

Article

Reliable Semantic Communication System Enabled by Knowledge Graph

Shengteng Jiang , Yueling Liu *, Yichi Zhang, Peng Luo, Kuo Cao *, Jun Xiong , Haitao Zhao  and Jibo Wei

College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China; jiangshengteng@nudt.edu.cn (S.J.); zhangyichi13@nudt.edu.cn (Y.Z.); pengluo.eric@outlook.com (P.L.); xj8765@nudt.edu.cn (J.X.); haitaozhao@nudt.edu.cn (H.Z.); wjbhw@nudt.edu.cn (J.W.)

* Correspondence: liuyueling16@nudt.edu.cn (Y.L.); caokuo18@nudt.edu.cn (K.C.)

Abstract: Semantic communication is a promising technology used to overcome the challenges of large bandwidth and power requirements caused by the data explosion. Semantic representation is an important issue in semantic communication. The knowledge graph, powered by deep learning, can improve the accuracy of semantic representation while removing semantic ambiguity. Therefore, we propose a semantic communication system based on the knowledge graph. Specifically, in our system, the transmitted sentences are converted into triplets by using the knowledge graph. Triplets can be viewed as basic semantic symbols for semantic extraction and restoration and can be sorted based on semantic importance. Moreover, the proposed communication system adaptively adjusts the transmitted contents according to channel quality and allocates more transmission resources to important triplets to enhance communication reliability. Simulation results show that the proposed system significantly enhances the reliability of the communication in the low signal-to-noise regime compared to the traditional schemes.



Citation: Jiang, S.; Liu, Y.; Zhang, Y.; Luo, P.; Cao, K.; Xiong, J.; Zhao, H.; Wei, J. Reliable Semantic Communication System Enabled by Knowledge Graph. *Entropy* **2022**, *24*, 846. <https://doi.org/10.3390/e24060846>

Academic Editors: Onur Günlü, Rafael F. Schaefer, Holger Boche and H. Vincent Poor

Received: 26 April 2022

Accepted: 18 June 2022

Published: 20 June 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/>).

Keywords: semantic communication; knowledge graph; semantic extraction; semantic restoration

1. Introduction

In recent years, wireless communication technology has developed rapidly, bringing great convenience to human life. Fifth-generation (5G) wireless communication technology has played an important role in smart cities, autonomous driving, telemedicine, and other fields [1]. However, with the gradual increase in the communication rate, the explosive growth of data has created enormous challenges for wireless communication technology [2]. According to the forecast from the International Telecommunication Union (ITU), the annual growth rate of the global mobile data stream will reach up to 55% by 2030 [3]. Moreover, the transmission rate of existing communication technologies has gradually approached the Shannon capacity [4], which cannot meet the continuously growing communication demands in the future 6G era. In the future, the 6G communication system will play an important role in remote holography [5], digital twin [6], and other application fields. Therefore, the sixth-generation wireless communication system needs to provide an ultra-high peak rate, ultra-large user experience rate, and ultra-low network latency, which will consume more limited available spectrum and power and bring huge challenges to communication technology. Semantic communication is one of the effective techniques used to overcome these challenges [7].

Semantic communication, as a revolution against traditional communication, is a new communication paradigm [8]. The concept of semantic communication was first proposed by Weaver (1949) [9]. After Shannon (1948) put forward the classical information theory [4], Weaver proposed that communication should be divided into three different layers, namely the technical layer, semantic layer, and effectiveness layer. The technical layer represents traditional communication, focusing on “how to accurately transmit communication symbols”.

The semantic layer focuses on “how to accurately convey the meaning of communication symbols”; the effectiveness layer focuses on “how the received meaning effectively affects the receiver’s behavior”. Compared with traditional communication, semantic communication aims to reduce the uncertainty of message understanding between the transmitter and the receiver. Moreover, semantic communication mainly transmits semantic-relevant information, which greatly reduces the amount of redundant data. Therefore, semantic communication is a suitable technology (against the scenarios) with limited communication bandwidth and a low signal-to-noise ratio (SNR) [10,11].

However, some fundamental problems of semantic communication have not been effectively solved. One of them is semantic representation, which limits the development of semantic communication [7]. Regarding semantic representation—existing research studies tend to use transmitted content features to represent the semantics. This representation lacks human language logic and cannot be interactive verification with human understanding [12]. To solve this problem, we considered using the knowledge graph instead of features to represent semantics. The knowledge graph can decompose text into multiple semantic units without losing semantics [13], ensuring the accuracy of semantic representation. The basic structure of the knowledge graph is a triplet in the form of an “entity-relation-entity” [13]. From the linguistic point of view, a single entity may have multiple types of semantic information. The specific semantic information can be determined after a relationship is formed between entities, so the triplet in the knowledge graph can be regarded as the smallest semantic symbol. There have been some research studies exploring the relationship between the knowledge graph and semantics. Jaradeh et al. (2019) proposed that the knowledge graph was the next-generation infrastructure for semantic scholarly knowledge [14]. Mosa (2021) proposed that the knowledge graph could help with semantic category prediction [15]. Zhou et al. (2022) combined the knowledge graph with semantic communication to improve the validity of communication [16]. Thus, the knowledge graph can effectively represent semantics; we investigated the semantic communication system based on the knowledge graph (SCKG) for improving communication reliability. The main contributions of this paper are summarized as follows:

- A semantic extraction method is proposed to extract triplets from transmitted text to represent its core semantic information, reducing the information redundancy of the transmitted text.
- A semantic restoration method based on text generation from the knowledge graph is proposed, which completes the semantic restoration process by reconstructing the text structure between entities and relations.
- A novel semantic communication system was developed, which can sort triplets based on semantic importance and adaptively adjust the transmitted contents according to the channel quality.

The rest of this paper is organized as follows. Section 2 briefly reviews the related work. Section 3 details the proposed system and the semantic extraction and restoration methods used in the model. Experimental results are presented in Section 4 to verify the performance of the proposed model. Finally, Section 5 concludes this paper.

2. Related Work

2.1. Semantic Communication Development

Due to technical limitations in the early stage of communication development, researchers have focused on solving engineering problems at the technical layer and postponed the study at the semantic layer. However, this does not mean that the research on semantic communication will be shelved. With the advancements in technology, the semantic problem has become an urgent problem that needs to be solved in the communication field [17].

In terms of theoretical research, Carnap et al. (1954) first proposed the concept of the semantic information theory to supplement the classical information theory [18]. They thought that the semantic information contained in the sentence should be defined based

on the logical probability of the content of the sentence. Floridi (2004) proposed a theory of strongly semantic information [19] and pointed out the problem that sentence contradictions will have infinite information. Bao et al. (2011) put forward a general model of semantic communication, using a factual statement in the propositional logic form to represent semantics [20]. Moreover, the semantic entropy, semantic noise, and semantic channel capacity were defined in [20]. Based on the literature [20], Basu et al. (2012) provided a detailed explanation of the relationship between semantic entropy and information entropy, and they defined the concepts of semantic ambiguity and semantic redundancy [21]. In [22], Lan et al. (2021) proposed that semantic communication can be divided into human-to-human, human-to-machine, and machine-to-machine sub-areas, which broadened the scope of semantic communication.

