Dict_2(c_i) = w_i$
  - 10: **end for**
  - 11: **return**  $Dict_1, Dict_2$
- 

### 3.2. Decoding Strategy

In this section, the details of the proposed decoding strategy are presented. Let  $\Theta$  be the inverse of the concatenation operator, then the received codewords could be divided into two parts, one of which is the Huffman codeword of the word and the other is the part of speech, shown as follows:

$$[\hat{v}_i, \hat{e}_i] = \Theta(\hat{c}_i). \quad (11)$$

According to the codeword dictionary  $Dict_1(\cdot)$ , Huffman codeword  $\hat{v}_i$  could be converted into a word. However, if the received Huffman codeword cannot be found in the dictionary  $Dict_1(\cdot)$ , it will be replaced with “[MASK]”.

To interpret “[MASK]”, we use language model BERT [29], POS tagging, and prior information concerning bit-flipping. The structure of the language model BERT is a multi-layer bidirectional Transformer encoder, which is pre-trained on an open-domain corpus. Because BERT can make use of the context information to interpret “[MASK]”, a whole

sentence containing “[MASK]” can be fed into BERT, and it can obtain a reasonable result for the position of “[MASK]”. Additionally, BERT could also interpret multiple “[MASK]” according to their context. Thus, we recover multiple “[MASK]” in a sentence together instead of recovering one by one. To obtain the context information, the conditional probability distribution of context  $P(\cdot|\tilde{s}, i; V)$  is obtained by the BERT model, where  $\tilde{s}$  represents the input sentence,  $i$  denotes the position of “[MASK]”, and  $V$  stands for the word list in the original language model. The interpreted results can be denoted as follows:

$$\hat{w}_i = \operatorname{argmax}\{P(\cdot|\tilde{s}, i; V)\}. \quad (12)$$

Nevertheless, BERT has a reasoning ability for general knowledge because of the training corpus. For a specific sentence, the decoded results only based on BERT are affected by the general knowledge, which leads to the sentence mismatch between the recovered sentence and the transmitted sentence. Therefore, we make use of the received wrong bits corresponding to “[MASK]” to produce the candidate set. Besides, this paper utilizes the candidate set as the prior information and POS tagging as the check information to improve decoding accuracy.

To make full use of the received wrong codeword corresponding to “[MASK]”, we flip the received codeword  $\hat{c}_i$  by 1 to  $N$  bits to produce flipped codewords  $\hat{c}_{ij}$ , and the total number of flipped codewords is shown as follows:

$$J = \sum_{n=1}^N C_{\operatorname{length}(\hat{c}_i)}^n, \quad (13)$$

where  $\operatorname{length}(\hat{c}_i)$  stands for the total number of bits of the word  $\hat{c}_i$ . For example, if the received codeword is 0000 ( $\operatorname{length}(\hat{c}_i) = 4$ ) and the maximum number of bit-flipping is set as 2 ( $N = 2$ ), then the number of flipped codewords for received codeword 0000 will be 10. Next,  $\hat{c}_{ij}$  is divided into two parts, which are given by

$$[\hat{v}_{ij}, \hat{e}_{ij}] = \Theta(\hat{c}_{ij}). \quad (14)$$

Converting Huffman codewords  $\hat{v}_{ij}$  into words and building the candidate set corresponding to each “[MASK]”, we then map the check information  $\hat{e}_{ij}$  to parts of speech and filter out the Huffman codeword  $\hat{v}_{ij}$  with the mismatch of its tag of the parts of speech. Next, we put sentences with “[MASK]” and their final corresponding candidate sets into the BERT model. Combing the candidate set as the new word list, BERT predicts the word  $w_i$  from the candidate set  $\bar{V}_i$ . Then, the interpreted results (12) can be approximately described as

$$\hat{w}_i = \operatorname{argmax}\{P(\cdot|\tilde{s}, i; \bar{V}_i)\}, \quad (15)$$

