

A Preliminary Study of Allusion Usage in Chinese Classical Poetry from the Perspective of Citation Analysis: Postprint

Authors: Cai Zhujuan, Zhao Danqun, Zhao Danqun

Date: 2022-11-26T00:00:00+00:00

Abstract

This study conducts theoretical analysis and empirical research on allusion usage in Chinese classical poetry from a citation analysis perspective, aiming to facilitate the transformation and expansion of traditional humanities research paradigms toward data-driven quantitative research paradigms in the big data era, and to promote the interdisciplinary integration and innovative development of Chinese classical poetry allusion research and bibliometric methods. First, through comparative analysis of the similarities and differences between allusion usage behavior (phenomenon) in Chinese classical poetry and citation behavior (phenomenon) in academic papers, theoretical analysis is conducted on the emergence and development of allusion phenomena, allusion motivations and behaviors, and allusion relationships and networks. Second, corpora such as the Complete Song Ci and the Dictionary of Allusions in Song Ci are employed to conduct quantitative empirical research on allusions in Song Ci poetry from the dimensions of allusion application and allusion source. The phenomenon of allusion usage in Chinese classical poetry has a long history with rich research topics, and there is an urgent need for innovation and transformation of research paradigms at the current stage. Theoretical analysis and empirical results demonstrate that citation analysis methods facilitate deeper exploration on large-scale corpora of classical poetry. The greatest difficulty in current research lies in the automatic identification and extraction of allusions. Future work must strengthen close integration with NLP technology, striving to propose more effective automated solutions and quantitative analysis strategies for issues such as allusion identification and extraction, classification of allusion motivations and behaviors, and construction of allusion networks and allusion knowledge graphs, thereby powerfully promoting the transformation of traditional humanities research paradigms from theoretical interpretation to empirical interpretation, from subjective judgment to objective revelation, and from

qualitative analysis to quantitative analysis, and generating innovative results across more digital humanities research topics.

Full Text

Preamble

Volume 66, Issue 20, October 2022

A Preliminary Study on the Use of Allusions in Ancient Chinese Poetry from the Perspective of Citation Analysis

Cai Zhujuan, Zhao Danqun Department of Information Management, Peking University, Beijing 100871

Abstract:

[Purpose/Significance] This paper conducts theoretical analysis and empirical research on the use of allusions in ancient Chinese poetry from the perspective of citation analysis, aiming to promote the transformation and expansion of traditional humanities research paradigms toward data-driven quantitative research paradigms in the era of big data, and to foster interdisciplinary integration and innovative development between Chinese poetry allusion studies and bibliometric methods. [Method/Process] First, through a comparative analysis of the similarities and differences between allusion usage behaviors (phenomena) in Chinese poetry and citation behaviors (phenomena) in academic papers, the paper examines the emergence and development of allusion usage, motivations and behaviors, and relationships and networks. Second, using corpora such as *Complete Song Ci* and *Dictionary of Allusions in Song Ci*, a quantitative empirical study of allusion usage in Song Ci poetry is conducted from both the citing and cited dimensions. [Result/Conclusion] The phenomenon of allusion usage in ancient Chinese poetry has a long history with rich research topics, yet urgently requires innovation and transformation of research paradigms. Theoretical analysis and empirical results demonstrate that citation analysis facilitates deeper exploration of large-scale poetry corpora. The greatest current challenge lies in the automatic recognition and extraction of allusions. Future research should strengthen integration with NLP technology to develop more effective automated solutions and quantitative analysis strategies for allusion identification, motivation classification, network construction, and knowledge graph development, thereby promoting the transformation of traditional humanities research from theoretical interpretation to empirical analysis, from subjective judgment to objective revelation, and from qualitative to quantitative analysis, ultimately generating innovative results across more digital humanities research topics.

Keywords: citation analysis; ancient Chinese poetry; allusions; allusion motivation; allusion behavior

Classification Numbers: G301; I207.2

DOI: 10.13266/j.issn.0252-3116.2022.20.009

1. Introduction

Chinese civilization has a long and profound history, with vast cultural heritage accumulated and transmitted over thousands of years, among which ancient poetry undoubtedly represents the most brilliant component. In the creation of ancient Chinese poetry, the use of allusions (用典) is a common rhetorical technique that scholars throughout history have employed with great ingenuity and deliberation. Since the *Book of Songs*, the phenomenon of allusion usage has continuously evolved, with topics such as the types and quantities of allusions, methods and stylistic preferences, sources and textual research, and the evolution of allusion imagery and meaning becoming focal points for humanities scholars and key areas in poetry studies, yielding abundant research 成果.

From a methodological perspective, research on allusion usage in ancient poetry can be categorized into three main types: (1) Traditional humanities research, which relies on scholars' extensive knowledge and dedicated study, focusing on literary, historical, and rhetorical perspectives; (2) Quantitative research, which employs mathematical and statistical methods to analyze the allusion behaviors of specific poets or works and the influence of classical literature on poetic creation; and (3) Automated analysis using Natural Language Processing (NLP) technologies to experiment with algorithmic approaches to various allusion-related questions.