On the other hand, the rapid development of neural networks and artificial intelligence technology promotes the progress of technical research in semantic communication. In terms of semantic coding, the authors of [23] proposed a joint source-channel coding for semantic information with a bidirectional long short-term memory model (BILSTM). As an extension of the literature [23], Rao et al. (2018) presented a variable-length joint source-channel coding of semantic information [24]. In [25], Liu et al. (2022) proposed a semantic encoding strategy based on parts-of-speech and context-based decoding strategies, which enhanced communication reliability from the semantic level. Based on the semantic communication framework, Xie et al. (2021) proposed a deep learning-based semantic communication model [26], which used word embedding technology to map text to semantic space and then performed source-channel joint encoding for semantic information by using the transformer framework [27]. Furthermore, the authors of [28] proposed a lightweight distributed semantic communication system for the application scenario of the internet of things (IoT), which reduced the cost of IoT devices. The authors of [29] proposed a semantic communication model based on reinforcement learning to investigate the impact of noisy environments on semantic information. In different information forms, Weng et al. (2021) proposed a semantic communication model for speech transmission [30]. In [31], Hu et al. (2022) proposed a robust end-to-end semantic communication system to combat the semantic noise for image transmission. Moreover, a semantic communication model based on multi-information modalities was developed in [32]. Regarding semantic representation, Zhou et al. (2022) used the transformer for semantic extraction and semantic restoration [33].

2.2. Performance Metrics

Semantic communication, different from traditional communication systems, does not emphasize the perfect recovery of the transmitted message, but rather on the receiver correctly understanding the message in the same way as the transmitter. As a result, performance metrics commonly used in traditional communication systems (e.g., bit error rate and symbol error rate) are no longer suitable for semantic communication. Hence, this paper uses the bilingual evaluation understudy (BLEU) score [34], a metric for evaluation of translation with the explicit ordering (METEOR) score [35], and the semantic similarity score [36], as performance metrics.

2.2.1. BLEU Score

BLEU is currently the most commonly used metric in text evaluation [37]. It evaluates the similarity by counting the number of the same n -grams between transmitted and received texts, where n -gram means n consecutive words in the text. The formula can be expressed as

$$\log \text{BLEU} = \min\left(1 - \frac{l_s}{l_t}, 0\right) + \sum_{n=1}^N \omega_n \log p_n \quad (1)$$

where s and \hat{s} denote the transmitted sentence and restored sentence, respectively. l_s and $l_{\hat{s}}$ are the lengths of the transmitted sentences s and restored sentence \hat{s} , respectively. ω_n represents the weight of n-grams, and p_n denotes the precision of n-grams.

2.2.2. METEOR Score

METEOR extends the synonym set by introducing external knowledge sources, such as WordNet [38]. Furthermore, it uses precision P_m and recall R_m to evaluate the similarity between transmitted and received texts. The formula is given as follows

$$F_{\text{mean}} = \frac{P_m R_m}{\alpha P_m + (1 - \alpha) R_m} \quad (2)$$

$$\text{METEOR} = (1 - \text{Pen}) F_{\text{mean}} \quad (3)$$

where α is the hyperparameter according to WordNet, F_{mean} represents the harmonic mean combining P_m and R_m , and Pen is the penalty coefficient.

2.2.3. Semantic Similarity Score

The semantic similarity score converts text into vectors by using the BERT model [39]. It evaluates the semantic similarity between sentences by comparing the degree of similarity between vectors. For the transmitted sentence's vector $v(s)$ and the received sentence's vector $v(\hat{s})$, the semantic similarity score can be expressed as

$$\text{sim}_v(s, \hat{s}) = \frac{v(s) \cdot (v(\hat{s}))^T}{\|v(s)\| \cdot \|v(\hat{s})\|} \quad (4)$$

All the performance metrics introduced above take values between 0 and 1. A higher score given by the performance metrics means that the received text's semantic is closer to the transmitted text's semantic; 0 means semantically irrelevant; 1 means semantically consistent.

3. System Model

As shown in Figure 1, the structure of the proposed system consists of a semantic extraction module, traditional communication architecture, and semantic restoration module. The proposed system can be divided into two levels, which are the semantic level and the technical level. The structure of the technical level is the same as that of the traditional communication system; thus, we mainly introduce the details at the semantic level. At the transmitter, the semantic extraction module can extract the knowledge graph (KG) of the transmitted sentence to represent its semantics. More importantly, the knowledge graph is sorted according to semantic importance. At the receiver, the semantic restoration module can recover the transmitted sentence according to the received knowledge graph.

Figure 2 shows examples of the proposed semantic communication system in different channel qualities. At the transmitter, the transmitted sentence is first converted into the knowledge graph through the semantic extraction module. Next, the transmitter adjusts the knowledge graph according to the channel quality. Then, the knowledge graph is transmitted through the channel. With the noisy knowledge graph received, the semantic is recovered through the semantic restoration module. In Figure 2a, when the channel quality is good, the transmitted sentence and the restored sentence convey the same semantics although they have different sentence structures. When the channel quality is poor, all triplets cannot be transmitted correctly. Therefore, the proposed semantic communication system chooses to transmit the most important triplet. When it comes to Steve Jobs, people tend to care about his relationship with Apple rather than the college he graduated from. As shown in Figure 2b, the transmitter only sends "< Steve Jobs-founder-Apple" when the channel quality is poor.

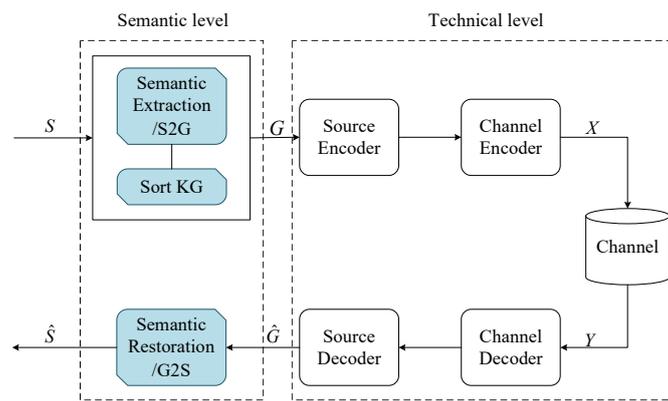


Figure 1. The structure of the proposed semantic communication system based on the knowledge graph, including the semantic extraction module, traditional communication architecture, and semantic restoration module.

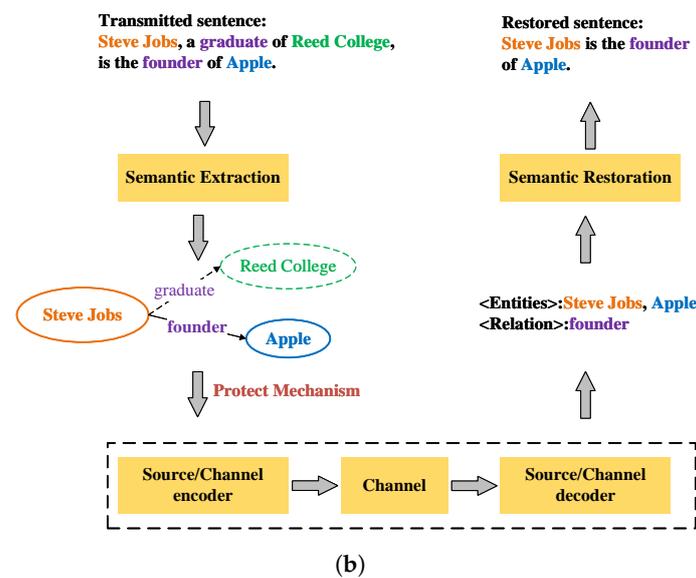
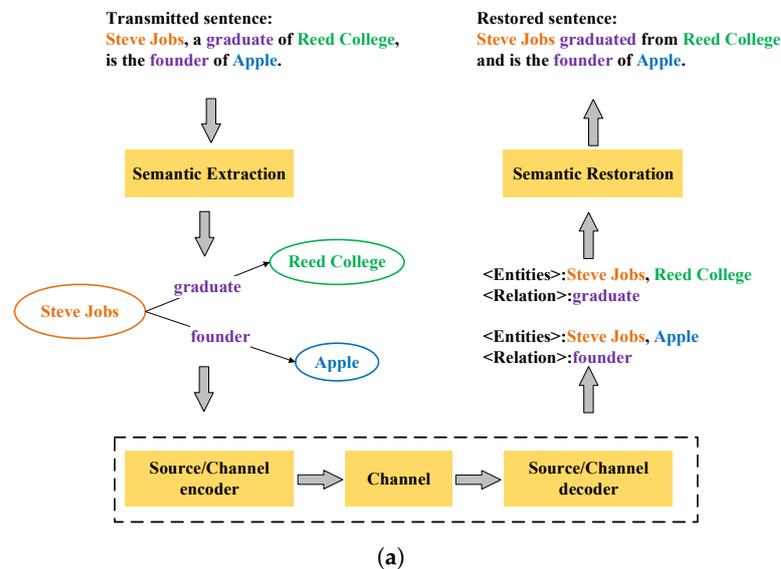