where  $P(\cdot|\tilde{s}, i; \bar{V}_i)$  represents the new conditional probability distribution of the context. According to the BERT model,  $P(\cdot|\tilde{s}, i; \bar{V}_i)$  can be obtained from the following formula:

$$P(\cdot|\tilde{s}, i; \bar{V}_i) = \operatorname{softmax}\{Emb(\operatorname{context})^T \cdot Emb(\bar{V}_i)\}, \quad (16)$$

where  $Emb(\operatorname{context})$  is the contextual representations produced by BERT at the  $i$ th position of the sentence and  $Emb(\bar{V}_i)$  is the word embedding matrix in candidate set  $\bar{V}_i$ .

Consequently, the estimated sentence can be obtained as  $\hat{s} = [\hat{w}_1, \hat{w}_2, \dots, \hat{w}_m]$ . The details of the proposed decoding process are summarized in Algorithm 2.

**Algorithm 2** Context- and prior-information-based decoding method

---

**Input:** codeword dictionary  $Dict_1$ , received codewords  $\hat{c}$ ;  
**Parameter:** number of bit-flipping  $N$ ;  
**Output:** the optimal sequence;

- 1: **for**  $\hat{c}_i$  in  $\hat{c}$  **do**
- 2:    $\hat{v}_i, \hat{e}_i = \Theta(\hat{c}_i)$
- 3:   **if**  $Dict_1(\hat{v}_i)$  exist **then**
- 4:      $\tilde{w}_i = Dict_1(\hat{v}_i)$
- 5:   **else**
- 6:      $\tilde{w}_i = [MASK]$
- 7:     **for**  $n$  in  $N$  **do**
- 8:       flip  $\hat{c}_i$  by  $n$  bits
- 9:        $C_{\text{length}(\hat{c}_i)}^n$  flipped codewords are obtained
- 10:    **end for**
- 11:     $J = \sum_{n=1}^N C_{\text{length}(\hat{c}_i)}^n$
- 12:    **for**  $j$  in  $J$  **do**
- 13:      $\hat{v}_{ij}, \hat{e}_{ij} = \Theta(\hat{c}_{ij})$
- 14:     **if**  $E(\hat{v}_{ij}) == \hat{e}_{ij}$  **then**
- 15:       pass
- 16:     **else**
- 17:       remove  $\hat{v}_{ij}$
- 18:     **end if**
- 19:    **end for**
- 20:    Build a candidate set  $\tilde{V}_i$  for  $i$ th word in the sentence
- 21: **end if**
- 22: **end for**
- 23: Put sentence  $\tilde{s} = [\tilde{w}_1, \dots, \tilde{w}_i]$  and final candidate set  $\tilde{V}_i$  into BERT model
- 24: **for** each [MASK] **do**
- 25:    $P(\cdot | \tilde{s}, i; \tilde{V}_i) = \text{softmax}\{Emb(\text{context})^T \cdot Emb(\tilde{V}_i)\}$
- 26:    $\hat{w}_i = \text{argmax}\{P(\cdot | \tilde{s}, i; \tilde{V}_i)\}$
- 27: **end for**
- 28: **return**  $\hat{s}$

---

#### 4. Simulation Results

In this section, numerical results are provided to verify the effectiveness of the proposed model. For the baseline, we adopted the Huffman coding and LDPC coding as the traditional source coding and channel coding methods, respectively (baseline: Huffman + LDPC). To further verify our proposed model, we compared the results with the scheme of [23], which similarly adopted a non-jointly designed communication system for text transmission. For brevity, the scheme in [23] “context-based semantic communication” is labeled as context-based SC in the following. As described above, our proposed model adopts language model BERT, POS tagging, and candidate sets concerning bit-flipping (named Method A). To study the impacts of the language model, one strategy considers introducing the language model into the traditional communication system (Method B: tradition+ language model). To investigate the effectiveness of our proposed model, all encoding and decoding experiments were based on words (e.g., “hello”) instead of characters (e.g., “h”, “e”, “l”, “l”, “o”). When calculating semantic scores for the baseline, the received Huffman codewords that cannot be found in the dictionary were replaced by random words in the dictionary, which helps reduce the semantic gap caused by the sentence alignment problem.

