International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 398-405

https://doi.org/10.70517/ijhsa46435

Exploring the Potential Impact of New Media Language on Chinese Semantic Change Based on Natural Language Processing Algorithms

Kangxin Yuan[1,*]
[1] College of Arts and Law, China University of Petroleum (East China), Qingdao, Shandong, 266580, China
Corresponding authors: (e-mail: zgsydxykx@163.com).

Abstract This study focuses on the potential impact of new media language on the semantic evolution of Chinese, constructs the Chinese Lexical Semantic Core Knowledge Base (CLSKB_Core), which integrates the resources of multi-source dictionaries, and puts forward a semantic similarity computation model based on conceptual graphs. The semantic evolution of Chinese in the new media context is explored through the analysis of ephemeral frequency and the validation of the co-conceptual map retrieval algorithm. The experiments show that the algorithm recognizes sentence types with simple structure such as definite middle, preposition, number, gerund, etc. with better efficiency, and the average correctness rate reaches 0.899. In eight types of problems, the average Fβ value reaches 67.025%. For concept maps with imbalanced attribute descriptions, the algorithm in this paper achieves an average compatibility of 0.886, which is 0.6% higher than the original algorithm. The semantic migration trajectory of the typical case “Tuhao” reveals the penetration mechanism of Internet buzzwords into the semantic field of traditional vocabulary, and reflects the feasibility of adopting natural language processing algorithms to study the semantic change of Chinese.

Index Terms natural language processing algorithm, concept map, semantic similarity, new media language, Chinese semantic change

I. Introduction

In the long history of China's development, many cultures and civilizations have been created one after another, including the Chinese language, and it has been used until now, occupying a very important position in Chinese culture [1], [2]. With the arrival of informationization reform, new media language is gradually prevalent, in this context, the change of Chinese language has ushered in a new challenge, and the birth of new media language is a double-edged sword to the change of Chinese semantics [3]-[5]. Therefore, how to correctly view the new media language, standardize its usage and promote its better development has become an important issue at present [6].

First of all, the development of new media has a positive impact on the development of Chinese language in China, especially the new media language, which can be considered as a very important information carrier [7], [8]. No matter which language should be as concise and powerful as possible, so that both parties can quickly understand each other's meaning, in the new media era, people's daily life is very fast-paced, the language should be concise and clear, and the new media language just meets all these requirements [9], [10]. Internet language can break the limitations of traditional Chinese semantics, both in terms of the overall structure and the network language, as far as the form of broad expression is concerned [11], [12]. This means that in practice, the language can be applied more directly to improve the efficiency and quality of text input and communication, which not only meets the needs of modern people, but also has the potential to effectively promote the development of the Chinese language [13]-[15]. However, new media language is not entirely favorable to the healthy development of the Chinese language [16]. Although the meaning of words can be expressed through new media language without seriously affecting the communication between people [17], [18]. Grammatical and spelling errors are very common in online languages, and these erroneous habits will further breed and exacerbate the stupidity, frivolity, and vulgarization of Chinese semantics, which is detrimental to the high level of development of the Chinese language [19]-[21].

In this paper, we integrate the modern Chinese grammar information dictionary, the modern Chinese dummy word usage knowledge base and the modern Chinese dictionary to construct a large-scale Chinese lexical semantic knowledge base CLSKB_Core. We propose a conceptual graph matching and inference algorithm suitable for the Chinese semantic computation, and emphasize the characteristics of the similarity of the conceptual graph. The similarity of the whole conceptual graph is obtained by calculating the similarity of nodes and arcs in the conceptual graph. The influence of new media on Chinese expressions is sorted out to explore the language changes in the new media environment. Eight common sentence patterns in Chinese are tested, and the performance level of the algorithm is evaluated in conjunction with the retrieval check rate. A control experiment is set up to examine the effectiveness of the improvement strategy in this paper. A case study is conducted to verify the semantic evolution effect of new media language based on the ephemeral corpus.