Figure 2. Examples of the proposed semantic communication system in different channel qualities. (a) An example of the proposed semantic communication system when the channel quality is good. (b) An example of the proposed semantic communication system when the channel quality is poor.

3.1. Semantic Extraction Method

To represent the semantic information correctly, the semantic extraction module at the transmitter uses a deep learning network to extract the knowledge graph from the transmitted sentence. Let $S2G_{\theta}(\bullet)$ be the function of the proposed semantic extraction method, which takes the sentence $S = [w_1, w_2, \dots, w_m]$ as input and its corresponding output is the knowledge graph G , where w_m is the m th word in the sentence. The deep learning network structure for the semantic extraction method is shown in Figure 3.

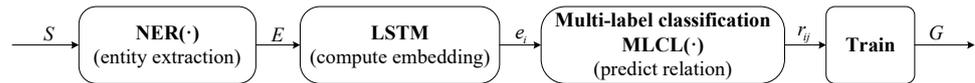


Figure 3. The deep learning network structure for the semantic extraction method.

In particular, we used the pipeline method to extract the knowledge graph, which means extracting the entities in S and then predicting the relations between entities. Firstly, we used a well-established named entity recognition model (NER) to extract the entities [40]. This model is based on the conditional random field classifier and Gibbs sampling. The conditional random field classifier combines the characteristics of the maximum entropy model and the hidden Markov model, and it is often used to deal with sequence labeling tasks, such as parts-of-speech tagging and named entity recognition. Gibbs sampling is a method of generating Markov chains that can be used for Monte Carlo simulations. Based on the conditional random field classifier and Gibbs sampling, NER is trained by using a large amount of manually annotated text and can recognize entities from given sentences. Therefore, the entities in the transmitted sentence can be expressed as

$$E = [en_1, en_2, \dots, en_i, \dots, en_L] = \text{NER}(S) \tag{5}$$

where en_i represents the i th entity in the sentence, L is the total number of entities contained in the sentence.

After extracting entities from S , we predict the relations between the two entities. Firstly, the embedding of each word w_j in the entity en_i is averaged to obtain the entity's embedding. The embedding of w_j can be obtained by using a long short-term memory model (LSTM) [41] to encode w_j and its context. The formula is given as follows

$$\text{emb}(w_j) = \text{LSTM_encode}(w_j, w_{<j}, w_{>j}) \tag{6}$$

Therefore, the i th entity's embedding e_i can be represented as

$$e_i = \frac{1}{\text{Len}(en_i)} \sum_{w_j \in en_i} \text{emb}(w_j) \tag{7}$$

where $\text{Len}(en_i)$ is the number of words in the entity en_i .

Then we feed the entity embeddings into a multi-label classification layer $\text{MLCL}(\bullet)$ to predict the relations. The multi-label classification layer $\text{MLCL}(\bullet)$ can take in two entities and predict the possible relation set. To prevent these two entities from being irrelevant, the relation set includes the "no-relation" type. The relation set between the i th entity and the j th entity can be represented as

$$r_{ij} = \text{MLCL}(e_i, e_j) \tag{8}$$

Since the knowledge graph is made of entities and relations, the probability of extracting a graph from a given sentence is equivalent to the product of the probability of extracting the relation set given any two entities. The formula can be expressed as

$$p(G | S) = \prod_{i=0}^L \prod_{j=0}^L p(r_{ij} | e_i, e_j, S) \tag{9}$$

Based on the probability $p(G | S)$, we can denote the loss function of the proposed semantic extraction method by using the negative log-likelihood loss, which can be formulated as

$$\begin{aligned} \mathcal{L}_{S2G}(\theta) &= \mathbb{E}[-\log p(G | S; \theta)] \\ &= \mathbb{E} \left[-\log \prod_{i=0}^L \prod_{j=0}^L p(r_{ij} | e_i, e_j, S; \theta) \right] \end{aligned} \tag{10}$$

where θ is the network parameter set of the deep learning network, which is shown in Figure 3.

Utilizing the loss function \mathcal{L}_{S2G} , the optimal parameter set θ^* can be easily found using the gradient descent method. Consequently, the details of the proposed semantic extraction method can be summarized in Algorithm 1.

Algorithm 1 The proposed semantic extraction method

Input: the transmitted sentence S

- 1: Build entity set E by Equation (5)
- 2: **for** each $en_i \in E$ **do**
- 3: Compute the embedding e_i by Equations (6) and (7)
- 4: **end for**
- 5: Construct the relation set according to Equation (8)
- 6: Compute loss function $\mathcal{L}_{S2G}(\theta)$ according to Equation (10)
- 7: Train $\theta \rightarrow \theta^*$

Output: The knowledge graph G

3.2. Semantic Restoration Method

The proposed semantic restoration method—similar to the proposed semantic extraction method—uses deep learning to generate sentences from the received knowledge graph. The generated sentence can help the receiver understand the semantics of the transmitted sentence. Let $G2S_\varphi(\bullet)$ be the function of the proposed semantic restoration. The input of $G2S_\varphi(\bullet)$ is the received knowledge graph \hat{G} and its output is the restored sentence \hat{S} . The deep learning network structure for the semantic restoration method is shown in Figure 4.

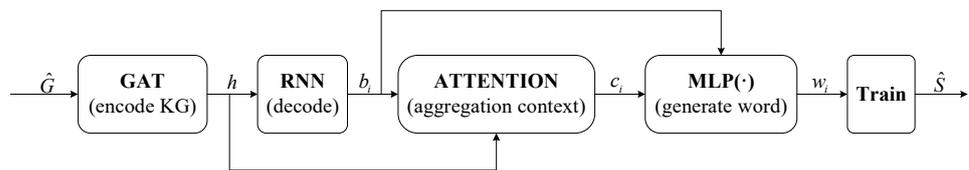


Figure 4. The deep learning network structure for the semantic restoration method.

