Research on the Evolution of Discipline Knowledge based on Concept Flows

Muzhe Han[1,a,*], Yu Li[2,b]

[1] School of Information Management, Central China Normal University, Wuhan, Hubei Province, China

[2] Sichuan University of Culture and Arts, Mianyang, Sichuan Province, China

[a] hmz_china@mails.ccnu.edu.cn, [b] 532053441@qq.com

Keywords: Concept flow, Discipline knowledge, Evolution, Logistic curve.

Abstract. In order to grasp the law of development and evolution of discipline knowledge from a more detailed level, this paper analyzes the movement form and law of objective knowledge world concept after analyzing the meaning of knowledge and concept. On this basis, a theoretical model of diachronic concept flow is proposed, and its practical significance is briefly described. Finally, taking the interdisciplinary field of "library & computer" as an example, according to the diachronic frequency data of sparse complex network intermediary concept, the diachronic evolution of discipline knowledge is visualized and analyzed, and the existence of sequential concept flow following the logic growth law is confirmed.

1. Introduction

Knowledge (explicit) is an orderly set of conceptual and symbolic information that the cognitive subject adapts and assimilates the information content of the cognitive object with its cognitive schema, and reappears after integration and reconstruction [1]. The basic knowledge unit with complete knowledge expression is called knowledge element [2], which is composed of concepts and corresponding relations. In a specific subject field, the diachronic development and change of concept itself and the relationship between them will continuously promote the evolution of discipline knowledge. Although the factors that cause the motion of concepts themselves and the relationships between them are extremely complex, the concepts in a limited subject field do show a certain diachronic growth law. This diachronic growth law can help us to sort out the attributes of different concepts, and the concept combination of different stages of discipline knowledge may further help to study the diachronic mobility of concepts. After a brief description of the meaning, motion form and motion law of concept, this paper puts forward the hypothesis of temporal and spatial flow of the concept, and analyzes the diachronic evolution mechanism of the discipline knowledge on this basis. Finally, taking the cross field of "library" and "computer" in Library and information science as an example, this paper verifies the existence and rationality of the above hypothesis of concept flow.

2. Foundations

The hypothesis of conceptual flow is based on the World 3 theory, that is, the existence of objective knowledge world. Before elaborating the concept in more detail, the meaning of the concept, the motion form and motion law of the concept should be explained briefly.

2.1 Concept and knowledge

The British philosopher Karl R. Popper divides the universe into three worlds, namely, the physical world (World 1), the subjective knowledge world (World 2) and the objective knowledge world (World 3) [3]. Among them, the World 3 it contains the products of the human mind as recorded in languages, the arts, the sciences, the technologies ---- in all the artefacts humans have stored or scattered around the Earth. Bertram C. Brookes, a British information scientist, built the foundation

equation of information science based on Popper's theory, $K[S] + \Delta I = K[S + \Delta S]$, believing that human cognitive process is reconstructing the new knowledge structure $K[S+\Delta S]$ after integrating the environmental information increment ΔI with the original knowledge structure $K[S]$. This kind of cognitive process is not the physical superposition of the original information reserve and the information increment, but the process of internalizing the original knowledge structure and the new information and reconstructing the new knowledge structure after the judgment, selection and decision of the human brain, which makes the human cognition subjective and fuzzy, so that the traditional mathematical method can never completely solve the related problems [5]. Different scholars in many fields have carried out researches on the structural description of knowledge and the dynamic evolution mechanism of knowledge, among which the representative achievements in the field of Library and information science include Jason Farradane's individual cognitive response model [6], Charles Cole's definition of knowledge structure [7], Ross Todd's cognitive constructivism theory [8]. On the knowledge element theoretical system put forward by Wen Youkui, etc. [2]; scholars on business management mainly develops from the perspective of knowledge dissemination and knowledge evolution in the organization and society, mainly represented by the knowledge spiral mode of Ikujiro Nonaka [9], and the knowledge flow model, which integrates the complex network structure [10].

The above research results are based on knowledge or the vehicle of knowledge. However, under the knowledge element with complete and independent knowledge expression function, there are some concepts with clear ideographic function but incomplete logic. These concepts are often confused with common words, or broadly referred to as data or information, few scholars carry out further research on the concept and the motion of concepts. Brookes views the knowledge as a structure of concepts linked by their relations, and information as a small part of such a structure. The basic knowledge unit with complete knowledge expression is called knowledge element, which consists of concepts and corresponding relations. The concept in the sense of philosophy refers to the abstract and universal idea or idea or the entity that serves as the category or class of entity, event or relationship. The concept in this paper is the basic expression unit with independent ideographic function. Vocabulary or phrase is the most common expression form of concept, but the concept is different from vocabulary. For example, the same word "apple" corresponds to two different concepts when it represents fruit variety and electronic product brand; different words "cellphone" and "mobile phone" usually correspond to the same concept when it represents electronic equipment; while the concept composed of phrase is general. For example, the concept "mobile Internet" is composed of "mobile" and "Internet", but the separated concept is no longer specific. Therefore, a concept can be regarded as a word, phrase, symbol and its combination expression unit fused with certain context logic and semantic connotation. Concepts are also different from knowledge. Although concepts usually have certain logic, they cannot express knowledge independently and completely. Concepts, together with other concepts and data without independent ideographic function, usually constitute knowledge in a certain logical form.