II. Semantic analysis of Chinese vocabulary based on concept maps

II. A. Construction of the core knowledge base CLSKB_Core

Lexical semantic knowledge base is an important resource in natural language processing, and its scale and quality largely determine the success or failure of natural language processing system. The completion of the construction of large-scale lexical semantic knowledge base is of great significance to the current research fields in natural language processing such as information retrieval, syntactic analysis, machine translation and so on.

Considering the information summarization of credible vocabulary volume, credibility and characteristics, the Modern Chinese Grammar Information Dictionary, Modern Chinese Dictionary and Modern Chinese Vocabulary Usage Knowledge Base are used as the resources for constructing CLSKB_Core.

The main work of constructing CLSKB_Core is to merge these three resources. CFKB, GKB and MCD5 are selected as the core libraries, from which the contents of fields such as words, lexical properties, full pinyin, denotations, pragmatics and example sentences are extracted. The corresponding synonyms, superlatives, associates and antonyms of the synonym fields are extracted from CLhit, CB and YYCD. Since different resources may contain the same words and the same terms, the project team developed an automatic word merge tool to automatically merge similar and identical terms from different resources, and then manually assisted to check the results of the automatic merge. The specific processing flow for building CLSKB_Core is shown in Figure 1.

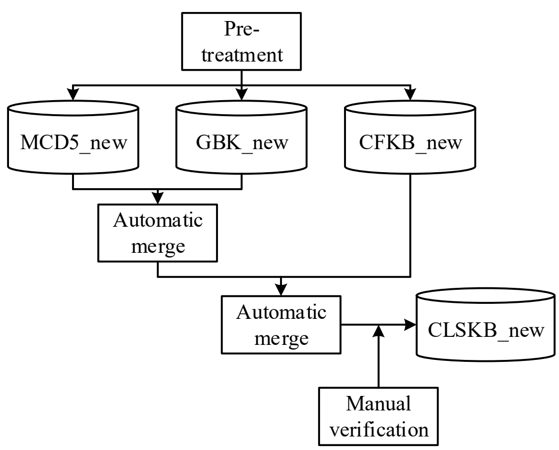


Figure 1: The construction process of the core knowledge base of lexical semantics

(1) Preprocessing. Convert the storage structure of the three to a table structure. For field attributes that exist in the original resource, migrate their values first to the new table. For fields that do not exist in the original resource but are required by the lexical semantic knowledge base table, their values are kept blank. The three core libraries after preprocessing are labeled as MCD5_new, GKB_new, and CFK_new, respectively.

(2) Automatic merging of knowledge bases. Since the same words are described differently in different resources with different focuses, and this information is not one-to-one correspondence, the word automatic merging algorithm is used to merge the same words in different resources. In two different resources, if there are two records with the same words and the same lexical properties, then it is considered that they represent the same word. On this basis, the similarity of the two terms is examined, and if the two terms are similar, the fields of pinyin, lexical properties and remarks are merged.

Firstly, MCD5_new and GKB_new are merged, because the entries from MCD5_new have more detailed explanations of the meanings of real words, so MCD5 is the main source, and the entries from GKB are used to supplement the entries and meanings of MCD5. After the merger is completed, the result is merged with CFKB, which is the most complete knowledge base for the interpretation and usage of dummy words. Based on this, CFKB is used to supplement the intermediate result of the merger of MCD5 and GKB to obtain the lexical semantic core knowledge base Cl. SKB_Core. The implementation of the automatic word merge algorithm is now introduced using MCD5 and GKB as examples.

For the word "hit", the record with ID 8995 in the new table is marked with "[CLhit]" before the serial number <1> in the "synonymous" column, indicating that the synonym is merged from CLhit.

There are 38002 words in the merged knowledge base with the number of 1, which will not make any error during the merging process and do not need to be modified manually, and these words account for 30.5% of CLSKB_Core. Besides, there are 80,351 polysemous words in total, and this kind of words may make mistakes during the automatic merging by machine because they contain more than one sense, so they need to be checked manually.

