

Retraction

Retracted: A Study of English Translation Theory Based on Multivariate Statistical Analysis of Random Matrix

Mathematical Problems in Engineering

Received 13 September 2023; Accepted 13 September 2023; Published 14 September 2023

Copyright © 2023 Mathematical Problems in Engineering. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] Y. Li, "A Study of English Translation Theory Based on Multivariate Statistical Analysis of Random Matrix," *Mathematical Problems in Engineering*, vol. 2022, Article ID 3077453, 6 pages, 2022.

Research Article

A Study of English Translation Theory Based on Multivariate Statistical Analysis of Random Matrix

Yongfang Li 

Department of General Education, Deyang Vocational College of Technology and Trade, Guanghan, Sichuan 618300, China

Correspondence should be addressed to Yongfang Li; li_yf@scmc.edu.cn

Received 11 May 2022; Revised 28 June 2022; Accepted 4 July 2022; Published 31 July 2022

Academic Editor: Ning Cao

Copyright © 2022 Yongfang Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the continuous advancement of artificial intelligence in natural language processing technology, machine translation based on machine learning technology has been fully transformed from traditional machine translation methods to neural network machine translation methods. In particular, the tremendous development of large families has made data-driven a reality. With deep learning as the research and design background, the neural network structure of random matrix multivariate statistical analysis is designed according to the language characteristics of English. The model was tested on a Chinese-English panning model, and the optimal model fused with a Bleu value of 39.53. The model results were applied to a real system to achieve language detection, multidirectional language translation, and manual correction of results to be able to learn long dependencies and overcome the limitations of recurrent neural networks to translate long sentences more fluently.

1. Introduction

The basic model of machine translation system is a natural language processing system, and its basic principle is the principle of element synthesis [1–3]. The process of machine translation can be simply understood as three stages: first, the source language text is decomposed into basic constituents such as words, phrases, and grammatical structures, and the original text is analyzed. Morphological analysis, syntactic analysis, and semantic analysis are used to form the internal representation of the source language, and then translation using the target language eventually produces the internal representation of the source language by transforming and ordering at the structural level using the compound rules of the target language [4].

Therefore, to improve the accuracy of machine translation, the processing of natural language is taken as the core problem of machine translation system research. The rapid development of artificial intelligence technology in natural language processing technology and the rapid development of the Internet, has led to a boom in machine translation research, which has become a hot topic not only in the field of scientific and technological research, but also in the keen research of linguists all over the world [5, 6].

Although the statistical machine translation system is also developing and improving, its statistical machine translation which ignores the grammar rules is highly dependent on the large-scale corpus and relies on statistical data for disambiguation and translation selection, while the translation effect ultimately depends on the probabilistic model and the coverage capacity of the corpus, [7] so the probabilistic matching results of words and phrases do not achieve the accuracy of the final translated sentences and the translation effect is poor. In order to improve the translation effect, neural network technology was introduced into machine translation [8]. A neural network is an artificial system for intelligent information processing developed by exploring models that simulate the function of the human brain's nervous system with functions such as learning, association, memory, and pattern recognition [9].

Research has shown that although recurrent neural networks are good at using global information, due to the flexibility, complexity, and diversity of natural language, they are unable to remember content that is too far ahead or too far behind, i.e., the interval between relevant information and the current predicted position is widened and recurrent neural networks become unable to learn how to connect information, a problem known as “long-term dependency.”

This is known as the “long-term dependency” problem [10]. [11] Long Short Term Memory Neural Network (LSTM), a variant of recurrent neural network, is able to learn long dependencies and overcome the limitations of recurrent neural networks, resulting in more fluent translation of long sentences. The attention mechanism (AM)-based neural machine translation can dynamically obtain the source language word information related to the generated words during decoding, and obtain word alignment information [12]. [13] For example, when generating “I”, the word “I” is dynamically calculated to be the most relevant, rather than the other words. Neural machine translation is rapidly replacing statistical machine translation as the mainstream machine translation technology in academia and industry [14].

2. Neural Network Structure for Chinese-English Translation

Unlike traditional statistical machine translation, which uses multiple modules to complete the translation task, neural machine translation uses an encoder-decoder architecture [15–19]. The encoder-decoder architecture was first proposed by Kalchbrenner and Blunsom at the University of Oxford in 2013.