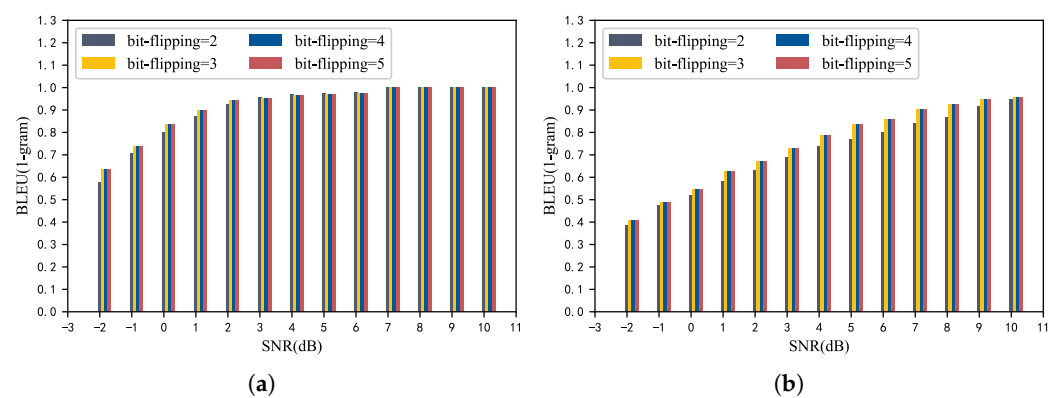
##### 4.1. Parameter Setup

For comparison, we considered the English Wikipedia as the text for transmission, and we only extracted the text passages with long contiguous sequences and ignored lists, tables, and headers. In Huffman encoding, we used the English Wikipedia as the corpus

to compute frequencies, which were then used to generate the Huffman codebook. Parts of speech were simply set as four kinds. The BERT model was set as “bert-base-uncased”, whose specific parameters are shown in Table 1. To investigate the impacts of the changes of the number of bit-flips on the performance of our proposed model, BLEU(1-g) scores under different numbers of bit-flips versus the SNR over the AWGN channel and the Rayleigh fading channel are drawn in Figure 2. From Figure 2, the BLEU performance of 1-g was optimal when the number of bit-flips of the codeword was greater than or equal to 3. Thus, the number of bit-flips of the codeword was set as 3 to obtain the candidate set.

**Table 1.** Parameters of the used BERT model.

The Number of Attention Heads	The Number of Layers	Hidden Size	Parameters
12	12	768	110 M