(3) Manual checking. Although the initial merging of resources has been completed by the machine, since the automatic merging tool cannot guarantee that the merging result is completely correct, it is still necessary to manually assist in checking the automatic merging result, which inevitably introduces a large number of manual operations. Therefore, we have developed the Trusted Resource Merger Tool, which is used to browse and check the automatic merging results manually after the automatic merging of entries, confirming the correct merging results and correcting the incorrect merging results, which is capable of directly modifying the database, simple to operate, and convenient for the merging of resources.

In the process of manual merging using the Trusted Resource Merge Tool, the following principles were followed: 1) CFKB was used as the main source of lexical entries for imaginary words, and MCD5 was used to supplement them. 2) MCD5 was used as the main source of lexical entries for real words, and GKB was used to supplement them. 3) When it was found that the values in the fields were not taken in a standardized manner, such as those that contained lexical properties in the paraphrase, they should be removed. 4) The field value of the “synonym” field contains <1> synonyms; <2> superlatives; <3> associates; and <4> antonyms.

II. B.Retrieval Algorithms for Concept Maps

II. B. 1) Semantic similarity of concept maps

In terms of concept map retrieval and similarity computation, this paper improves the existing algorithms to make them more suitable for Chinese semantic computation. There are two main types of existing studies on matching concept maps, one is the exact matching between two concept maps, and the other is to calculate the similarity between two concept maps. This paper is mainly a study of the latter.

Semantic matching is mainly to solve the problems caused by synonyms and homonyms, when matching, not only the structural similarity of the two should be considered, but also their relationship should be considered from the semantic point of view, if the semantics can't be matched exactly, the semantic proximity of the two has to be compared, so it is expressed by the semantic similarity.

In this paper, it is elaborated that the semantic similarity between concepts is obtained through the semantic distance of two concepts, and then the similarity of each node and relationship is accumulated by weights to finally arrive at the similarity between two graphs. In order to avoid recursion into an endless loop during computation, it is stipulated that the user specifies an entry node for querying the conceptual graph, and the existing searched conceptual graph also has an entry node, which only compares the similarity of concepts that are equally located in the two conceptual graphs. The entry node, here, is the first node in the linearized representation of the graph, and this knowledge is described in detail in the linearization techniques for graphs. Here is the formula for calculating the similarity of a graph.

$$SoG(nQ,nR)=W(nQ,n)\cdot sim_n(nQ,nR) + \max_{for\ all\ combination}\left\{\left(\sum_j w(nQ,j)\right)\cdot sim_a(a_Q^j,a_R^j)\cdot SoG(n_Q^{a_Q^j},n_R^{a_Q^j})\right\} \tag{1}$$

$SoG(nQ,nR)$ is used to compute the similarity between two graphs, where n_Q denotes the entry node of the query conceptual graph and n_R denotes the entry node of the resource conceptual graph. $sim_n(n_Q,n_R)$ denotes the similarity between the two nodes, and $sima(a_Q^j,a_R^j)$ computes the similarity between the j th arc from nodes n_Q and n_R, where $W(n_Q,n)+\sum_j w(n_{Q,j})=1$ means that the sum of the weights of the individual arcs departing from a node and that node is 1, and that the weights of the individual arcs and nodes are equal.

II. B. 2) Similarity of nodes

The nodes of the conceptual graph built up through the above steps should be called a word rather than a concept in the real sense. So in this paper, in the link of calculating node similarity, we consider more about the similarity calculation of words.

There are two types of common calculation methods for word similarity, one is based on some kind of world knowledge, and the other is to utilize a large-scale corpus for statistics, in this paper, we mainly use the word similarity calculation method based on CLSKB_Core. Generally, a synonym dictionary is used to organize all the words in one or several tree-like hierarchical structures, and there is one and only one path between any two nodes. The length of this path is then used as a measure of the semantic distance between these two concepts. Another method of calculating word similarity is large-scale corpus to count, which is not studied in this paper.