To date, the first category constitutes the vast majority and dominates the field, while the latter two have emerged more recently with paradigms still in formation. Further literature review reveals that although traditional humanities research has produced fruitful results covering extensive topics, most focus on micro-level analysis with qualitative examination and illustrative examples. While diverse schools of thought have emerged, such studies generally lack support from large-sample data. Moreover, constrained by individual scholars' capabilities, conclusions tend to be subjective and limited in scale, rarely addressing meso- or macro-level research or covering all poets and texts, making them increasingly inadequate for deepening research in the current big data era.

Citation analysis, established in the 1960s, is a typical "data-driven" research method dedicated to quantitative study of citation relationships in academic literature. As a core method of bibliometrics, it has evolved from simple citation counting to advanced stages of citation relationship clustering, network analysis, and knowledge graph visualization, finding wide application across numerous fields. Notably, allusion usage in poetry and citation in academic papers share natural similarities, yet also exhibit inevitable differences due to their distinct origins and unique characteristics. This paper introduces citation analysis into Chinese poetry allusion studies, conducting comprehensive comparative analysis with academic citation practices and employing corpora such as *Complete Song Ci* for large-scale quantitative analysis and empirical research. This endeavor

aims to facilitate the transformation of traditional humanities paradigms toward data-driven quantitative research, promoting deep interdisciplinary integration and innovative development between poetry allusion studies and bibliometrics.

2. Theoretical Analysis of Allusion Usage in Ancient Poetry

This section conducts theoretical analysis of allusion usage in ancient Chinese poetry from the perspective of citation analysis, attempting a comprehensive summary of similarities and differences between poetic allusion usage and academic citation.

2.1 Emergence and Evolution of Allusion Phenomena

Citation phenomena primarily emerged with the rapid development of modern science and disputes over scientific discovery priority. The advent of academic journals and the formation of reference traditions facilitated the establishment and refinement of citation mechanisms, gradually making them a fundamental academic norm. In the early 1960s, Dr. E. Garfield discovered the retrieval and evaluation value inherent in the widespread citation relationships within academic literature, fully exploiting these through the creation of the *Science Citation Index* (SCI). Since then, citation analysis has emerged and continues to develop.

Compared with citation phenomena, the phenomenon of allusion usage in Chinese poetry has a much longer history, traceable to the *Book of Songs* [1]. As the origin of ancient Chinese poetry, the *Book of Songs* served not only as the source of poetic allusions but also as source literature for numerous later works, establishing a tradition of quoting the *Book of Songs* that persisted through the pre-Qin, Han, and Wei-Jin periods [3-4]. Beyond this, works such as *Chu Ci*, *Zhou Yi*, *Laozi*, *Zhuangzi*, *Xunzi*, *Huainanzi*, *Lüshi Chunqiu*, *Wu Yue Chunqiu*, *Han Shi Wai Zhuan*, *Shuo Yuan*, as well as yuefu, ballads, and proverbs were gradually adopted as source literature [5]. The increase in Daoist allusions during the Eastern Jin Dynasty, the popularity of allusion techniques in Southern Dynasties literature, poets' enhanced ability to condense allusions [6], and rhetoricians' examination of poetic allusions [7] further promoted the development of allusion usage, significantly increasing the quantity, frequency, and density of allusions.

By the Tang and Song dynasties, poetic creation flourished unprecedentedly, with allusion usage becoming an extremely important creative practice. From early to high to late Tang, Tang poetry drew from an extremely broad range of sources, encompassing numerous Confucian, Buddhist, Daoist texts, and mythological fiction [8]. In the Song Dynasty, poetry exhibited characteristics of "using words as poetry, using argumentation as poetry, using erudition as poetry,"

namely the prosaic, argumentative, and allusion-rich nature of Song poetry [9-10]. Allusion techniques became more sophisticated, as exemplified by Huang Tingjian's concepts of "seizing the embryo and changing the bone" (夺胎换骨) and "turning iron into gold" (点铁成金) [11]. With the development of ci poetry, allusion usage in Song Ci gradually increased [12], with poets particularly favoring the adaptation of Tang poetry, especially late Tang and Five Dynasties poetry [13]. Additionally, Wang Anshi's promotion of "collection ci" (集句词) and Su Shi's development of "kuo ci" (!括词, which adapts other poems by editing them primarily "to fit musical patterns and facilitate singing" [14-15]) pushed allusion usage (particularly of textual allusions) to its peak. Among poets, Xin Qiji of the Southern Song was a master of allusion usage, with scholars summarizing his methods as seven types: direct use, reverse use, flexible use, transformative use, implicit use, consecutive use, and collection [16].

In the subsequent Yuan, Ming, and Qing dynasties, the peak of Chinese poetry creation had passed, and allusion usage remained largely consistent with Tang and Song practices, with few major innovations [17]. This brief review reveals that the emergence and evolution of allusion usage in Chinese poetry represents a process of continuously increasing allusion quantity, frequency, and density, expanding source ranges, and enriching techniques. Based on investigation of the Chinese Poetry Classical Corpus Database (CPDB) [18], I found that allusion usage in Song Ci was extremely common, appearing not only in main texts but also in titles and prefaces, with special phenomena such as "one allusion, multiple sources," "one allusion, multiple surfaces," "allusion within allusion," and "every line an allusion" (collection ci and kuo ci).