At first, we encoded the received knowledge graph \hat{G} to convert it to the embedding, which could be processed by the deep learning network. Specifically, we used the graph attention network (GAT) [42] to calculate the embedding of the received knowledge graph \hat{G} . GAT is a representative graph convolutional network that can encode the knowledge graph by introducing the attention mechanism into the knowledge graph. Therefore, the embedding of \hat{G} can be represented as

$$h = \text{GAT}(\hat{G}) \tag{11}$$

After obtaining the embedding h , we used the recurrent neural network (RNN) and the attention mechanism to generate the sentence word by word. Each step of RNN can produce a word embedding. In the i th step, the embedding b_i can be represented as

$$b_i = \text{RNN}(b_{i-1}, w_{i-1}) \quad (12)$$

where w_{i-1} is the $i - 1$ th word in the generated sentence, b_{i-1} is the embedding produced in the $i - 1$ th step. To improve the accuracy of the generated sentence, the attention mechanism was used to obtain the embedding of contextual information. The formula can be described as

$$c_i = \text{ATTENTION}(b_i, h) \quad (13)$$

where c_i denotes the contextual information of the i th word. Then we fed the word embedding b_i and the contextual information c_i into a multilayer perceptron (MLP) to generate the i th word w_i .

Consequently, the generation of w_i based on the received knowledge graph \hat{G} and all previously generated words $w_{<i}$ was fulfilled by predicting the word w_i through MLP with the assistance of the word embedding b_i and the contextual information c_i . Thus, the probability of recovering word w_i can be represented as

$$p(w_i | w_{<i}, \hat{G}) \propto \exp(\text{MLP}([b_i; c_i])) \quad (14)$$

In summary, the probability of generating a sentence from the received knowledge graph \hat{G} is equivalent to the product of the probability of generating each word. The probability can be described as

$$p(\hat{S} | \hat{G}) = \prod p(w_i | w_{<i}, \hat{G}) \quad (15)$$

Similarly, we used the negative log-likelihood loss to denote the loss function of the proposed semantic restoration method according to the probability $p(\hat{S} | \hat{G})$. The loss function can be represented as

$$\begin{aligned} \mathcal{L}_{G2S}(\varphi) &= \mathbb{E}[-\log p(\hat{S} | \hat{G}; \varphi)] \\ &= \mathbb{E}[-\log \prod p(w_i | w_{<i}, \hat{G}; \varphi)] \end{aligned} \quad (16)$$

where φ is the network parameter set of the deep learning network, which is shown in Figure 4. Finally, the gradient descent can be used to find the optimal parameter set φ^* for minimizing the loss function $\mathcal{L}_{G2S}(\varphi)$.

The details of the proposed semantic restoration process are summarized in Algorithm 2.

Algorithm 2 The proposed semantic restoration method

Input: the received knowledge graph \hat{G}

- 1: Compute the embedding of \hat{G} by Equation (11)
- 2: **while** w_i is not the satisfied end feature **do**
- 3: Compute b_i by Equation (12)
- 4: Compute the contextual information c_i by Equation (13)
- 5: Generate word w_i according to Equation (14)
- 6: **end while**
- 7: Compute the loss function $\mathcal{L}_{G2S}(\varphi)$ according to Equation (16)
- 8: Train $\varphi \rightarrow \varphi^*$

Output: the knowledge graph \hat{S}

3.3. System Process

In this section, we introduce the overall process of the proposed semantic communication system. Let $S = [w_1, w_2, \dots, w_m]$ be the transmitted sentence, where w_m is the m th word in the sentence. As shown in Figure 5, with the help of the proposed semantic extraction method $S2G_\theta(\bullet)$, the transmitter converts the transmitted sentence S to the knowledge graph G , which can be represented as $G = S2G_\theta(S)$. The knowledge graph G consists of n triplets and it can be formulated as $G = [g_1, g_2, \dots, g_n]$.

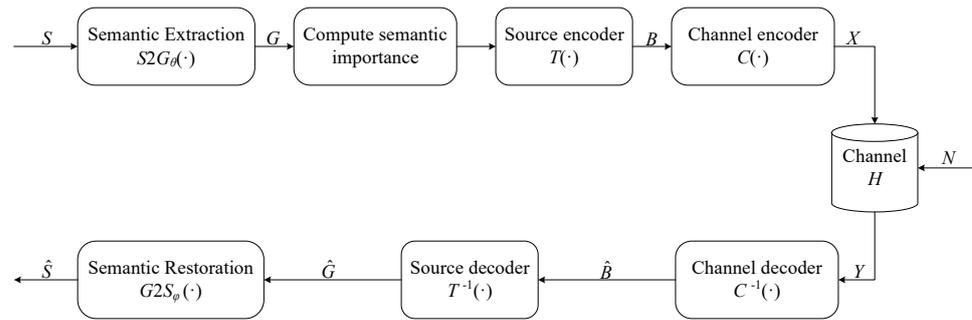


Figure 5. The overall process of the proposed semantic communication system based on the knowledge graph, combining the proposed semantic extraction method, the proposed semantic restoration method, and the traditional communication architecture.

Using the proposed semantic extraction method, the transmitted sentence is converted into a series of triplets. In this process, the semantics of the transmitted sentence are extracted without losing semantics [13]. During transmission, these triplets are independent of each other, which means that errors in some triplets will not affect other triplets. However, in Markov models, once there is a transmission error, the whole transmitted sentence will be affected. Therefore, the proposed semantic communication system is more robust under a low SNR. Moreover, different semantic basic symbols (triplets) have semantic importance in semantic communication, unlike bits or symbols that are treated equally in traditional communication, such as longer-range models and Markov chain-based probabilistic models. These triplets (with semantic importance) should be treated differently. The triplets with important semantics should be allocated with many time slots and bandwidth resources. When the channel quality is extremely poor, instead of transmitting all triplets, which cannot be guaranteed by the channel, it is better to ensure that the most important triplet can be transmitted correctly. When the channel quality is better, the system can adjust the sending content according to semantic importance. Motivated by the different triplets with semantic importance, we sort these triplets according to their semantic similarity scores:

$$sim_v(s, g_i) = \frac{v(s) \cdot (v(g_i))^T}{\|v(s)\| \cdot \|v(g_i)\|} \tag{17}$$

where g_i denotes the i th triplet in G . Table 1 shows an example of semantic importance. From Table 1, “< Steve Jobs – founder-Apple>” is more important than “< Steve Jobs – graduate-Reed College >”, which is also in line with human perception.

Table 1. An example of semantic importance.

Sentence	Triples of Knowledge Graph	Semantic Similarity
Steve Jobs, a graduate of Reed College, is the founder of Apple	Steve Jobs – graduate-Reed College Steve Jobs – founder-Apple	0.56 0.73

Based on the sorted triplets, we can adaptively adjust the number of transmitted triplets according to the channel quality. When the channel quality is extremely poor, we only transmit the most significant triplet and use the communication resources of triplets not transmitted to protect it. As the channel quality improves, we increase the number of transmitted triplets.

After the transmitted knowledge graph G is obtained, the transmitter first maps it into a binary bit stream $B = T(G)$, and then feeds the binary bit stream into the channel encoder to cope with the effects of channel noise and distortion. Therefore, the whole process of the transmitter can be represented as

$$X = C(T(G)) \quad (18)$$

where $T(\bullet)$ and $C(\bullet)$ denote the source encoder and the channel encoder, respectively. If X is sent, the received signal can be represented as

$$Y = HX + N \quad (19)$$

where H is the channel coefficient and $N \sim \mathcal{CN}(0, \sigma_n^2)$ denotes the additive white Gaussian noise.