For two Chinese words W_1 and W_2, if W_1 has n denotations (concepts): $S_{11}, S_{12}, \ldots\ldots, S_{1n}$, W_2 has m denotations (concepts): $S_{21}, S_{22}, \ldots\ldots, S_{2m}$, and we stipulate that the similarity of W_1 and W_2 is the maximum of the similarity of the individual concepts, that is:

$$Sim(W_1, W_2) = \max_{i=1,\ldots,n,\, j=1,\ldots,m} Sim(S_{1i}, S_{2j}) \tag{2}$$

In this way, we reduce the similarity problem between two words to the similarity problem between two concepts. Unlike traditional semantic dictionaries, in CLSKB_Core, instead of corresponding each concept to a node in a tree-like hierarchy of concepts, a concept is described by means of a series of justification principles, using some kind of knowledge description language. Our goal is to find a way to compute the similarity between two semantic expressions represented in this knowledge description language.

Since all the justification principles form a tree-like hierarchy of justification principles based on the contextual relationship, the simple approach of calculating similarity by semantic distance is used here. Assume that the path distance between two justification principles in this hierarchy is d and the semantic distance between these two justification principles is:

$$Sim(p_1, p_2) = \frac{\alpha}{d + \alpha} \tag{3}$$

where p_1 and p_2 denote the two righteousness primitives, and d is the path length of p_1 and p_2 in the hierarchy of the righteousness primitives, a positive integer. a is an adjustable parameter.

Since the concepts are described by a semantic expression, and the size of the support information that individual justification principles can provide in the expression varies, this paper adopts the method of weighting the classification of justification principles to reflect the similarity of the words more accurately.

For the semantic expressions of word concepts, we divide them into four parts:

(1) The first independent semantic origin descriptor: we notate the similarity of this part of two concepts as $Sim_1(S_1, S_2)$.

(2) Other independent justifier descriptor: all other independent justifiers (or specific words) in the semantic expression except the first independent justifier, we denote the similarity of this part of the two concepts as $Sim_2(S_1, S_2)$.

(3) Relational proto-descriptor: all of the semantic expressions are described by the relational proto-descriptor, and we notate the similarity of this part of the two concepts as $Sim_3(S_1, S_2)$.

(4) Symbolic proto-descriptor: all of the semantic expressions are described by symbolic proto-descriptors, and we write down the similarity of this part of the two concepts as $Sim_4(S_1, S_2)$.

So the total similarity formula can then be found by weighting the individual parts:

$$Sim(S_1, S_2) = \sum_{i=1}^{4} \beta_i \prod_{j=1}^{i} Sim_j(S_1, S_2) \tag{4}$$

where $\beta_i (1 \le i \le 4)$ is an adjustable parameter and there are: $\beta_1 + \beta_2 + \beta_3 + \beta_4 = 1$, $\beta_1 \ge \beta_2 \ge \beta_3 \ge \beta_4$. The latter reflects the decreasing role played by Sim_1 to Sim_4 for the overall similarity. Since the first independent sense of the original descriptor reflects the most important features of a concept, we should define its weights to be relatively large, which should generally be above 0.5. The reason for choosing the concatenated multiplication form later is to emphasize that the similarity values of the main parts play a constraining role on the similarity values of the secondary parts.

II. B. 3) Similarity of arcs

There are many types of arc relationships in the conceptual graph, so in order to reduce the complexity of the recursive computation process, this paper utilizes the computation of the similarity of arcs by defining that two arcs are similar if they are the same concept i.e., if the two arcs are represented by the same word, then their similarity is 1, and vice versa.

$$sim_a(a_Q^j, a_R^j) = \begin{cases} 1 & a_Q^j = a_R^j \\ 0 & a_Q^j \neq a_R^j \end{cases} \tag{5}$$