2.2 Analysis of Allusion Motivations and Behaviors

Psychologists believe that motivation arises from various hierarchical needs and forms the basis of most human behaviors, with close connections between motivation and behavior. Citation motivation is a social motivation arising from academic research needs, with citation behavior representing its external manifestation.

Allusions in poetry are primarily divided into "event allusions" (事典) and "textual allusions" (语典). Event allusions refer to ancient stories such as historical facts, operas, myths, legends, and institutional objects, while textual allusions refer to "established phrases with origins," such as influential lines from existing poetry and prose [2]. Generally, from pre-Qin through Han to Wei-Jin periods, poetic allusions continued the tradition of quoting the *Book of Songs* [3-4]. Additionally, texts like *Chu Ci*, *Zhou Yi*, *Laozi*, *Zhuangzi*, *Xunzi*, *Huainanzi*, *Lüshi Chunqiu*, *Wu Yue Chunqiu*, *Han Shi Wai Zhuan*, *Shuo Yuan*, as well as yuefu, ballads, and proverbs were gradually adopted as source literature [5]. The increase in Daoist allusions during the Eastern Jin Dynasty, the popularity of allusion techniques in Southern Dynasties literature, poets' enhanced ability to condense allusions [6], and rhetoricians' examination of poetic allusions [7] further promoted the development of allusion usage, significantly increasing the quantity, frequency,

and density of allusions.

Citation motivation has been a central concern as it relates to the validity of research assumptions and theoretical completeness. E. Garfield initially identified 15 types of citation motivations in academic writing through observation [19], with numerous subsequent discussions based on different methods and analytical frameworks [20-23]. Recently, automatic identification and classification of citation motivations using large-scale corpora and NLP technology has become a cutting-edge research topic.

The following discussion examines allusion motivations and behaviors in ancient Chinese poetry, with brief explanations of their special characteristics.

2.2.1 Allusion Motivations and Types Overall, allusion motivations in poetry are as complex and variable as citation motivations. As a literary rhetorical device, poets use allusions to express emotions, illustrate principles, satirize contemporary issues through historical references, convey aspirations through objects, or demonstrate erudition and originality, with vastly different purposes and preferences. Broadly, allusion motivations can be categorized into four types [24]: (1) **Testimony**: Quoting predecessors' words or deeds to prove one's own viewpoints or provide reasons for actions; (2) **Enhancement**: Comparing allusions with current matters to show superiority/inferiority or express emotions and preferences; (3) **Substitution**: Using allusion surfaces to replace current matters or objects; (4) **Representation**: Contrasting allusion contexts with poetic contexts to generate implied meanings and express true intentions.

Upon closer examination, "testimony" most resembles citation motivations, while the other three types arise more from literary creation needs. "Testimony" best reflects the Chinese humanities tradition of "quoting classics and authorities" and "no word without origin." Research on citation phenomena in ancient texts like *Guwen Guanzhi* and *Guwenqi Leizuan* also found this motivation to be most common [25].

Like citation motivation research, which faces the challenge of "getting inside the author's mind," allusion motivation research encounters similar difficulties. Combining historical context, personal circumstances, poetic content, creative emotions, and allusion sources and meanings to analyze allusion motivations using NLP technology presents even greater challenges than citation motivation research, as academic citations are more objective and rational, while literary allusions are more subjective and emotional.

2.2.2 Analysis of Allusion Behaviors Allusion behaviors, like citation behaviors, externalize motivations. Whether event or textual allusions, an allusion behavior consists of three components: (1) **Allusion user** (the poet or work itself, called "citing literature"); (2) **Allusion surface** (also called "典面"), the textual expression of the allusion in the poem; (3) **Allusion source**, including the source story/origin and the source literature (the earliest text recording the

source) [26].

Academic citations are divided into explicit and implicit citations, with the vast majority being explicit—authors use standardized citation markers. Similarly, poetic allusions have “explicit use” (明引) and “implicit use” (暗用) [27], with explicit use indicating clear source identification. In stark contrast, most poetic allusions are implicit, with only a small number using explicit methods (such as indicating source literature in prefaces or annotating sources in collection ci). Thus, identifying allusions relies primarily on the “allusion surface,” which can be understood as an identifier for poetic allusions.