After obtaining the received signal, the receiver will decode it to recover the transmitted knowledge graph. Defining $C^{-1}(\bullet)$ and $T^{-1}(\bullet)$ as the channel decoder and the source decoder, respectively, the received knowledge graph \hat{G} can be represented as

$$\hat{G} = T^{-1}(C^{-1}(Y)) \quad (20)$$

Then we use the proposed semantic restoration method $G2S_\varphi(\bullet)$ to obtain the restored sentence \hat{S} .

$$\hat{S} = G2S_\varphi(\hat{G}) \quad (21)$$

The process of the proposed semantic communication system is shown in Algorithm 3.

Algorithm 3 Process of the proposed semantic communication system.

Input: The transmitted sentence S

- 1: **Transmitter:**
- 2: Extract the knowledge graph by Algorithm 1
- 3: **for** $i = 1$ to n **do**
- 4: Compute the semantic importance of g_i by Equation (17)
- 5: **end for**
- 6: Sort the knowledge graph according to the semantic importance
- 7: Adjust the number of transmitted triplets according to the channel quality
- 8: $C(T(G)) \rightarrow X$
- 9: Transmit X over the channel
- 10: **Receiver:**
- 11: Receive Y
- 12: $T^{-1}(C^{-1}(Y)) \rightarrow \hat{G}$
- 13: Restore the sentence \hat{S} by Algorithm 2

Output: The restored sentence \hat{S}

4. Experimental Results

In this section, we compare the proposed SCKG with other traditional models under different channels, including the AWGN channel and the Rayleigh fading channel to verify the effectiveness of SCKG. In Table 2, we introduce the models used in the experiments, including their general features and technical methods. It is worth noting that the traditional communication models are not the only ones mentioned in Table 2. The source coding

can also choose arithmetic coding, L–Z coding, and other coding methods. Identically, the channel coding can also choose turbo code, polar code, and other coding methods.

Table 2. Introduction to the proposed model and other traditional models.

Model	General Features	Technical Methods
SCKG	(1) Adding the semantic extraction module and semantic restoration module into traditional communication architecture. (2) Using triplets as semantic basic symbols for semantic extraction and restoration.	(1) Semantic extraction—network structure using NER + LSTM. (2) Semantic restoration—network structure using GAT + RNN + ATTENTION.
Huffman [43] + LDPC [44]	(1) Using the traditional communication architecture from Shannon’s information theory. (2) Using Huffman coding as source coding and using LDPC coding as channel coding.	(1) Convert transmitted sentences to bit sequences by using Huffman coding. (2) Using LDPC coding to combat channel distortion.
DeepNN [23]	(1) Using the deep neural network for source-channel joint coding. (2) Replacing source encoding and channel encoding with the encoder of the deep neural network. (3) Replacing source decoding and channel decoding with the decoder of the deep neural network.	(1) Encoder—network structure using BILSTM. (2) Decoder—network structure using LSTM.

4.1. Experimental Settings

In the simulation, the adopted dataset was the WebNLG dataset [45], which is usually used to generate sentences from knowledge graphs. Each data in the dataset consists of multiple triplets and their corresponding sentences. After preprocessing the dataset, we obtained 12,597 training data, 1746 validation data, and 2493 test data. The training and testing environment was Ubuntu16.04+CUDA10.1, the selected deep learning framework was PyTorch 1.6.0. The training settings of the semantic extraction method and the semantic restoration method are shown in Table 3.

Table 3. Training settings for semantic extraction and restoration method.

Type	Parameters for Semantic Extraction Method	Parameters for Semantic Restoration Method
Epochs	50	50
Batch size	32	32
Optimizer	Adam	Adam
Learning rate	5×10^{-5}	2×10^{-4}
Drop	0	0.1

In the experiment, the test data of WebNLG were transmitted sentence-by-sentence to the transmitter. Then we obtained the restored sentences by using the above-mentioned methods at the receiver. After the restored sentences were obtained, the experimental results could be calculated according to the performance metrics.

For the benchmark, we adopted the traditional communication architecture with source coding and channel coding, where source coding could use Huffman coding, arithmetic coding, L–Z coding, etc., and channel coding could use LDPC coding, turbo code, polar code, etc. For simplicity, we adopted the combination of Huffman coding and LDPC

coding (named “Huffman + LDPC”). Moreover, we considered another two methods as ablation experiments to validate the effectiveness of the proposed model. One involved using the proposed model without adaptive transmission and semantic restoration (named the “Proposed model without AT and SR”), and the other involved using the proposed model without adaptive transmission (named the “Proposed model without AT”).

4.2. Experimental Result Analysis

4.2.1. Performance of the Proposed Semantic Communication System

First, we investigated the effects of the number of triplets on the semantic performance under different SNRs. Here, we considered three strategies, one strategy was to send the first triplet (named “Send the 1st triplet”), and the other two schemes involved sending 50% triplets (named “Send 50% triplets”) and 100% triplets (named “Send 100% triplets”), respectively. Moreover, we compared these three strategies with the benchmark and an end-to-end deep learning-based communication system proposed in [23] (named DeepNN). Figure 6 shows the performance of the semantic similarity versus the SNR in this experiment. From Figure 6, “Send the 1st triplet” has the best semantic similarity under a low SNR because it uses the most resources to protect the first triplet. With the SNR becoming better, “Send 50% triplets” has better performance because “Send the 1st triplet” transmits limited semantics, and the accuracy of the scheme “Send 100% triplets” cannot be guaranteed due to the channel distortion. The semantic similarity of “Send 100% triplets” is above the others at a high SNR, which is reasonable due to the error-free transmission when the channel quality is good. Meanwhile, all three strategies outperformed the benchmark and DeepNN in their superior SNR regions. According to Figure 6, it is reasonable to send the most important triplet in the low SNR region, send 50% triplets in the medium SNR region, and send 100% triplets in the high SNR region.

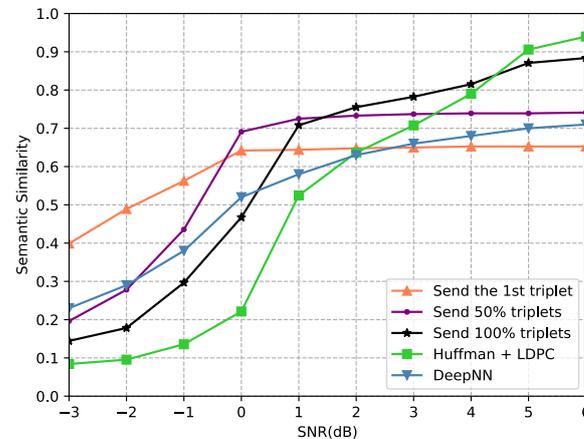


Figure 6. Semantic similarity versus the SNR under the AWGN channel, with send the 1st triplet; send 50% triplets; send 100% triplets; Huffman + LDPC; DeepNN.

Figure 7 demonstrates the relationship between the SNR and the BLEU score under the AWGN channel. From Figure 7, the proposed model performs better under a low SNR in terms of the 1-gram BLEU score or 2-gram BLEU score due to the protection of important triplets. Moreover, after converting the received triplets into sentences by using the proposed semantic restoration method, “Proposed model without AT” outperforms “Proposed model without AT and SR” for all SNR regimes. However, the performance of the proposed model is inferior to the traditional communication system in the high SNR region in Figure 7. This is because the proposed semantic restoration method attempts to recover the same semantic rather than the same sentence structure. For example, the transmitted sentence is “Steve Jobs was the founder of Apple”, and the restored sentence is “Steve Jobs founded Apple”. Although the two sentences are semantically consistent, the BLEU score of the proposed scheme is poor.