2.1. Basic Network Structure of Chinese-English Translation.

The encoder-decoder model is that encoder converts the input language sequence into a fixed length intermediate vector [20, 21]. Figure 1 shows the overall structure of the encoder-decoder model for translating from Chinese to English. The decoder then obtains the output Chinese-English sequence based on the fixed vector C .

Among them, the length of the parallel utterances in the training set is not unique and the words in the utterances will have sequence information; a recurrent neural network (RNN) is used in order to obtain a robust translation system; the structure of which is shown in Figure 2.

As you can see from the expansion diagram, the current moment is influenced by the previous moment. This feature is very suitable for language sequences, because no matter if it is Chinese-English, English, or linguistic, the meaning of the words in the sentence will depend on the context. The RNN not only reads the sequence information from x_0 to x_n and not only reads the sequence information from front to back, but also from back to front, i.e., from x_n to read information, so as to ensure the integrity of the input sequence information.

The backpropagation process can lead to gradient explosion or gradient disappearance, which can cause oscillation in the parameter update of the model learning, while gradient disappearance can make the learning very slow and lead to ineffective learning in the end.

2.2. Improvement of the Underlying Network Structure. This is to improve the translation effect. The improved model structure is shown in Figure 3.

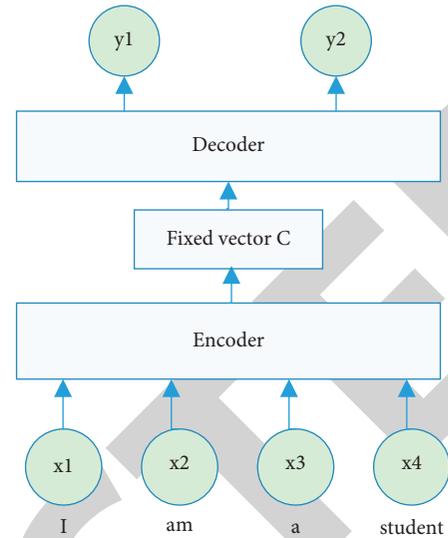


FIGURE 1: Encoder-decoder model structure.

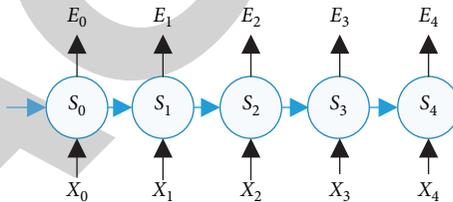


FIGURE 2: RNN structure unfolding.

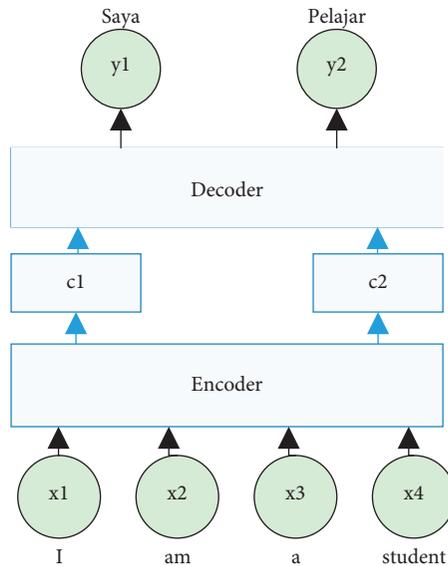


FIGURE 3: Model structure for generating dynamic vectors.

In the case of the translation of English I am a student into Chinese Saya pelajar, when parsing the Chinese English word, so that after the learning of the neural network, it will generate the vector C_2 .

As shown in Figure 4, the stronger the relationship between the source language words corresponding to each target word, i.e., the higher the weight, the darker the colour

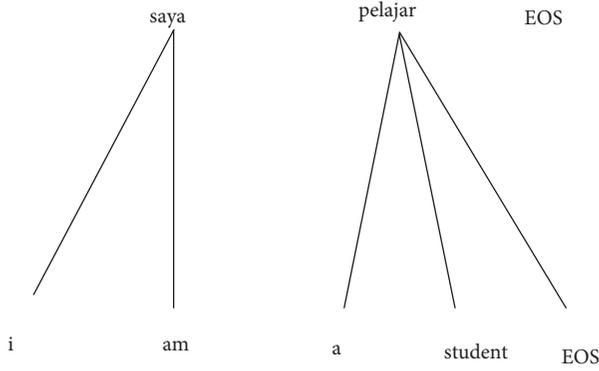


FIGURE 4: Schematic representation of the source and target language sequence.

of the connecting line. The EOS represents the end-of-sentence marker for each utterance, and also calculates the weight, because some words end when the sentence is read, and need to be fed back to the neural network to tell the model which word is going to end when the target language is generated.