Unlike standardized, semantically clear, and prominently positioned citation markers, allusion surfaces in poetry are typically variable, formed through complex and lengthy processes resulting from the interaction between “cohesion” and “extroversion.” “Cohesion” refers to the relative stabilization of allusion surfaces, while “extroversion” refers to poets’ adjustments to existing surfaces or creation of new ones to better meet their needs or satisfy metrical requirements [24]. Consequently, one allusion may have multiple different surfaces (“one allusion, multiple surfaces”). Du Fu’ s line “A single petal flies, and spring diminishes” exemplifies this, often directly quoted by Song poets but also modified through character substitution or deletion in other works. These phenomena—primarily implicit use and unstable surfaces (“one allusion, multiple surfaces” or “different surfaces, same source”)—pose significant challenges for analysis. Even with extensive existing research (e.g., *Dictionary of Allusions in Song Ci* [2], *Dictionary of Allusions in Tang Poetry* [28], *Dictionary of Poetic Allusions* [29], *Great Dictionary of Chinese Allusions* [30]), automatic allusion recognition remains limited to literal matching, with “one allusion, multiple surfaces” or “different surfaces, same source” problems requiring domain ontology construction, synonym 对照 tools, or expert assistance [31-33].

The following special phenomena in poetic allusions warrant deeper investigation by both literary scholars and bibliometricians:

- (1) **“One allusion, multiple sources”** : Beyond multiple surfaces, some allusions have multiple different sources, establishing “one-to-many” relationships with source literature. For example, the allusion “a vine branch” has two sources: Li Shangyin’s “North Green Vine” and Li Dong’s “Sending Someone to Tiantai.”
- (2) **“Same surface, different sources”** : Different poems may use identical allusion surfaces but derive from different sources. For instance, “wine star” in Zhu Dunru’ s work sources from *Records of the Three Kingdoms* and *Book of Jin*, while the same surface in Wu Yong’ s work sources from Pi Rixiu’ s Tang dynasty poetry.
- (3) **“Stacked citation”** : Using multiple identical or different allusions within one line or line group [24], categorized as pure stacking, frame stacking, and layered stacking. “Layered stacking” is particularly special, essentially “allusion within allusion” or nested allusions.

- (4) **“Repeated allusion usage”**: A special case of stacked citation where the same allusion (with identical or different surfaces) appears multiple times in one poem. Xin Qiji’s “Congratulations to the Groom: On Tengwang Pavilion” repeats the same allusion twice with different surfaces.
- (5) **“Collection allusion usage”**: Creating new poems by collecting existing lines from others’ works [34]. Since Wang Anshi of the Northern Song pioneered this practice, it’s also called “Jinggong style” [35]. Collections may draw from multiple authors or specialize in one, such as the Song dynasty “Collected Du Fu Poems” focusing exclusively on Du Fu’s lines, with Wen Tianxiang achieving the highest mastery [36].

Notably, collection ci gradually adopted explicit citation methods. Su Shi annotated the poets’ names for collected lines in his “Nanxiangzi • Changwang Songchun Bei” [37], while Wen Tianxiang annotated each line with its original title in his *Collected Du Fu Poems*, abbreviating lengthy titles [38].

Debate has long surrounded collection allusion usage, with some viewing it as “poetry theft” categorized as “stealing words,” “stealing meaning,” or “stealing momentum,” with “stealing words” considered the most egregious [39]. By this view, collection ci represents the ultimate “poetry theft.” However, unlike modern academic citation norms, creating seamless adaptations or successful collection poems required not only extensive knowledge but also considerable creative talent. Moreover, the “poetry theft” concept overlooks the philological value of collection ci in recovering lost texts and textual collation [40].

2.3 Allusion Relationships and Network Construction

Allusion relationship chains and networks in poetry can be constructed by analogy with academic citation relationships and networks, enabling related empirical and quantitative analyses.

2.3.1 Analysis of Allusion Relationships A basic allusion relationship chain is described as follows (boxes indicate nodes, parentheses indicate relationships):

Author (Poet) → (Creates) → Poem → (Uses/Quotes) → Allusion → (Derives from) → Source Literature

This chain contains two major relationship types:

- (1) **Relationships between different nodes**: Such as “creation/created” between author and poem, “use/used” between poem and allusion, and “derive from/originate” between allusion and source literature. Due to phenomena like “one allusion, multiple sources,” “same surface, different sources,” and “different surfaces, same source,” multiple “one-to-many” relationships exist among allusions, surfaces, and sources.

Ancient scholars, particularly in the “annotation of allusions” tradition, extensively studied these relationships. Annotators’ work involved identifying sources and explaining meanings [41]. For example, Li Shan’s commentary on Cao Zhi’s “Seven Sorrows” in *Wenxuan* [42] established allusion relationships while creating a new “explanation/explained” relationship. However, such annotations represent individual interpretations, and different annotators may provide different explanations for the same line [43].

(2) **Relationships between similar nodes:** Including:

- **Coupling relationships:** If two poems use the same allusion, they have allusion coupling, measurable by shared allusion count. If they use allusions from the same source literature, they have source literature coupling. The former is more restrictive since different allusions may share sources. When calculating coupling strength, source duplication must be considered.
- **Co-citation relationships (or co-occurrence):** If two different allusions appear together in one or more poems, they have allusion co-citation, with co-occurrence frequency as strength. Similarly, source literature co-citation can be measured.

Notably, academic co-citation relationships typically change dynamically over time (monotonically increasing), whereas poetic allusion co-citation is generally fixed because both the poems and classical texts are historical corpora that no longer expand.