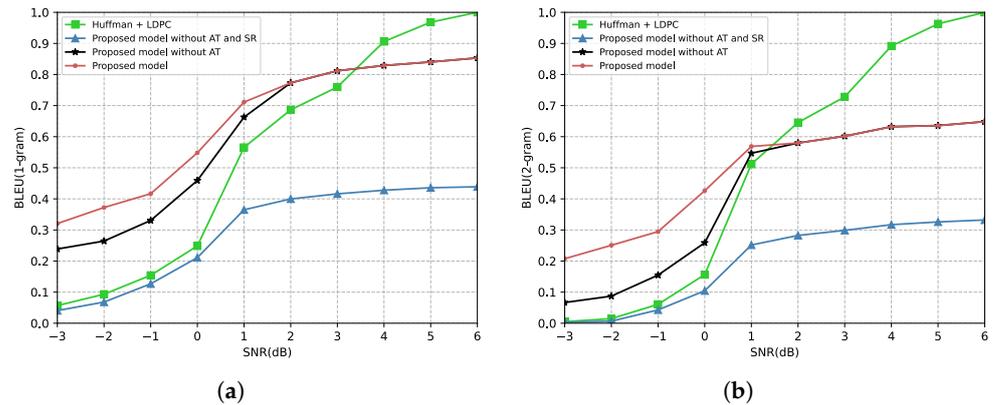


Figure 7. BLEU score versus the SNR over the AWGN channel. (a) BLEU(1-gram) score over the AWGN channel. (b) BLEU(2-gram) score over the AWGN channel.

Figure 8 shows the relationship between the SNR and the BLEU score under the Rayleigh fading channel. All scores in Figure 8 are lower than the scores in Figure 7 because of the severe impacts of Rayleigh fading. However, the proposed model significantly improves performance compared to the benchmark. From Figure 8, the proposed model outperforms the benchmark across the SNR range over the Rayleigh fading channel, either the 1-gram BLEU score or the 2-gram BLEU score. It reflects that our proposed model is more robust to complex communication environments. Meanwhile, since “Proposed model without AT” and “Send 100% triplets” are identical in the high SNR region, the results of the proposed model and “Proposed model without AT” are the same when the SNR is higher than 2 dB.

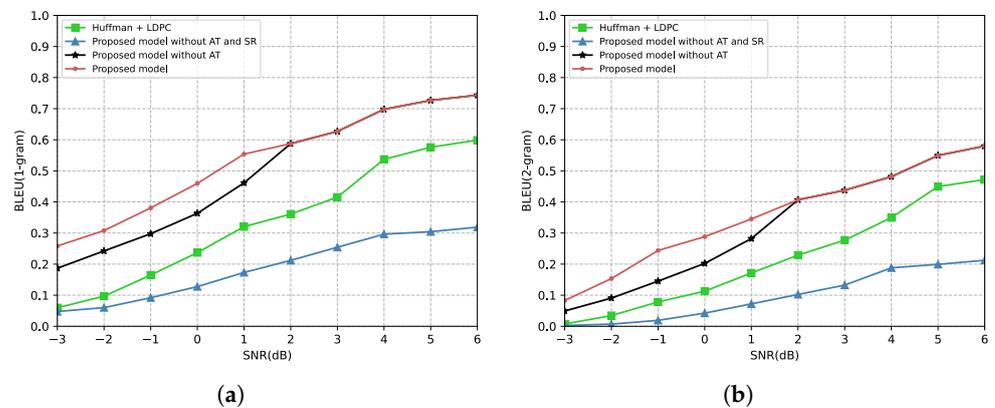


Figure 8. BLEU score versus the SNR over the Rayleigh fading channel. (a) BLEU(1-gram) score over the Rayleigh fading channel. (b) BLEU(2-gram) score over the Rayleigh fading channel.

Since BLEU is an evaluation metric that calculates scores based on word matching, sentence sizes can affect the performance of our proposed model. To research this, we divided the transmitted sentences into three groups—sentence length between 0 and 15, sentence length between 15 and 30, and sentence length greater than 30. Figure 9 shows the relationship between the SNR and the (1-gram) BLEU score under the AWGN channel and the Rayleigh fading channel, respectively. From Figure 9a, “Sentence Length (0, 15)” is significantly higher than the other two groups. This is because the proposed model only transmits the most important triplet in the low SNR, and the length of the restored sentence is limited. In the low SNR region, the BLEU score decreases as the sentence length increases. With the SNR increasing, the number of the transmitted triplets increases, and the gaps between the different groups narrow. In Figure 9b, the gaps between the different groups are not obvious due to the effects of Rayleigh fading.

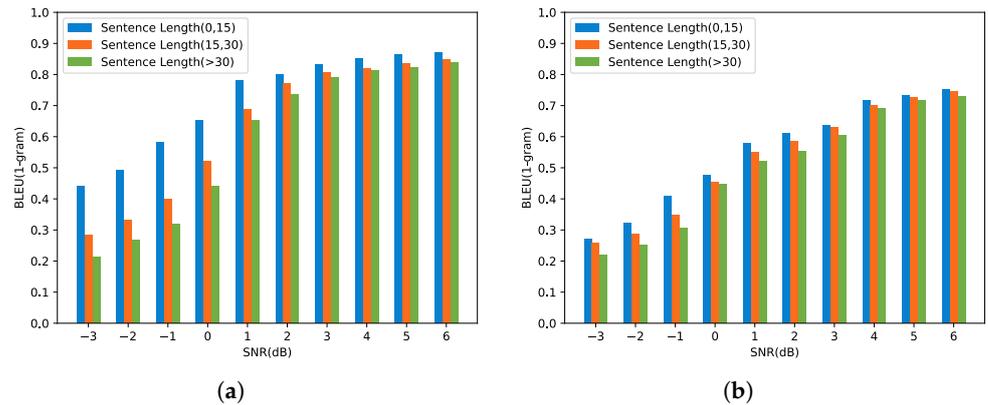


Figure 9. BLEU (1-gram) score versus the SNR with sentence length (0, 15). Sentence Length (15, 30); sentence length (>30). (a) BLEU (1-gram) score over the AWGN channel. (b) BLEU (1-gram) score over the Rayleigh fading channel.

Figure 10 shows the METEOR score versus the SNR over the AWGN channel and the Rayleigh fading channel. From Figure 10a, the score of the benchmark is close to 1 and higher than our proposed model when the SNR is above 4 dB. This is because the few errors that occurred during the transmission were corrected by the channel coding at a high SNR; the benchmark could restore the transmitted sentence without distortion. However, our proposed model discards the information of sentence structure during transmission. When the SNR is less than 4 dB, the channel coding cannot correct all errors during transmission. In this situation, the METEOR score of the benchmark degrades rapidly. However, the proposed model reduces the number of transmitted triplets and protects important triplets in this case, which leads to a better performance in the low SNR region. From Figure 10b, even under the Rayleigh fading channel, our model outperforms the benchmark in all SNR regions.