The body dynamic vector C_i , is calculated as follows:

$$C_i = \sum_{j=1}^{T_x} a_{ij} h_j, \quad (1)$$

where T_x represents the length of the input sentence, i.e., the length of the English sentence, plus the stop character, T_x here being 5.

$h_1 = f(I), h_2 = f(am), h_3 = f(a), h_4 = f(student), h_5 = f(\langle EOS \rangle)$. Here, the f function represents encoder's transformation function for the input English word. a_{ij} is the weight value, which is continuously optimized during the learning process of the neural network.

Therefore,

$c_1 = g(a_{11} * h_1, a_{12} * h_2, a_{13} * h_3, a_{14} * h_4, a_{15} * h_5)$, and the g function is the weighted summation.

2.3. Improvement of the Decoder Hidden Layer Unit. Since the introduction of mathematical methods that simulate actual human neural networks, people have slowly got used to referring to such artificial neural networks directly as neural networks. Neural networks have a wide and attractive prospect in the fields of system identification, pattern recognition, and intelligent control, etc. Especially in intelligent control, people are particularly interested in the self-learning function of neural networks and regard this important feature of neural networks as one of the key keys to solve the difficult problem of controller adaptability in automatic control. The sequence information is read from front to back, and from back to front, i.e., from reading information, thus ensuring the integrity of the input sequence information.

We use the encoder-decoder neural network structure to conduct experiments on Chinese-English machine translation. After the source and target language texts are pre-trained with word vectors.

The specific structure of the co-decoder's implicit unit is shown in Figure 5.

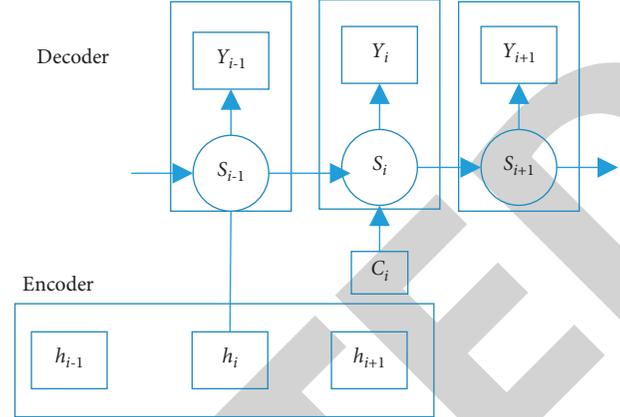


FIGURE 5: Encoder-decoder network hidden layer cell structure.

In the Chinese-English translation experiments in the computation of decoder decoding, an improvement was made to the hidden layer unit in it: that is, the GRU module is composed of the prehidden layer state S_{t-i} [22, 23].

3. Structural Design of Chinese-English Translations

Based on the above analysis, an integrated translation system can be designed, which is based on file configuration and can complete the whole translation system. The architecture of the Chinese-English machine translation system is shown in Figure 6.

Before training starts, a parallel training corpus is loaded and the model is pretrained with word vectors to obtain word vectors for the source and target languages, respectively [24, 25].

4. Google Translate Example Analysis

Language is flexible and diverse, and many words and expressions are related to specific contexts and can sometimes be difficult to understand. The first two words "how much" and the last two words "how much" mean exactly the opposite. In the following, I will analyse the application of Google's neural network translation in technical English with examples.

Example 1. Where the separation of the two light components and the two recombinant components takes place, column means column; but in the context of chemical texts, this refers specifically to the distillation column, where the components are separated by distillation, and cannot be translated as column. Compared to human translators, the AI translation results are not only readable and fluent when dealing with complex, specialised or technical passages, but also still intolerably flawed in terms of correctness as a basic requirement [26, 27].