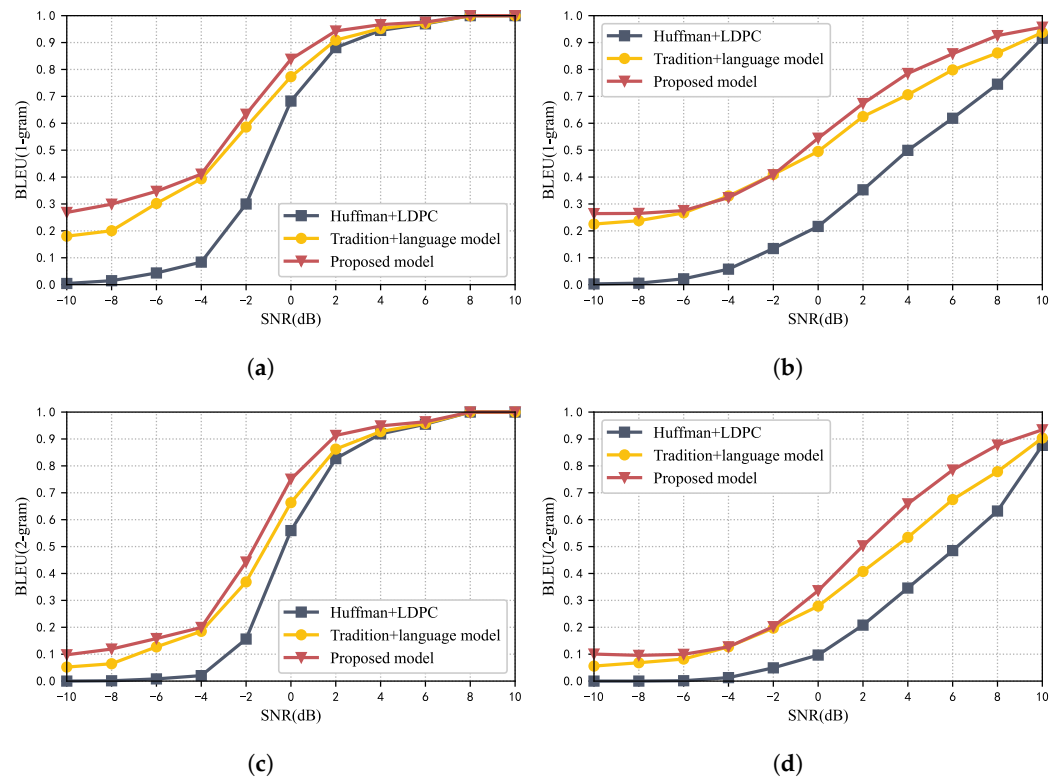
**Figure 2.** BLEU(1-g) scores under different numbers of bit-flips versus SNR over the AWGN channel and the Rayleigh fading channel. (a) AWGN channel. (b) Rayleigh channel.

#### 4.2. Performance Analysis

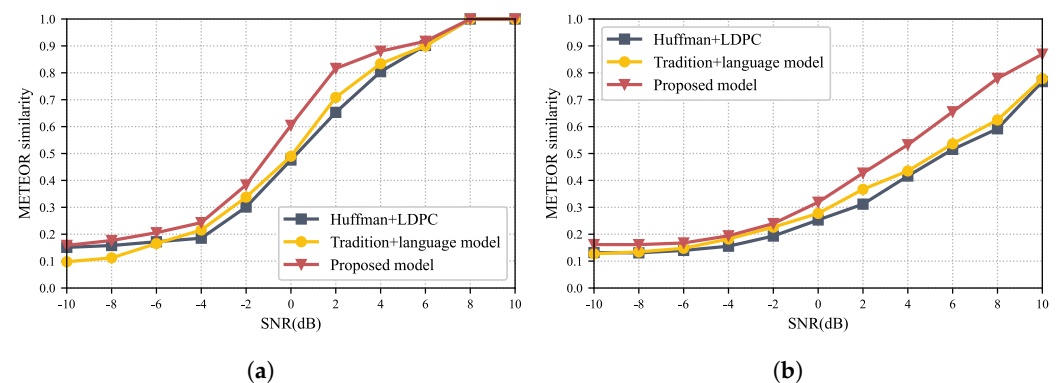
Figure 3 shows the BLEU score versus the SNR over the AWGN channel and the Rayleigh fading channel, respectively. As observed, the 1–2 g BLEU scores improved with the increasing of the SNR, and Methods A and B outperformed the baseline across the entire SNR range, especially under a low SNR. For both channel conditions, our proposed model outperformed the baseline in terms of the BLEU score and achieved the largest performance gain with the 1-g BLEU score. Besides, the introduction of the language model brought great improvements in performance. It is worth noting that the joint addition of parts of speech and the candidate set also produced higher 1–2 g BLEU scores, especially for 2-g BLEU. In Figure 3a,c, due to the protection of the channel coding, the scores of all methods approached 1 when the SNR was above 7 dB. On the other hand, with the severe impacts of Rayleigh fading, the BLEU scores in Figure 3b,d rose slowly with the SNR increasing, and the scores of all methods over the Rayleigh fading channels did not reach 1, even though the SNR was 10 dB. However, benefiting from our proposed strategies, the performances improved significantly when compared with the baseline. Specifically, both the 1-g BLEU score and 2-g BLEU score of our proposed model in Rayleigh fading channels reached above 0.8 when the SNR was above 6 dB.