III. Impact of new media on Chinese language expressions

(1) Simplification of language and extensive use of emoticons

In the new media environment, the immediacy and information overload of social media platforms have prompted people to pursue more concise and intuitive communication. Linguistic simplicity is reflected in users' tendency to use short sentences and vocabulary to convey information quickly. This trend has changed the traditional expression habits of the Chinese language to a certain extent, making the language more direct and efficient. The earliest and fastest manifestation of these linguistic changes is in the new media, which is figuratively called the “fifth media”. It is a form of communication and media that uses digital technology as the main channel through the Internet, in addition to the four traditional media, namely, paper-printed newspapers and magazines, outdoor messaging carriers, radio and television. The most widely used forms are the social platforms carried by smartphones, which have the functions of powerful interactivity and instantaneousness, massiveness and sharing, multimedia and hypertext, and personalization and communalization. The majority of the people who need these features are urban youth, and it is they who create the language simplicity and spread it widely so that it becomes a language that spreads rapidly from online social platforms to almost the whole society, which is a form of rapid turnover and long-term development.

(2) Popularity of network terms and neologisms

Social media, as an open communication platform, facilitates the rapid spread of online terms and neologisms. These new words are often a reflection of real social phenomena or an innovative use of existing words. Internet terms usually have a sense of the times and group identity, and they can quickly spread among users and be widely accepted. The creation and popularity of new words not only reflect social and cultural changes, but also the vitality and innovation of language. The extensive use of these words in social media has led to the continuous updating and expansion of the Chinese vocabulary base, and at the same time has provided rich materials for language research. The popularity of Internet terms and neologisms is an important feature of language development in the new media environment, and they have changed the face of the Chinese language to a certain extent. This change can be reflected just by taking the example of simplified terms that have been widely circulated in the recent period, such as the following, which include simplified terms, new network terms, regional specialties, and so on.

IV. Potential impact of new media language on the semantic change of Chinese language

IV. A. Algorithmic validity

IV. A. 1) Performance testing

In order to evaluate the performance of the algorithm, this paper first conducts a test on the correct rate of converting text into conceptual graphs. Eight common sentence types in Chinese were tested, 100 sentences were selected for each, and the rate of correctly recognizing sentence relations was tested as shown in Table 1. The algorithm is more efficient in recognizing simple sentence types such as definite middle, preposition, number, gerund, etc. The average correct rate reaches 0.899. It is relatively poor in recognizing movable-object, subject-predicate, cognate, and movable-complement structures, with an average correct rate of 0.731. This is mainly related to the characteristics of the Chinese language, because some commonly used sentence types, with a large selective nature of the front and back items, a large span of definition levels, and a high degree of dispersion, make it difficult to summarize the laws.

Table 1: Tests the correct ratios of common sentence patterns

Sentence pattern	Ratio	Sentence pattern	Ratio
Subject-predicate relationship	0.718	Verb-object relationship	0.735
Deterministic relationship	0.852	Quantitative relationship	0.966
Dynamic and complementary relationship	0.767	Intermediary relationship	0.938
Relationship within the state	0.839	Allogeneic relationship	0.704

Further three commonly used metrics are Precision, Recall and F-measure, which is the ratio of the number of relevant documents retrieved to the total number of documents retrieved and measures the accuracy of the retrieval.

Recall is the ratio of the number of relevant documents retrieved to the number of all relevant documents in the document library, and measures the search accuracy of the retrieval system. 500 simple questions were selected and tested on 8 types of questions, and their experimental results are shown in Table 2. Among the 8 types of questions, the average Fβ value of this paper's algorithm reaches 67.025%, in which the best retrieval effect is achieved for the number of inquiries, with a high Fβ value of 76.1%. The experimental results prove that this paper's algorithm is effective for Chinese semantic computation.

Table 2: Evaluation of Retrieval Results (%)