Example 2. A Chinese technology company with 80,000 employees is banned by the Trump administration from dealing with American firms. Google translation follows the

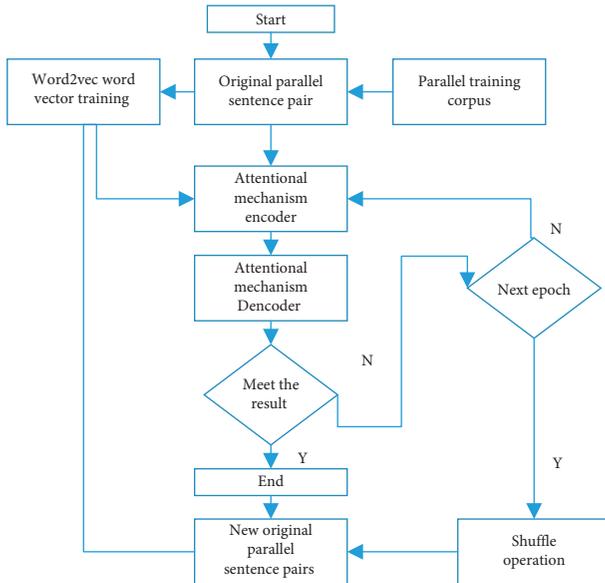


FIGURE 6: Architecture of the translation model.

original language directly and uses passive sentences, which is not difficult for the target readers to understand, but the manual translation changes passive to active, which makes the expression more in line with the habits of the target audience and more authentic.

Human translation: genetically modified plants can be cultivated to possess enhanced stress behaviour. Stress behavior is a dynamic response completed in a short time, not an adaptive inertia behavior. The source text uses the passive voice, but the Google translation avoids the rigidity of the expression by not using the word “to be,” but by translating the subsequent purpose clause “to possess improved stress behaviour.” This is not clear what is meant by “to possess improved stress behaviour,” which creates ambiguity and leads to confusion in understanding. The translation in the human translation adds “its” to the object of reference, translating it as to have improved stress behaviour, which means the transgenic plant, avoiding ambiguity [25].

Example 3. Momentum is building for American regulators to catch up with Europe in promoting “biosimilars,” which are generic approximations of patented drugs. Google Translate: US regulators are building for American regulators to catch up with Europe in promoting “biosimilars,” which are generic approximations of patented drugs.

Human translation: US regulators are catching up with Europe in the promotion of generic approximations of patented drugs, “biosimilars.” Momentum is variously understood to mean, momentum, impetus, etc., and this case is meant to indicate the efforts made by US regulators, so momentum is more appropriate. The translation of approximations in this sentence is inaccurate. Google does not affect the understanding of the meaning of the original text too much, but there is a certain difference with the Chinese language expression. However, there is a certain gap with the Chinese language, especially when translating long and difficult sentences, and due to technical limitations, the

TABLE 1: Results of Chinese-English translation of Chinese seed 1.

Seed1	94W	103W	109W	117W
Dev_BLEU	35.72	35.99	35.88	35.75

TABLE 2: Results of English translation of Chinese seed 2.

Seed2	82W	88W	89W	95W
Dev_BLEU	36.72	35.91	36.88	35.96

translation may only achieve formal equivalence and basic sentence fluency.

Example 4. Science has cracked the genetic information code. Google Translate: Science has cracked the genetic information code.

Human translation: Science has cracked the genetic information code. Code has multiple semantic meanings, such as code, password, encoding, and codex. The deciphering of the genetic code in biology is a great milestone in the history of biology; genetic engineering technology is not a code, and machine translation cannot identify the exact meaning of words in a particular application context.

5. Tests on the Horizontal Model of Chinese-English Translation

5.1. Test of Chinese-English and Indo-English Translation Models. In the experiments, two seeds were selected for model training in each translation direction, and then the model with the better round was selected for ensemble, i.e. fusion model, at the end of training, and the specific experimental results are shown in Tables 1, 2, and 3.

- (1) Model test results of the better round of two seeds
- (2) The result of the better model after model fusion, with the best value in bold.

The table shows that the BLEU of the optimal model after fusion can reach 37.99.

- (1) The model test results for the two seeded better rounds are shown in Tables 4 and 5.
- (2) The results of the better model after model fusion are shown in Table 6.

The table shows that the two models are equally effective, with a BLEU of 34.44..

- (1) The results of the model testing for the two seeded better rounds are shown in Tables 7 and 8.
- (2) The results of the better model after model fusion are shown in Table 9.

The table shows that the optimal model after fusion has a high BLEU of 39.53.

5.2. Competitive Translation Reviews. If the translation model is not available in the corresponding language direction, it cannot be evaluated, and it is necessary to check

TABLE 3: Results of the fusion of Chinese-English translation models.

1d2cn_ensemble	94w_103w_82w_88w	94w_103w_82w_95w	94w_117w_82w_88w	94w_103w_82w_95w
Dev_BLEU	37.82	38.01	37.99	37.99

TABLE 4: Results of Chinese translation of seed 1.