2.3.2 Allusion Network Construction Theoretically, an allusion network comprises node sets and relationship sets, classifiable as:

- (1) **Homogeneous networks:** All nodes and relationships are of the same type, such as poem coupling networks, allusion co-citation networks, or source literature co-citation networks, which can be weighted or unweighted.
- (2) **Heterogeneous networks:** Containing multiple node and relationship types, such as “author-poem-allusion” networks, “poem-allusion-source literature” networks, or “author-poem-allusion-source literature” networks, which essentially function as semantic networks or knowledge graphs.

While empirical research on allusion networks based on coupling and co-citation remains rare, citation analysis techniques are fully capable of addressing these issues—precisely the essence of citation analysis applications. In this sense, introducing citation analysis represents both an innovation and a paradigm shift for poetry allusion studies.

2.4 Summary of Theoretical Analysis

The above theoretical analysis reveals that poetic allusions and academic citations share natural similarities while exhibiting numerous differences. A com-

parative summary is provided in .

3. Empirical Research on Allusion Usage in Ancient Poetry

Given that the Tang and Song dynasties represent the peak of allusion usage, they provide ideal cases for empirical study. Limited by space, this section focuses on Song Ci poetry using *Complete Song Ci* and *Dictionary of Allusions in Song Ci*.

3.1 Corpus Selection and Preprocessing

3.1.1 Corpus Selection The empirical study requires three main corpora: *Complete Song Ci*, *Dictionary of Allusions in Song Ci*, and *Author Index of Song Ci*.

- (1) ***Complete Song Ci*** (Primary corpus): Tang Guizhang' s compilation [45] represents the definitive collection and indispensable resource. The study uses the JSON-formatted third edition from CPDB [18], containing poet names, texts, and tune patterns.
- (2) ***Dictionary of Allusions in Song Ci*** (Supporting corpus): Among several specialized dictionaries, Fan Zhilin' s edition [2] was selected for its 10,516 entries arranged by stroke count, each including source, explanation, and example sentences, as shown in [Figure 1: see original paper]. The electronic version from CNKI was crawled for local storage.
- (3) ***Author Index of Song Ci*** (Auxiliary corpus): This appendix to *Complete Song Ci* lacks an electronic version and required scanning and OCR processing (using <https://ocr.wdku.net/>) with manual correction for conversion to structured data.

3.1.2 Corpus Preprocessing Preprocessing addressed structuring, error correction, and noise removal:

- (1) ***Complete Song Ci* preprocessing:**
 - Converted JSON format to structured data, yielding 21,051 poems and excluding 1,568 anonymous works, resulting in 19,483 poems for analysis with fields: poem ID, poet, tune pattern, and text.
 - Standardized poet names against the author index and Baidu Chinese (<https://hanyu.baidu.com/>).
 - Disambiguated poets and assigned poet IDs.
 - Segmented poems into lines and calculated absolute and relative line positions.
- (2) ***Dictionary of Allusions in Song Ci* preprocessing:**
 - Removed redundant source column content.

- Extracted allusion surfaces and corresponding source literature using regex based on “see” and “refer to” relationships.
- Separated similar entries.

This yielded a structured table with fields: allusion surface, allusion ID, and source literature.

- (3) **Mapping table:** The processed *Complete Song Ci* and dictionary were matched by allusion surface using forward maximum matching, producing 48,454 records with fields including poem ID, poet, poet ID, tune pattern, line, line position, allusion surface, allusion ID, and source literature, as partially shown in .
- (4) **Author Index preprocessing:** In addition to digitization, the index was enhanced with temporal periods using Wang Zhaopeng’s six-stage method [46]: (1) Peaceful era (1017-1067); (2) Reform era (1068-1125); (3) War era (1110-1162); (4) Revival era (1163-1207); (5) Appeasement era (1208-1265); (6) Subjugation era (1252-1310). The final index includes poet ID, name, period, and notes, as partially shown in .

3.2 Research Design

Complete empirical research includes dimensional analysis and integrated analysis. The latter, involving complex network integration, will be addressed separately. This section focuses on dimensional analysis, subdivided into citing dimension (active side) and cited dimension (passive side).

3.2.1 Citing Dimension Analysis As the active side of allusion relationships, this analyzes “allusion-using poets” and “allusion-using poems,” including: (1) Allusion quantity/frequency to measure prevalence; (2) Allusion intensity to measure repeated usage; (3) Allusion position to analyze distribution; (4) Coupling strength to measure inter-poem connections.

3.2.2 Cited Dimension Analysis As the passive side, this analyzes “allusions” and “source literature,” including: (1) Citation frequency to measure usage impact; (2) Citation intensity for repeated citations; (3) Citation position for usage patterns; (4) Source literature co-citation strength.