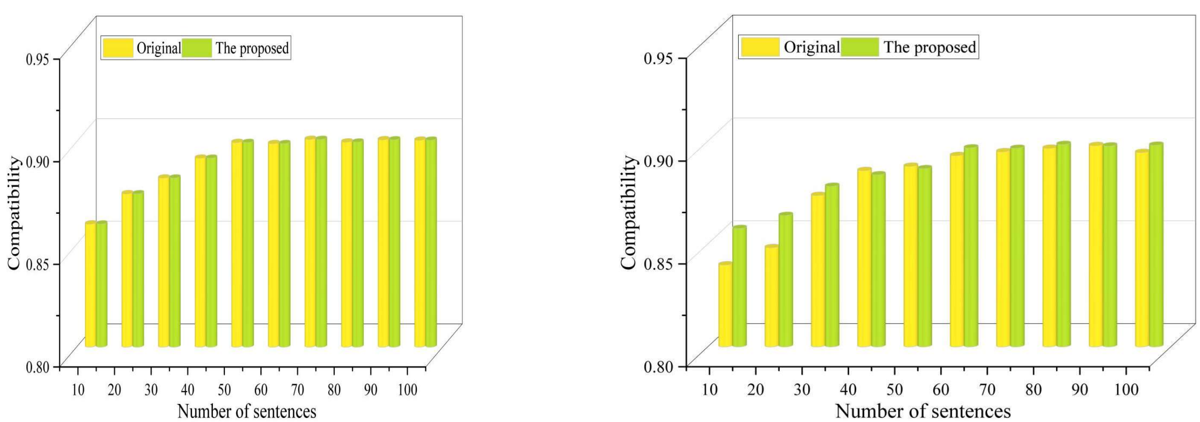
Type	P	R	Fβ	Type	P	R	Fβ
Ask the person	70.3	68.5	69.4	Inquiry definition	60.2	71.3	65.3
Inquire about the location	76.2	72.3	74.2	Inquiry method	59.4	72.2	65.2
Inquire about the time	73.1	68.9	70.9	Ask for the reason	51.6	63.3	56.9
Inquire about the quantity	72.4	80.2	76.1	Inquire about others	49.7	70.2	58.2

IV. A. 2) Analysis of controlled trials

In the experiment, 2 groups of 200 sentences were selected from the user query logs provided by Sogou Labs, and then these 200 sentences were modified separately to form 200 pairs of sentences as the experimental dataset. Among them, 100 pairs are sentences with balanced attribute descriptions and 100 pairs are sentences with partially imbalanced attribute descriptions. Then these sentences are used to obtain 2 groups of 200 pairs of concept maps using the composite concept map construction method, and this paper uses these 200 pairs of composite concept maps to experiment and analyze the algorithm. In this paper, the original algorithm is introduced in the concept map matching process after the selected sentences are preprocessed, syntactically analyzed, semantically analyzed, and composite concept map constructed, and then compared with the improved algorithm proposed in this paper. The evaluation index uses the compatibility function Compat.

The similarity values of a total of 200 pairs of concept maps in these 2 groups were calculated using the original algorithm based on the original algorithm and the improved algorithm in this paper, respectively, and then these calculations were utilized to compare with the subjective empirical values (by separating the 2 groups for the comparison), resulting in the algorithmic compatibility degree of the 2 groups as shown in Fig. 2(a~b) (in the figure, randomly selected 10 pairs, 20 pairs, ..., 90 pairs, and 100 pairs of sentences are randomly selected in the figure to calculate the value of algorithmic compatibility).

From Fig. 2(a), it can be seen that the original algorithm has the same degree of compatibility with the algorithm of this paper, and it can be analyzed that: the two algorithms have the same calculation results for conceptual graphs with balanced attribute descriptions. As can be seen from Fig. 2(b), for conceptual graphs with partially imbalanced attribute descriptions, the improved algorithm of this paper shows its instability in dealing with such problems. Because there are basically no attributes that need to be specially represented in this kind of conceptual graphs, and people will do some descriptions of certain attributes in their semantic expressions, which causes some errors in the results of semantic computation. However, its calculation results are still better than the original algorithm as a whole, with an average compatibility of 0.886, which is 0.6% better than the original algorithm.



(a) Attribute description balance (b) The attribute description is partially unbalanced

Figure 2: Comparison of the compatibility of the two algorithms

IV. B. Case Study on the Trend of Chinese Semantic Evolution

Based on the CLSKB_Core core knowledge base, the term "Tuohao" is used as an example, and the stratified sampling technique is applied for its frequency statistics over time. The frequency distribution of "Tuhao" in different periods is shown in Figure 3. It can be seen that from the Tang Dynasty to the Song Dynasty, the frequency of this word has increased greatly, and the frequency rises to the highest in the period of 2011-2020, reaching 0.0000042.