Figure 4 draws the relationship between the METEOR score and the SNR over the AWGN channel and the Rayleigh fading channel, respectively. From Figure 4, the trend of the METEOR scores for all methods was similar to the trend of the BLEU scores, but the METEOR scores of our methods showed more advantages than the baseline since METEOR pays more attention to the fluency of sentences compared to BLEU. The joint introduction of parts of speech and the candidate set brought a noticeable performance gain over the traditional methods, especially in the Rayleigh fading channel.





**Figure 3.** BLEU score versus SNR over the AWGN channel and the Rayleigh fading channel. (a) BLEU (1-g) score over the AWGN channel. (b) BLEU (1-g) score over the Rayleigh channel. (c) BLEU (2-g) score over the AWGN channel. (d) BLEU (2-g) score over the Rayleigh channel.

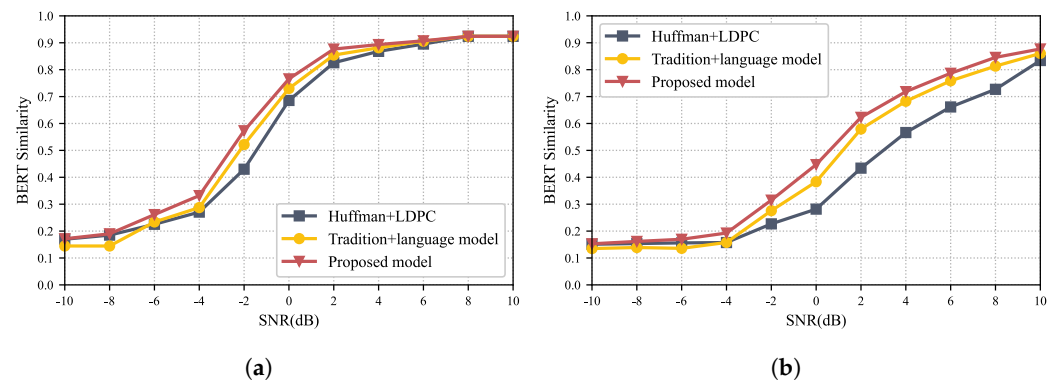


**Figure 4.** (a) METEOR score versus SNR over the AWGN channel. (b) METEOR score versus SNR over the Rayleigh fading channel.

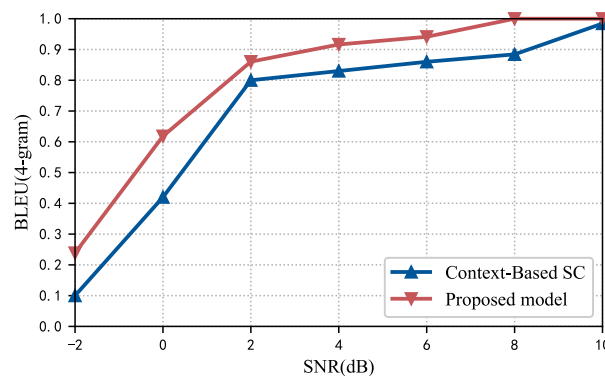
Figure 5 plots the relationship between the similarity score based on BERT and the SNR over the AWGN channel and the Rayleigh fading channel, respectively. Being trained on large corpora, the similarity score based on BERT could capture semantic relationships among words and be more relevant to human judgments. Therefore, the introduction of the language model resulted in higher scores in terms of the similarity score based on BERT. It can be seen from Figure 5 that the proposed model had excellent performances in both the AWGN and Rayleigh channel, particularly over the Rayleigh fading channel.

Figure 6 shows the 4-g BLEU scores of our proposed model and context-based SC versus the SNR over the AWGN channel. From Figure 6, our model outperformed context-based SC across the entire SNR range over the AWGN channel. In our proposed model, the parts of speech were utilized as the check information to reduce the computational complexity of constructing candidate sets and improving the semantic accuracy. However, the parts of speech in context-based SC were used for data compression. That is the main

reason why the proposed model achieved better performance than context-based SC in terms of the 4-g BLEU scores.