3.3 Empirical Analysis and Discussion

Based on the above, dimensional analyses yielded these key findings:

- (1) **Allusion usage is universal in Song Ci:** Among 19,483 poems, 15,974 (81.99%) contain allusions, with 3,509 (18.01%) containing none. Allusion-using poems average 3.03 allusions each, with approximately 80% using fewer than 4 allusions and 20% using between 5-16 (see [Figure 2: see original paper]). The maximum of 16 allusions appears in four poems by Su Shi, Liu Kezhuang, Zhang Yuangan, and Chen He, demonstrating

remarkable allusive density. Temporal analysis shows a clear upward trend in average allusions per poem across periods ($R^2 = 0.83$).

- (2) **Allusion positions follow fixed patterns:** Absolute position peaks at line 3 (6,935 instances, 14.31%), dropping sharply after line 8. Relative position shows a “three peaks, three valleys” pattern (“//”), consistent across all six periods (see [Figure 3: see original paper]).
- (3) **Repeated allusion usage is rare:** Only 729 poems (4.56%) repeat the same allusion in the main text. One extreme case repeats an allusion 8 times in a single poem (Hong Shi’ s “Tiaoxiaoling • Fisherman’ s Drink”).
- (4) **Source literature shows enduring influence:** The 2,186 source texts are led by *Wenxuan*, *Records of the Grand Historian*, *Book of Han*, *Shishuo Xinyu*, *Chu Ci*, *Book of Jin*, *Book of Later Han*, *Book of Songs*, *Su Shi Poetry Collection*, *Zhuangzi*, *Old Book of Tang*, *Miscellaneous Records of the Western Capital*, *History of the Southern Dynasties*, *Liezi*, *Tao Yuanming Collection*, *Book of Song*, *Records of the Three Kingdoms*, *Book of Rites*, *Huainanzi*, and *Song of Everlasting Regret*. *Wenxuan* remains the most influential across all periods, cited by 3,855 poems (24.13% of allusion-using poems), demonstrating the enduring impact of classic texts.
- (5) **Coupling relationships are common but weak:** Among 15,974 allusion-using poems, only 169 lack coupling relationships with others. The remaining 15,805 poems form 1,608,792 coupling pairs, with 98.8% having strength of 1. Among the 1.2% with strength ≥ 2 , higher coupling strength indicates greater similarity between poems. An extreme case shows Li Qingzhao’ s and Su’ s “Changshoule • Wei Han Ying Hou” sharing 9 allusions with only 9 differing characters.
- (6) **Co-citation relationships among sources are common and strong:** Among 2,186 source texts, only 152 lack co-citation relationships. The remaining 2,034 texts form 53,372 co-citation pairs: 95.66% with strength < 10 , 2.62% in $[10, 20)$, and 1.72% ≥ 20 (see [Figure 4: see original paper], where node size indicates number of co-cited partners and edge thickness indicates strength).

Conclusion

This paper has comprehensively analyzed allusion usage in ancient Chinese poetry from the perspective of citation analysis, conducting empirical research on Song Ci poetry. Preliminary findings demonstrate that citation analysis offers theoretical innovation and empirical feasibility for poetry allusion studies, representing a valuable exploratory approach worthy of broader application.

After more than half a century of development, citation analysis has entered

the era of full-text 计量 analysis, achieving continuous breakthroughs in citation content, sentiment, and context analysis through integration with NLP and visualization tools. In this context, the latter two research categories mentioned in the introduction can be organically integrated and deepened within citation analysis. Citation analysis is poised to become a new paradigm for poetry allusion studies, parallel to and complementary with traditional humanities approaches, fostering a new model combining “theory + empiricism,” “subjective + objective,” and “qualitative + quantitative” analysis.

However, the greatest challenge in applying citation analysis to poetry allusion studies remains accurate allusion recognition and extraction. Three main reasons account for this: (1) Poetic allusions belong to literary creation, reflecting subjective emotions and personal preferences that resist objective, consistent interpretation; (2) Special phenomena like “one allusion, multiple sources,” “same surface, different sources,” “different surfaces, same source,” and “allusion within allusion” complicate even expert annotation; (3) Current text reuse detection and allusion recognition algorithms remain exploratory and insufficiently robust.

Presently, citation analysis of poetic allusions must rely on existing expert-compiled dictionaries, with future breakthroughs in automatic allusion recognition needed to expand research scope and depth, and to develop practical applications in knowledge graph construction, multi-dimensional querying, and 计量 analysis.