Seed1	166W	170W	172W	174W	187W
Dev_BLEU	32.01	31.61	31.49	31.30	31.43

TABLE 5: Results of Chinese translation of Chinese-English seed 2.

Seed2	135W	136W	139W	144W	153W
Dev_BLEU	31.56	31.59	31.49	31.30	31.43

TABLE 6: Results after fusion of translated Chinese-English models.

cn2id_ensemble	166W_172W_139w_153w	172W_187W_135w_144w
Dev_BLEU	34.44	34.44

TABLE 7: Results of Chinese-English translation of English seed 1.

Seed1	120W	124W	126W	127W
Dev_BLEU	36.41	36.69	36.70	37.14

TABLE 8: Results of Chinese-English translation of English seed 2.

Seed2	101W	102W	103W	108W	110W	124W
Dev_BLEU	35.86	36.22	36.20	36.5	36.17	36.13

TABLE 9: Results of the fused Chinese-English translation English model.

cn2id_ensemble	120W_124W_110w_108w	120W_124W_102w_110w	120W_126W_110w_103w	120W_124W_102w_108w
Dev_BLEU	39.42	39.42	39.53	39.45

TABLE 10: Competitive review data for Chinese-to-English translation of English.

En2id	Bing	Google	4model_ensemble
Dev	28.50	33.29	39.53

TABLE 11: Competitive review data for English translation from Chinese to English.

En2id	Bing	Google	4model_ensemble
Dev	25.14	28.27	32.19

whether the language pair exists on the website of the corresponding company in advance.

The evaluation method is to use our test data to crawl the online translation results of Baidu, Microsoft, and Google on the cluster, and then perform BLEU evaluation on the crawled data. The evaluation data from the experiment are shown below. Only the data of English and Indonesian translation are available here, while the online translation results of Chinese and Indian translation are not available for evaluation, as shown in Tables 10 and 11.

This project is embedded in the system, and after the model has been tested, the results of the competing translations are calculated and then saved as one file output in the same language translation direction. The data obtained are also much clearer and more convincing.

6. Conclusions

Artificial intelligence has greatly facilitated the development of translation services, allowing people to quickly understand the general content of languages other than their native language without having to rely to some extent on human translation. The whole training and testing results are very simple, just start loading our collated bilingual material cat algorithm, run the script and then no longer need excessive human cost to test the BLEU value of each model in real time, and finally the results can be obtained. However, due to the differences between the source and target languages, it is found that the translation results of the above-mentioned technical English translation examples have some semantic deviations from the original sentences, and there is a gap with the human translation, and machine translation still cannot achieve perfect translation. Therefore, for the time being,

machine translation has some application value, but it cannot completely replace human translation yet.

Data Availability

The dataset used in this paper are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