**Figure 5.** (a) Similarity score based on BERT versus SNR over the AWGN channel. (b) Similarity score based on BERT versus SNR over the Rayleigh fading channel.



**Figure 6.** BLEU (4-g) score of our proposed model and context-based SC versus SNR over the AWGN channel.

To ensure a fair comparison, we calculated the execution time of the baseline (Huffman + LDPC), Method B (tradition + language), and our proposed model, respectively. All simulations were conducted on the same platform and were performed by a computer with Inter Xeon Silver 4110 CPU@2.1 GHz and NVIDIA GeForce GTX 2080TI. We transmitted 300 sentences using different schemes from the transmitter to the receiver for the test. The average execution times for each sentence on three methods are shown in Table 2. From Table 2, the execution time of the baseline was the shortest, and our proposed model increased only 6.9% compared to the execution time of the Method B algorithm. It can be concluded that our proposed model enhances the communication reliability by increasing the computational complexity.

**Table 2.** The average execution time for each sentence on three methods.

Methods	Huffman + LDPC	Tradition + Language	Proposed Model
Time complexity/s	4.8561	7.7494	8.2832

### 5. Conclusions

In this paper, a novel semantic communication model that incorporates the language model, prior information, and parts of speech was proposed to improve the reliability. We used the language model and prior information to enhance the semantic reasoning ability of the receiver. Moreover, the semantic accuracy and computational complexity were improved by using the parts of speech as the check information. The numerical results

illustrated that our proposed model improved the performance in terms of semantic metrics, especially in the low SNR regime. However, the proposed decoding scheme heavily relies on the assumption that the received context that can be found in the dictionary is correct. In fact, the received codewords that can be found in the dictionary may be wrongly decoded as other words, leading to the incorrect context. Based on the above considerations, we will consider potential errors at each position of a sentence in the future work, to achieve better semantic accuracy.

**Author Contributions:** Conceptualization, Y.L.; methodology, Y.L. and Y.Z.; formal analysis, P.L. and S.J.; investigation, K.C.; supervision, H.Z.; writing—original draft preparation, Y.L. and Y.Z.; writing—review and editing, K.C., H.Z. and J.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded in part by the National Natural Science Foundation of China under Grant Nos. 61931020, U19B2024, and 62001483 and in part by the science and technology innovation Program of Hunan Province under Grant No. 2021JJ40690.