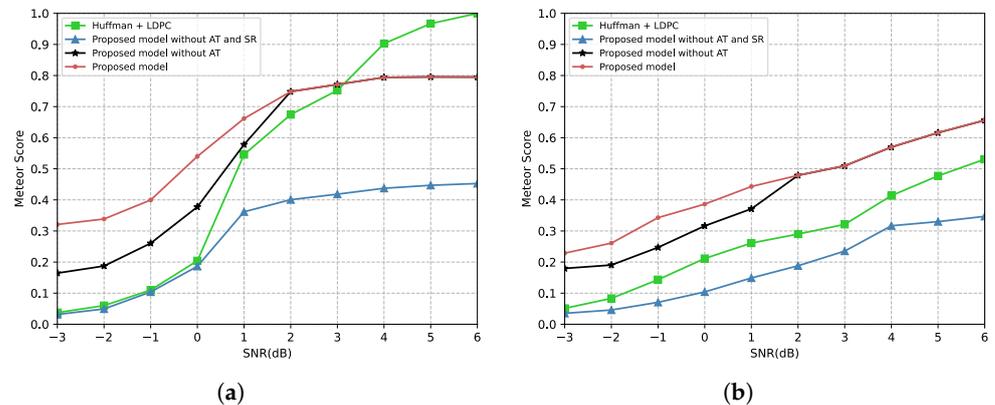


Figure 10. (a) METEOR score versus the SNR over the AWGN channel. (b) METEOR score versus the SNR over the Rayleigh fading channel.

Figure 11 draws the relationship between the SNR and the semantic similarity under the AWGN channel and the Rayleigh fading channel. From Figure 11, the “Proposed model without AT and SR” outperforms the benchmark in the low SNR region under the AWGN channel, while it outperforms the benchmark in all SNR regions under the Rayleigh fading channel. This is because our proposed model splits the transmitted sentence into multiple independent triplets, leading to that, the wrongly transmitted triplets will not affect the semantics of other triplets. However, the benchmark model transmits the sentence as a whole, and if errors occur in the transmission, then the semantics of the sentence are affected. Therefore, when the channel quality is poor, our proposed model can preserve partially correct semantics. Meanwhile, since the semantic similarity based on the BERT

model can capture semantic relationships among words, the proposed scheme obtains a higher similarity compared with the BLEU score and METEOR score.

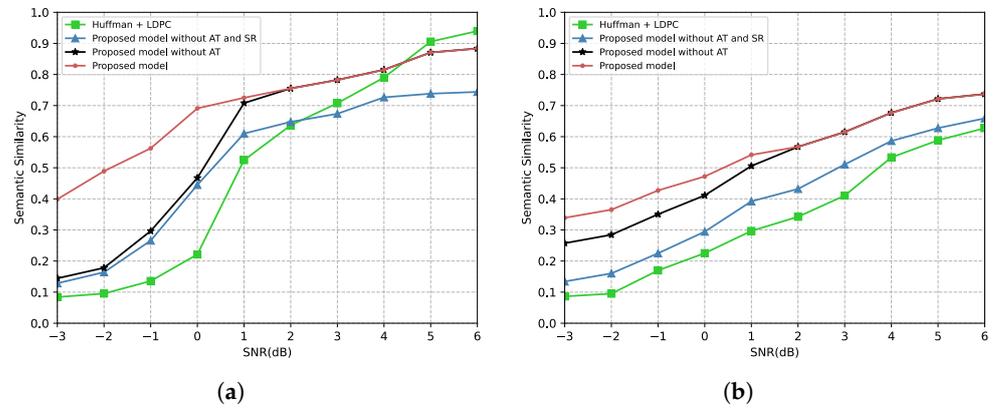


Figure 11. (a) Semantic similarity versus the SNR over the AWGN channel. (b) Semantic similarity versus the SNR over the Rayleigh fading channel.

To ensure the fairness of the comparison of experimental results, we computed the time complexities of all strategies. We transmitted 1000 sentences from the transmitter to the receiver by using different strategies and calculated the average execution time. All tests were run on Python and were performed by the computer with AMD Ryzen 7 4800H and NVIDIA GeForce GTX 3060. The results are shown in Table 4. From Table 4, our proposed model increases the computational complexity and improves communication reliability.

Table 4. The time complexity of all strategies.

Strategies	Time Complexity/s
Huffman + LDPC	2.7324
Proposed model without AT and SR	3.1638
Proposed model without AT	3.7742
Proposed model	3.8539

4.2.2. Comparison with Other Semantic Communication Models

To validate that our proposed model is more competitive than existing research, we compared it with the scheme from [23], which adopts an end-to-end deep learning-based communication system for text transmission (named DeepNN). Figure 12 shows the relationship between the SNR and the semantic similarity performance over the AWGN channel. From Figure 12, our proposed model outperforms DeepNN across the entire SNR region. The reasons are two-fold. First, by using triplets as semantic basic symbols, our proposed model can extract lossless semantics. Moreover, the important triplets are allocated more transmission resources in our proposed model, which effectively protects the importance of the semantics. However, DeepNN uses a fixed bit length to encode sentences of different lengths, resulting in a partial loss of semantics.

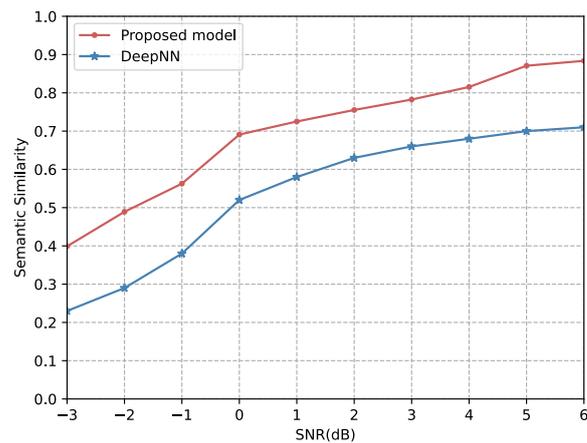


Figure 12. Semantic similarity of our proposed model and DeepNN versus the SNR over the AWGN channel.

5. Conclusions

In this paper, the reliable semantic communication assisted by the knowledge graph was studied, which overcomes the problem that the meaning of data represented by the features of the deep learning model cannot be explainable [26,28]. Specifically, we proposed a semantic extraction scheme that transforms the transmitted sentence into multiple triplets with semantic importance. Moreover, an adaptive transmission scheme is proposed, in which the important triplets are allocated more communication resources to combat channel distortion. Moreover, a semantic restoration scheme was designed to reconstruct the sentence and recover the whole semantic at the receiver. The simulation results show that the proposed system outperforms the traditional schemes in improving communication reliability, especially in the low SNR regime. However, the optimal number of triplets transmitted over a specific channel is still an ‘open question’. In the future, more work is needed to analyze the relationship between the number of triplets and the channel quality.

Author Contributions: Conceptualization, S.J.; methodology, S.J. and Y.L.; formal analysis, Y.Z. and P.L.; investigation, K.C. and J.X.; supervision, H.Z.; writing—original draft preparation, S.J. and Y.L.; 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.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

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

References

1. Sah, D.K.; Kumar, D.P.; Shivalingagowda, C.; Jayasree, P.V.Y. 5G applications and architectures. In *5G Enabled Secure Wireless Networks*; Jayakody, D., Srinivasan, K., Sharma, V., Eds.; Springer: Cham, Switzerland, 2019; pp. 45–68.
2. Zhang, Y.; Zhang, P.; Wei, J. Semantic communication for intelligent devices: Architectures and a paradigm. *Sci. Sin. Inform.* **2022**, *52*, 907–921. [\[CrossRef\]](#)
3. International Telecommunication Union. *Report on the Implementation of the Strategic Plan and the Activities of the Union for 2019–2020*; ITU: Geneva, Switzerland, 2020.
4. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [\[CrossRef\]](#)
5. Pencheva, E.; Atanasov, I.; Asenov, I. Toward network intellectualization in 6G. In *Proceedings of the 2020 XI National Conference with International Participation (ELECTRONICA)*, Sofia, Bulgaria, 23–24 July 2020; pp. 1–4.