- [1] C. Huo, X. Du, and W. Gu, "A study of English translation of literature on diet in taihang mountainous areas from the perspective of skopos theory of translation," *Creative Education*, vol. 11, no. 07, pp. 979–984, 2020.
- [2] W. J. Zhao, "A practical study on college English translation teaching based on cat technology," *Journal of Physics: Conference Series*, vol. 1533, no. 2, Article ID 022062, 2020.
- [3] T. Zhao and Y. Cai, "Improvement of English key competences based on machine learning and artificial intelligence technology," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2069–2081, 2021.
- [4] X. I. E. Tao, C. Zhang, and Y. Xu, "Collaborative parameter update based on average variance reduction of historical gradients [J]," *Journal of Electronics and Information Technology*, vol. 43, no. 4, pp. 956–964, 2021.
- [5] Y. . Shi, "Study on intelligent construction of English-Chinese bilingual context model based on CBR," *Journal of Intelligent and Fuzzy Systems*, vol. 2021, no. 1, pp. 1–8, 2021.
- [6] A. M. Bogush, T. M. Korolova, and O. V. Popova, "Teaching machine translation to the students majoring in the humanities [j]," *Information Technologies and Learning Tools*, vol. 71, no. 3, p. 122, 2019.
- [7] S. . Walde, "Human associations and the choice of features for semantic verb classification [J]," *Research on Language and Computation*, vol. 6, no. 1, pp. 79–111, 2008.
- [8] G. Stevens, P. Monachesi, and A. Bosch, "A pilot study for automatic semantic role labeling in a Dutch[J]," *Lot Occasional*, vol. 7, pp. 99–114, 2007.
- [9] X. Wang, Y. Kawabe, R. Kato et al., *Journal of Bioscience and Bioengineering*, vol. 124, no. 5, pp. 583–590, 2017.
- [10] Z. W. Zhang, D. Wu, and C. J. Zhang, "Study of cellular traffic prediction based on multi-channel sparse LSTM [J]," *Computer Science*, vol. 48, no. 6, pp. 296–300, 2021.
- [11] B. K. Kang and J. E. Lee, "The operating principles of neural machine translation and the accuracy of translation - focusing on the Chinese-Korean translation," *Korea Journal of Chinese Language and Literature*, vol. 73, pp. 253–295, 2018.
- [12] S. Zhang, M. W. Morris, C. Y. Cheng, and A. J. Yap, "Heritage-culture images disrupt immigrants' second-language processing through triggering first-language interference," *Proceedings of the National Academy of Sciences*, vol. 110, no. 28, pp. 11272–11277, 2013.
- [13] X. Fan and M. Sun, "A method of recognizing entity and relation," in *Proceedings of the International Conference on Natural Language Processing*, Jeju Island, Korea, October 2005.
- [14] J. Brooke, A. Hammond, and G. Hirst, "Using models of lexical style to quantify free indirect discourse in modernist fiction [J]," *Literary and Linguistic Computing*, vol. 32, no. 2, pp. 234–250, 2017.
- [15] P. An, Z. Wang, and C. Zhang, "Ensemble unsupervised autoencoders and Gaussian mixture model for cyberattack detection," *Information Processing & Management*, vol. 59, no. 2, Article ID 102844, 2022.
- [16] H. Singh and I. Daly, "Translational algorithms: the heart of a brain computer interface [J]," *Intelligent Systems Reference Library*, vol. 74, pp. 97–121, 2015.
- [17] S. A. Zonouz, R. Berthier, H. Khurana, W. H. Sanders, and T. Yardley, "Seclius: an information flow-based, consequence-centric security metric," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 2, pp. 562–573, 2015.
- [18] M. Cordella and S. Musgrave, "Oral communication skills of international medical graduates: assessing empathy in discourse," *Communication and Medicine*, vol. 6, no. 2, pp. 129–142, 2009.
- [19] M. I. Tiwana, S. J. Redmond, and N. H. Lovell, "A review of tactile sensing technologies with applications in biomedical engineering," *Sensors and Actuators A: Physical*, vol. 179, no. none, pp. 17–31, 2012.
- [20] D. Jiang, F. Wang, Z. Lv et al., "QoE-aware efficient content distribution scheme for satellite-terrestrial networks," *IEEE Transactions on Mobile Computing*, vol. 2021, p. 1, 2021.
- [21] G. Cai, Y. Fang, J. Wen, S. Mumtaz, Y. Song, and V. Frascolla, "Multi-carrier M-ary DCSK system with code index modulation: an efficient solution for chaotic communications," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 6, pp. 1375–1386, 2019.
- [22] K. Chandra, A. S. Marcano, S. Mumtaz, R. V. Prasad, and H. L. Christiansen, "Unveiling capacity gains in ultradense networks: using mm-wave NOMA," *IEEE Vehicular Technology Magazine*, vol. 13, no. 2, pp. 75–83, June 2018.
- [23] J. Du, C. Jiang, Z. Han, H. Zhang, S. Mumtaz, and Y. Ren, "Contract mechanism and performance analysis for data transaction in mobile social networks," *IEEE Transactions on Network Science and Engineering*, vol. 6, no. 2, pp. 103–115, 1 April–June 2019.
- [24] T. Yang, M. Mahdavi, R. Jin, and S. Zhu, "An efficient primal dual prox method for non-smooth optimization," *Machine Learning*, vol. 98, no. 3, pp. 369–406, 2015.
- [25] I. Porto-Díaz, D. Martínez-Rego, A. Alonso-Betanzos, and O. Fontenla-Romero, "Information Theoretic Learning and local modeling for binary and multiclass classification," *Progress in Artificial Intelligence*, vol. 1, no. 4, pp. 315–328, 2012.
- [26] C. R. Shalizi and J. P. Crutchfield, "Computational mechanics: pattern and prediction, structure and simplicity [J]," *Journal of Statistical Physics*, vol. 104, no. 3/4, pp. 817–879, 2001.
- [27] Q. Wang and H. Wang, "Deep learning model-based machine learning for Chinese and Japanese translation," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–8, Article ID 8762125, 2022.