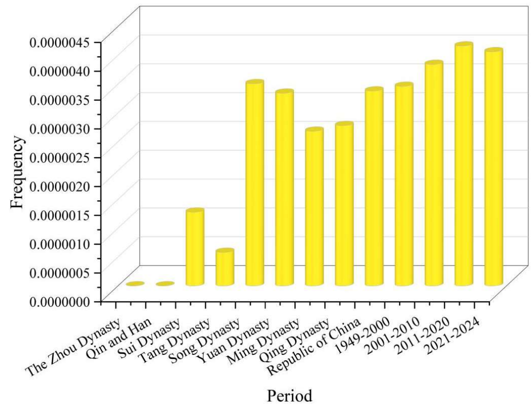


Figure 3: Frequency distribution in different periods

At the level of semantic evolution analysis, the cosine similarity calculation reveals that the trend of semantic evolution of "Tuohao" in different periods is shown in Figure 4. It can be seen that in the period of 2011-2020, the cosine similarity of the word has dropped sharply, from 0.87 to 0.68, indicating that the word "Tuohao" may have undergone semantic changes in this period. It can be seen that around 2010, with the popularization of the Internet and the rise of social media in China, the term "Tuhao" was reconstructed. In 2013, "Tuhao" became the hot Internet term of the year, and brands took advantage of the situation to market their products, and the implementation of "Tuhaojin" transformed the original derogatory meaning of the word into a high-end, eye-catching symbol, which dissolved the negative connotation of the word itself. The case study verifies the effectiveness of the algorithm in exploring the potential impact of new media language on the semantic change of Chinese.

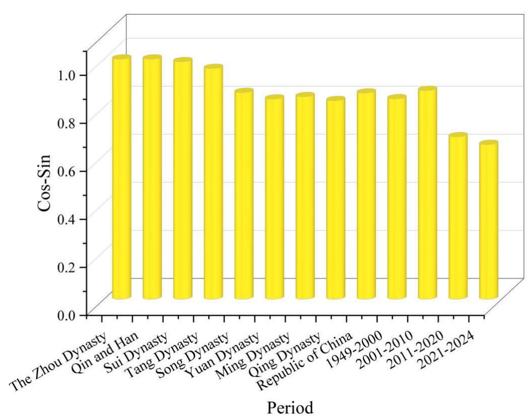


Figure 4: Semantic Evolution Trend

V. Conclusion

In this paper, we build the core knowledge base CLSKB_Core, improve the concept map retrieval algorithm suitable for Chinese semantic computation, and analyze the impact of new media language on Chinese semantic change through case studies.

The algorithm recognizes sentence types with simple structures such as definite middle, preposition, number, gerund, etc. with good efficiency, and the average correct rate reaches 0.899. It recognizes verb-object, subject-predicate, congruent, and verb-complement structures relatively poorly, with an average correct rate of only 0.731. Among eight types of questions, the algorithm in this paper achieves an average Fβ value of 67.025%, among which it retrieves the best results for the questions that ask about the number of questions, and the Fβ value reaches as high as 76.1%. For the conceptual map with partially imbalanced attribute descriptions, the improved algorithm in this paper computes overall better results than the original algorithm, with an average compatibility of 0.886, which is 0.6% better than the original algorithm.

Taking the word “Tuhao” as an example to start the case study, the frequency of this word has increased greatly from the Tang Dynasty to the Song Dynasty, and the frequency rises to the highest in the period of 2011-2020, reaching 0.0000042. The cosine similarity of the word declined sharply from 0.87 to 0.68 in the period 2011-2020, suggesting that the word “Tuhao” may have undergone a semantic change during that period.

References

[1] Li, D. (2016). On Chinese Culture. Springer Singapore.