References

- [1] Chen Dongbiao. Investigation of Allusion Usage in Pre-Tang Poetry [J]. *Journal of Changsha Hydroelectric Normal College (Social Science Edition)*, 1996(4): 57-62.
- [2] Fan Zhilin, Liu Yangzhong, Xu Yan. *Dictionary of Allusions in Song Ci* [M]. Wuhan: Hubei Dictionary Press, 2001: 1.
- [3] Lin Gang. On Quoting Poetry [J]. *Theoretical Studies in Literature and Art*, 2007(4): 43-49.
- [4] Wang Sihao. A Neglected Phenomenon in Style Structure Selection: On the Omission of “Shi Yue” in Han Fu’ s Quoting of the *Book of Songs* [J]. *Literary Heritage*, 2013(4): 55-61.
- [5] Li Pengfei. *Aesthetic Research on Allusion Usage in Medieval Poetry* [M]. Wuhan: Wuhan University Press, 2016: 29.
- [6] Ma Cao. Poetic Function and Literary Consciousness: An Investigation of the Aesthetic Principles of Allusion Usage in Wei-Jin Southern and Northern Dynasties Literature [J]. *Journal of Harbin Institute of Technology (Social Sciences Edition)*, 2018, 20(6): 75-81.
- [7] Hu Dalei. On Allusion Usage in Six Dynasties Poetry: Also on “Poetry Expresses Intent” and Collective Unconsciousness [J]. *Literary Review*, 2014(5): 154-160.
- [8] Qin Hongli. *Research on Allusion Art in Late Tang Poetry* [D]. Nanjing: Nanjing Normal University, 2019.
- [9] Dong Chaogang. Also on Song Poetry Characteristics: Reflections on “Using Words as Poetry, Using Argument as Poetry, Using Erudition as Poetry” [J]. *Journal of Tangshan Teachers College*, 2004(4): 10-12, 61.
- [10] Liu Ying. Words, Erudition, Argument: Reflections on Su Shi and Huang Tingjian’ s “Using Words as Poetry, Using Argument as Poetry, Using Erudition as Poetry” [J]. *Anhui*

Literature (Second Half), 2008(5): 56-57. [11] Mo Lifeng. Distinguishing Huang Tingjian' s "Seizing the Embryo and Changing the Bone" [J]. *Social Sciences in China*, 1983(5): 187-198. [12] Chen Fusheng. On Allusion Usage in Ci Poetry and the Development of Tang-Song Ci [J]. *Inner Mongolia Social Sciences (Chinese Edition)*, 2003(S1): 106-107. [13] Li Dingguang. On the Close Relationship Between Northern Song Ci and Late Tang Poetry: Also on Correctly Interpreting the Cultural Implications of Song Ci' s Adaptation of Tang Poetry [J]. *Seeker*, 2006(11): 170-172, 187. [14] Wu Chengxue. On Song Dynasty Kuo Ci [J]. *Literary Heritage*, 2000(4): 10. [15] Zheng Yuan. On Su Shi' s Kuo Ci [J]. *Literary Heritage*, 2006(3): 147-149. [16] Fan Jinjun. On Xin Qiji' s Allusion Usage Methods [J]. *Chuanshan Academic Journal*, 2006(1): 111-114. [17] Zhang Yuelun. A Brief Discussion on Allusion Usage in Classical Chinese Poetry [J]. *Journal of Hunan University of Humanities, Science and Technology*, 2007(1): 138-140. [18] Chinese Poetry Classical Corpus Database (CPDB) [EB/OL]. [2022-06-13]. <https://github.com/chinese-poetry>. [19] GARFIELD E. Can Citation Indexing Be Automated? [C] // STEVENS M E, GIULIANO V E, HEILPRIN L B. *Statistical Association Methods for Mechanized Documentation: Symposium Proceedings*. Washington DC: The National Bureau of Standards, 1965: 189-192. [20] WEINSTOCK M. Citation Indexes [M] // KENT A. *Encyclopedia of Library and Information Science: Vol. 5*. New York: Marcel Dekker, 1971: 16-40. [21] THORNE F C. The Citation Index: Another Case of Spurious Validity [J]. *Journal of Clinical Psychology*, 1977, 33(4): 1157-1161. [22] BROOKS T A. Private Acts and Public Objects: An Investigation of Citer Motivations [J]. *Journal of the American Society for Information Science*, 1985, 36(4): 223-229. [23] BORGMANN L, DANIEL H. What Do Citation Counts Measure? A Review of Studies on Citing Behavior [J]. *Journal of Documentation*, 2008, 64(1): 45-80. [24] Luo Jiyong. *Research on Allusion Usage* [M]. Wuhan: Wuhan University Press, 2005. [25] Shu Fei, Feng Lixuan, Qiu Junping, et al. Citation Analysis Research Based on Classical Chinese Texts' Quoting of Classics [J]. *Journal of the China Society for Scientific and Technical Information*, 2021, 40(12): 1338-1346. [26] Zhu Dong. *Research on Allusion Usage in Tang Dynasty Examination Poems* [D]. Wuhan: Wuhan University, 2013. [27] Chen Wangdao. *Rhetoric Outline* [M]. Shanghai: Shanghai People' s Publishing House, 1976: 97. [28] Fan Zhilin, Wu Gengshun. *Dictionary of Allusions in Tang Poetry* [M]. Wuhan: Hubei Dictionary Press, 1989. [29] Peng Qingsheng, Qu Lingqi. *Dictionary of Poetic Allusions* [M]. Taiyuan: Shuhai Publishing House, 1990. [30] Zhao Yingduo. *Great Dictionary of Chinese Allusions (Dictionary Edition)* [M]. Shanghai: Shanghai Dictionary Press, 2014. [31] Yao Rui. Automatic Analysis System for Poetic Allusions Based on Allusion Domain Ontology [J]. *Software Guide*, 2011, 10(6): 80-82. [32] Souyun [EB/OL]. [2022-08-07]. <https://sou-yun.cn/index.aspx>. [33] TANG X, LIANG S, ZHENG J, et al. Automatic Recognition of Allusions in Tang Poetry Based on BERT [C] // *2019 International Conference on Asian Language Processing (IALP)*. Piscataway: IEEE, 2019. [34] Zhang Minghua. The Development and Characteristics of Collection Poetry [J]. *Journal of Nanjing Normal University (School of Chinese Language and Literature)*, 2006(4): 24-29. [35] Liu Chengguo.