6. Khan, L.U.; Saad, W.; Niyato, D.; Han, Z.; Hong, C.S. Digital-Twin-Enabled 6G: Vision, architectural trends, and future directions. *IEEE Commun. Mag.* **2022**, *60*, 74–80. [[CrossRef](#)]
7. Zhang, P.; Xu, W.J.; Gao, H.; Niu, K.; Xu, X.D.; Qin, X.Q.; Yuan, C.X.; Qin, Z.J.; Zhao, H.T.; Wei, J.B.; et al. Toward wisdom-evolutionary and primitive-concise 6G: A new paradigm of semantic communication networks. *Engineering* **2021**, *8*, 60–73. [[CrossRef](#)]
8. Shi, G.M.; Gao, D.H.; Song, X.D.; Chai, J.X.; Yang, M.X.; Xie, X.M.; Li, L.D.; Li, X.Y. A new communication paradigm: From bit accuracy to semantic fidelity. *arXiv* **2021**, arXiv:2101.12649.
9. Weaver, W. Recent contributions to the mathematical theory of communication. *ETC: Rev. Gen. Semant.* **1953**, *10*, 261–281.
10. Güler, B.; Yener, A.; Swami, A. The semantic communication game. *IEEE Trans. Cogn. Commun. Netw.* **2018**, *4*, 787–802. [[CrossRef](#)]
11. Popovski, P.; Simeone, O.; Boccardi, F.; Gündüz, D.; Sahin, O. Semantic-effectiveness filtering and control for post-5G wireless connectivity. *J. Indian Inst. Sci.* **2020**, *100*, 435–443. [[CrossRef](#)]
12. Weng, Z.Z.; Qin, Z.J.; Li, G.Y. Semantic communications for speech signals. In Proceedings of the ICC 2021-IEEE International Conference on Communications, Montreal, QC, Canada, 14–23 June 2021.
13. Ji, S.X.; Pan, S.R.; Cambria, E.; Marttinen, P.; Yu, P.S. A survey on knowledge graphs: Representation, acquisition and applications. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 494–514. [[CrossRef](#)]
14. Jaradeh, M.Y.; Oelen, A.; Farfar, K.E.; Prinz, M.; D’Souza, J.; Kismihók, G.; Stocker, M.; Auer, S. Open research knowledge graph: Next generation infrastructure for semantic scholarly knowledge. In Proceedings of the 10th International Conference on Knowledge Capture, CA, USA, 19–21 November 2019; pp. 243–246.
15. Atef Mosa, M. Predicting semantic categories in text based on knowledge graph combined with machine learning techniques. *Appl. Artif. Intell.* **2021**, *35*, 933–951. [[CrossRef](#)]
16. Zhou, F.H.; Li, Y.H.; Zhang, X.Y.; Wu, Q.H.; Lei, X.F.; Hu, R.Q. Cognitive semantic communication systems driven by knowledge graph. *arXiv* **2022**, arXiv:2022.11958v1.
17. Shi, G.M.; Xiao, Y.; Li, Y.Y.; Xie, X.M. From semantic communication to semantic-aware networking: Model, architecture, and open problems. *IEEE Commun. Mag.* **2021**, *59*, 44–50. [[CrossRef](#)]
18. Rudolf, C.; Bar-Hillel, Y. An outline of a theory of semantic information. *J. Symb. Log.* **1954**, *19*, 230–232.
19. Floridi, L. Outline of a theory of strongly semantic information. *Minds Mach.* **2004**, *14*, 197–221. [[CrossRef](#)]
20. Bao, J.; Basu, P.; Dean, M.; Partridge, C.; Swami, A.; Leland, W.; Hendler, J. 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.
21. 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; pp. 58–63.
22. 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.
23. 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; pp. 2326–2330.
24. 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.
25. Liu, Y.L.; Zhang, Y.Z.; Luo, P.; Jiang, S.T.; Cao, K.; Zhao, H.T.; Wei, J.B. Enhancing communication reliability from the semantic level under low SNR. *Electronics* **2022**, *11*, 1358. [[CrossRef](#)]
26. Xie, H.Q.; Qin, Z.J.; Li, G.Y.; Juang, B.H. Deep learning enabled semantic communication systems. *IEEE Trans. Signal Process.* **2021**, *69*, 2663–2675. [[CrossRef](#)]
27. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008.
28. Xie, H.Q.; Qin, Z.J. A lite distributed semantic communication system for Internet of Things. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 142–153. [[CrossRef](#)]
29. Lu, K.; Li, R.P.; Chen, X.F.; Zhao, Z.F.; Zhang, H.G. Reinforcement learning-powered semantic communication via semantic similarity. *arXiv* **2021**, arXiv:2108.12121.
30. Weng, Z.Z.; Qin, Z.J. Semantic communication systems for speech transmission. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 2434–2444. [[CrossRef](#)]
31. Hu, Q.; Zhang, G.; Qin, Z.; Cai, Y.; Yu, G. Robust semantic communications against semantic noise. *arXiv* **2022**, arXiv:2202.03338v1.
32. Xie, H.Q.; Qin, Z.J.; Li, G.Y. Task-oriented multi-user semantic communications for VQA task. *arXiv* **2021**, arXiv:2108.07357.
33. Zhou, Q.Y.; Li, R.P.; Zhao, Z.F.; Peng, C.H.; Zhang, H.G. Semantic communication with adaptive universal transformer. *IEEE Wirel. Commun. Le.* **2022**, *11*, 453–457. [[CrossRef](#)]
34. Papineni, K.; Roukos, S.; Ward, T.; Zhu, W.J. 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; pp. 311–318.
35. Banerjee, S.; Lavie, A. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In Proceedings of the Second Workshop on Statistical Machine Translation, Prague, Czech Republic, 23 June 2007; pp. 228–231.

36. Agirre, E.; Banea, C.; Cer, D.; Diab, M.; Gonzalez-Agirre, A.; Mihalcea, R.; Rigau, G.; Wiebe, J. SemEval-2016 Task 1: Semantic textual similarity, monolingual and cross-lingual evaluation. In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), San Diego, CA, USA, 16–17 June 2016; pp. 497–511.
37. Mathur, N.; Baldwin, T.; Cohn, T. Tangled up in BLEU: Reevaluating the evaluation of automatic machine translation evaluation metrics. *arXiv* **2020**, arXiv:2006.06264.
38. Kilgarriff, A.; Fellbaum, C. WordNet: An electronic lexical database. *Language* **2000**, *76*, 706–708. [[CrossRef](#)]
39. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019), Minneapolis, MN, USA, 2–7 June 2019.
40. Qi, R.; Zhang, Y.H.; Zhang, Y.H.; Bolton, J.; Manning, C.D. Stanza: A python natural language processing toolkit for many human languages. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, Online, 5–10 July 2020; pp. 101–108.
41. Greff, K.; Srivastava, R.K.; Koutník, J.; Steunebrink, B.R.; Schmidhuber, J. LSTM: A search space odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 2222–2232. [[CrossRef](#)]
42. Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. In Proceedings of the 6th International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
43. Huffman, D.A. A method for the construction of minimum-redundancy codes. *Proc. IRE* **1952**, *40*, 1098–1101. [[CrossRef](#)]
44. Gallager, R. Low-density parity-check codes. *IRE Trans. Inf. Theory* **1962**, *8*, 21–28. [[CrossRef](#)]
45. Gardent, C.; Shimorina, A.; Narayan, S.; Perez-Beltrachini, L. Creating training corpora for NLG micro-planners. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Vancouver, BC, Canada, 30 July–4 August 2017; pp. 179–188.