[2] Zhang, Q. (2015). An introduction to Chinese history and culture. Heidelberg: Springer.

[3] Varis, P., & Van Nuenen, T. (2016). THE INTERNET, LANGUAGE. The Oxford handbook of language and society, 473.

[4] McCulloch, G. (2019). Because internet: Understanding how language is changing. Random House.

[5] TOMA, A. M. (2021). Linguistic Changes and Conventions in Netspeak. ACROSS Journal of Interdisciplinary Cross-border Studies, 4(1).

[6] Winarto, E. R. (2019). Modelling Abbreviation in Internet Slang: a Comparison Study of Indonesian Internet Slang and English Internet Slang. ETERNAL (English Teaching Journal), 10(2).

[7] Zabotnova, M., & Bohdanova, O. V. (2018). Internet Slang as Key means of Interaction in Cyberspace. Development of Philological Sciences in Countries of the European Union taking into Account the Challenges of XXI Century, 146-64.

[8] Vlasov, M., Sychev, O., Toropchina, O., Isaeva, I., Zamashanskaya, E., & Gillespie, D. (2024). The Effects of Problematic Internet Use and Emotional Connotation on Internet Slang Processing: Evidence from a Lexical Decision Task. Journal of Psycholinguistic Research, 53(3), 39.

[9] Liu, J., Zhang, X., & Li, H. (2023). Analysis of Language Phenomena in Internet Slang: A Case Study of Internet Dirty Language. Open Access Library Journal, 10(8), 1-12.

[10] Teodorescu, H. N., & Saharia, N. (2015, October). An internet slang annotated dictionary and its use in assessing message attitude and sentiments. In 2015 International Conference on Speech Technology and Human-Computer Dialogue (SpeD) (pp. 1-8). IEEE.

[11] Jumanazarova, N., & Khamitov, E. (2024). A LINGUISTIC ANALYSIS OF THE EVOLUTION OF INTERNET SLANG. Central Asian Journal of Multidisciplinary Research and Management Studies, 1(17), 133-135.

[12] Odacı, H., & Çıkrıkçı, Ö. (2014). Problematic internet use in terms of gender, attachment styles and subjective well-being in university students. Computers in Human Behavior, 32, 61-66.

[13] Wahid, R., & Farooq, O. (2022). Uses and abuses of Netspeak. International Journal of Social Sciences & Educational Studies, 9(1), 53-59.

[14] Tupamahu, M. S., Uktolseja, L. J., & Gaspersz, S. (2023). The Analysis of Netspeak Used on Instagram. INTERACTION: Jurnal Pendidikan Bahasa, 10(2), 680-691.

[15] Sun, Y. (2020, February). Changes from the Internet Language to Emoji. In 6th International Conference on Education, Language, Art and Inter-Cultural Communication (ICELAIC 2019) (pp. 508-511). Atlantis Press.

[16] Aleksić-Maslać, K., Borović, F., & Biočina, Z. (2018). Comparative content analysis of the Netspeak elements among pupils and students in asynchronous discussion"professor-student". Diacritics, 10, G2.

[17] Augustyn, R., & Prażmo, E. (2020). The Spread of Chinese Virus in the Internet Discourse: A Cognitive Semantic Analysis. Gema Online Journal of Language Studies, 20(4).

[18] Glushkova, S., & Voronina, M. (2017). Structural and semantic analysis of neologism in Chinese language. The Turkish Online Journal of Design, Art and Communication.

[19] Yang, B. (2024). Seven Types of Meaning and Seven Attributes: A Study of Chinese Anti-Pandemic Neologisms Under the Dual Perspective of Lexical Semantics and Lexical Pragmatics. 3L, Language, Linguistics, Literature, 30(1), 158-174.

[20] Xueran, D. (2023). A Study of Code-mixing in Netspeak: from the Perspective of the Meme Theory. Lecture Notes on Language and Literature, 6(7), 1-5.

[21] Jurida, S. H. (2015). Netspeak: Linguistic Properties and Aspects of Asynchronous Online Communication. Bosanski jezik, (12), 21-38.