**Data Availability Statement:** Data are available from the authors, on request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Technol. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
2. Bao, J.; Basu, P.; Dean, M.; Partridge, C.; Swami, A.; Leland, W.; Hendler, J.A. Towards a theory of semantic communication. In Proceedings of the IEEE Network Science Workshop, West Point, NY, USA, 22–24 June 2011; pp. 110–117.
3. Weaver, W. Recent contributions to the mathematical theory of communication. *ETC A Rev. Gen. Semant.* **1953**, *10*, 261–281.
4. Lan, Q.; Wen, D.; Zhang, Z.; Zeng, Q.; Chen, X.; Popovski, P.; Huang, K. What is semantic communication? A view on conveying meaning in the era of machine intelligence. *arXiv* **2021**, arXiv:2110.00196.
5. Qin, Z.; Tao, X.; Lu, J.; Li, G.Y. Semantic communications: Principles and challenges. *arXiv* **2021**, arXiv:2201.01389.
6. Carnap, R.; Bar-Hillel, Y. An outline of a theory of semantic information. *J. Symb. Logic* **1954**, *19*, 230–232.
7. Bar-Hillel, Y.; Carnap, R. Semantic information. *Br. J. Philos. Sci.* **1953**, *4*, 147–157. [[CrossRef](#)]
8. Barwise, J.; Perry, J. Situations and attitudes. *J. Philos.* **1981**, *78*, 668–691. [[CrossRef](#)]
9. Floridi, L. Outline of a theory of strongly semantic information. *Minds Mach.* **2004**, *14*, 197–221. [[CrossRef](#)]
10. Basu, P.; Bao, J.; Dean, M.; Hendler, J. Preserving quality of information by using semantic relationships. In Proceedings of the 2012 IEEE International Conference on Pervasive Computing and Communications Workshops, Lugano, Switzerland, 19–23 March 2012.
11. Güler, B.; Yener, A.; Swami, A. The semantic communication game. *IEEE Trans. Cogn. Commun. Netw.* **2018**, *4*, 787–802. [[CrossRef](#)]
12. Farsad, N.; Rao, M.; Goldsmith, A. Deep learning for joint source–channel coding of text. In Proceedings of the 2018 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018.
13. Rao, M.; Farsad, N.; Goldsmith, A. Variable length joint source–channel coding of text using deep neural networks. In Proceedings of the 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Kalamata, Greece, 25–28 June 2018.
14. Bourtsoulatzé, E.; Kurka, D.B.; Gündüz, D. Deep joint source–channel coding for wireless image transmission. *IEEE Trans. Cogn. Commun. Netw.* **2019**, *5*, 567–579. [[CrossRef](#)]
15. Kurka, D.B.; Gündüz, D. DeepJSCC-f: Deep joint source–channel coding of images with feedback. *IEEE J. Sel. Areas Commun.* **2020**, *1*, 178–193. [[CrossRef](#)]
16. Kurka, D.B.; Gündüz, D. Bandwidth-Agile image transmission with deep joint source–channel coding. *IEEE Trans. Wirel. Commun.* **2021**, *20*, 8081–8095. [[CrossRef](#)]
17. Jankowski, M.; Gündüz, D.; Mikolajczyk, K. Wireless image retrieval at the edge. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 89–100. [[CrossRef](#)]
18. Xie, H.; Qin, Z.; Li, G.Y.; Juang, B.-H. Deep learning enabled semantic communication systems. *IEEE Trans. Signal Process.* **2021**, *69*, 2663–2675. [[CrossRef](#)]
19. Xie, H.; Qin, Z. A lite distributed semantic communication system for Internet of Things. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 142–153. [[CrossRef](#)]
20. Weng, Z.; Qin, Z. Semantic communication systems for speech transmission. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 2434–2444. [[CrossRef](#)]
21. Hu, Q.; Zhang, G.; Qin, Z.; Cai, Y.; Yu, G. Robust semantic communications against semantic noise. *arXiv* **2022**, arXiv:2202.03338v1.
22. Xie, H.; Qin, Z.; Li, G.Y. Task-oriented multi-user semantic communications for VQA. *IEEE Wirel. Commun. Lett.* **2022**, *11*, 553–557. [[CrossRef](#)]

23. Zhang, Y.; Zhang, P.; Wei, J. Semantic communication for intelligent devices: Architectures and a paradigm. *Sci. Sin. Inform.* **2021**. [[CrossRef](#)]
24. Wang, Y.; Chen, M.; Saad, W.; Luo, T.; Cui, S.; Poor, H.V. Performance optimization for semantic communications: An attention-based learning approach. In Proceedings of the 2021 IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, 7–11 December 2021.
25. Papineni, K.; Roukos, S.; Ward, T.; Zhu, W. BLEU: A method for automatic evaluation of machine translation. In Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, Philadelphia, PA, USA, 7–12 July 2002.
26. Banerjee, S.; Lavie, A. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization, Ann Arbor, MI, 29 June 2005.
27. Available online: <https://pypi.org/project/semantic-text-similarity/> (accessed on 14 March 2022).
28. Bird, S. NLTK: The natural language toolkit. In Proceedings of the Joint Conference of the International Committee on Computational Linguistics and the Association for Computational Linguistics (COLING/ACL) on Interactive Presentation Sessions, Sydney, Australia, 17–18 July 2006.
29. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT), Stroudsburg, PA, USA, 2–7 June 2019.