Alternative Interpretation of “Jinggong Style” [J]. *Literary Heritage*, 2006(4): 149-151. [36] Zuo Hanlin, Gao Fujun. Three Development Stages and Characteristics of Song Dynasty Collection Du Fu Poetry [J]. *Journal of Weinan Normal University*, 2012, 27(11): 88-92. [37] Tang Guizhang, Miao Yie, Ye Jiaying, et al. *Tang-Song Ci Appreciation Dictionary (Tang, Five Dynasties, Northern Song)* [M]. Shanghai: Shanghai Dictionary Press, 1988: 657. [38] Wen Tianxiang. *Complete Works of Mr. Wenshan* [M]. Beijing: Beijing Chinese Bookstore, 1985: 387. [39] Jiaoran. *Poetry Style with Annotations* [M]. ZHOU Weide, annotator. Hangzhou: Zhejiang Ancient Books Publishing House, 1993: 31. [40] Zhang Minghua. On the Philological Value and Relativity of Collection Poetry [J]. *Literary Heritage*, 2012(2): 91-100. [41] Bian Renhai. Li Shan’ s Citation-Style Commentary [J]. *Journal of Xinyang Normal University (Philosophy and Social Sciences Edition)*, 2005(3): 78-82. [42] Xiao Tong. *Wenxuan* [M]. LI Shan, annotator. Beijing: Zhonghua Book Company, 1977: 329. [43] Wang Yijia. Allusion Usage and “Allusion Annotation” : The *Shishuo Xinyu* in Song Dynasty Du Fu Commentary and Its Academic Significance [J]. *Beijing Social Sciences*, 2021(3): 13-24. [44] Liu Yingzhu. On Citation Norm Violations in Academic Papers [J]. *Acta Editologica*, 2014, 26(1): 7-9. [45] Tang Guizhang. *Complete Song Ci* [M]. Beijing: Zhonghua Book Company, 1965. [46] Wang Zhaopeng. On the Development of Song Ci [J]. *Jinan Journal (Philosophy and Social Sciences)*, 2000(6): 14-23, 29.

Author Contributions:

Cai Zhujuan: Topic selection, paper writing, data processing, and empirical analysis.

Zhao Danqun: Guidance on topic selection, paper writing, and revision.

Research on Allusions Usage in Ancient Chinese Poetry from the Perspective of Citation Analysis

Cai Zhujuan, Zhao Danqun

Department of Information Management, Peking University, Beijing 100871

Abstract: [Purpose/Significance] From the perspective of citation analysis (CA), this paper theoretically analyzes and empirically studies the use of allusions in ancient Chinese poetry in order to promote the transformation and expansion of traditional humanities research paradigms to data-driven quantitative research paradigms in the era of big data, and to foster integration and innovative development of ancient Chinese poetry allusion studies and bibliometric methods. [Method/Process] Firstly, the paper gave a completely comparative discussion focusing on the similarities and differences between the allusion behaviors (phenomena) in ancient Chinese poetry and the citation behaviors (phenomena) among academic papers from the following three aspects: ancient Chinese poetry allusion phenomenon and its emergence & development; motivations and behaviors while using allusions; allusion relationships and allusion networks. Secondly, a quantitative empirical study about allusions using based on *The Complete Song Ci & The Allusions’ Dictionary of Song Ci* was carried out in two dimensions: the literatures of using allusions and the lit-

eratures of producing allusions. [Result/Conclusion] The phenomenon of using allusions in ancient Chinese poetry has a long history along with rich topics to be studied, the innovation and transformation of research paradigm are needed urgently at present stage. Through theoretical analysis and empirical results, it can be seen that the citation analysis method is helpful to carry out more in-depth exploration on a large scale of ancient Chinese poetry corpus. The biggest difficulty in current research lies in automatic recognition and extraction of allusions. In the future, it is necessary to strengthen the close combination with NLP technology, and strive to put forward more effective automatic solutions and quantitative analysis strategies in allusion identification and extraction, motivation and behavior classification of using allusions, allusion network and allusion knowledge graph construction, so as to effectively promote the transformation of traditional humanities research paradigm from theoretical analysis to empirical interpretation, from subjective judgment to objective disclosure, from qualitative analysis to quantitative analysis, and produce innovative results on more Digital Humanities (DH) research topics.

Keywords: citation analysis (CA); ancient Chinese poetry; allusion(s); motivation of using allusions; allusion behaviors

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv –Machine translation. Verify with original.