2.2 Motion forms of concept

As mentioned above, as an ideographic entity integrating certain semantics, concepts also have corresponding motion form and motion law in the objective knowledge world. In the course of the development of social science, many philosophy or scientific concepts involved have been constantly division, integration, migration and transformation.

Concept division, a typical example is "knowledge management", which originated in 1950s, is still popular. However, from the previous statistics on domestic research, it can be seen that the research on knowledge management has divided into Management science and Information science from a very early time due to different research purposes. Although it has the same origin and similar basic theories, the concept of "knowledge management" has a wide gap in the key points defined in the two disciplines, and the research topics are not similar. The direct communication between scholars in the two fields is not even as dense as the indirect communication [11]. This conceptual differentiation is also reflected in Keping's monograph on knowledge management, in

which he also divides knowledge management into two fields of "enterprise knowledge management" and "social knowledge management" for discussion [12]. Sometimes, this phenomenon of concept division extends from the same theme to two different directions of communication, and has their own independent trend. From the perspective of D. Price's view of science history, the differentiation of concepts is an important reason for the formation of "Big Science" system in basic science.

Concept fusion is the opposite of concept division. Different concepts are fused into new concepts due to certain needs, which corresponds to what price called integration. Such cases abound, such as "physicochemistry" formed by the combination of "physics" and "Chemistry", which is not redundant here.

Concept migration, which refers to the phenomenon that a certain concept is migrated from one discipline to another and expresses different meanings on the premise of retaining part of its connotation. Like the concept "entropy" used to represent the disorder state of system in thermodynamics is introduced into information field by Shannon to measure the uncertainty of things. Similarly, the philosophy concept "ontology" has migrated to informatics and computer science as some kind of formal representation.

Concept transformation could also be understood as evolution. The connotation of the two concepts may not have changed much, but there is some extension or change. After the transformation, the concepts tend to inherit or replace the position of the original concepts, sometimes the two concepts will also continue to coexist as a synonym. For example, with the measurement granularity becoming finer, the coarse-grained bibliometrics gradually changes to the medium-grained informatrics and the fine-grained knowmetrics.

2.3 Motion law of concept

Literature is the vehicle of knowledge, and concept is the basic ideographic unit of knowledge. The diachronic motion law of literature can reflect the motion law of concept to a certain extent. In order to explore the motion law of discipline concepts, the author verifies the unity of exponential growth law, cubic polynomial growth law and logical growth law with the diachronic growth accumulation law of hot concepts in Library and information science, and the effectiveness of logical growth law in the description of information growth or knowledge growth [13].

Logistic function is a mathematical model put forward by P.F. Verhulst, a Belgian mathematician, when he studied the relationship between population growth and resource carrying capacity. Its curve shape is S-shaped, and it is called logical function because of its rich logical reasoning meaning [14]. The research shows that the cumulative law of the diachronic growth of discipline concept obeys the law of logical growth, and its mathematical expression is as follows:

$$f(t) = \frac{K}{1 + ae^{-bt}}, a > 0, b > 0$$

In the above formula, t is an independent variable, representing time; $f(t)$ is a dependent variable, representing the amount of literature at time t; $K = \lim_{t \to \infty} f(t)$, that is, the expected value of the total cumulative amount of literature in the target area; a and B are positive parameters.

The logical function is a monotonic increasing function with the growth trend of "slow-fast-slow" three segments. The derivation of the logical function (formula 1) can obtain the speed function of the logical function. Its expression is as follows:

$$v(t) = f(t)' = \frac{Kabe^{-bt}}{(1 + ae^{-bt})^2}$$

Continue the second derivative of the velocity function (formula 2) and make the second derivative 0. Three key points in the growth process of the logic function can be calculated, which are two inflexion points and one extreme point of the velocity function.

The abscissa of the extreme point is:

$$t_H = \frac{\ln a}{b}$$

The abscissa of the inflection point is as follows:

$$t_1 = \frac{\ln a - 1.317}{b}, t_2 = \frac{\ln a + 1.317}{b}$$

These three key points correspond to the beginning, peak and end periods of social and economic growth reflected by the logic function, and the two inflexion points can divide the growth process into three stages: gradual increase$(0, t_1)$, rapid increase (t_1, t_2) and slow increase$(t_2, +\infty)$. At the same time, there is $f(t_H) \approx 0.5K$, that is, when the cumulative value of growth reaches half of the expected scale, the growth rate of logic function is the fastest, and then its growth rate begins to slow down gradually.
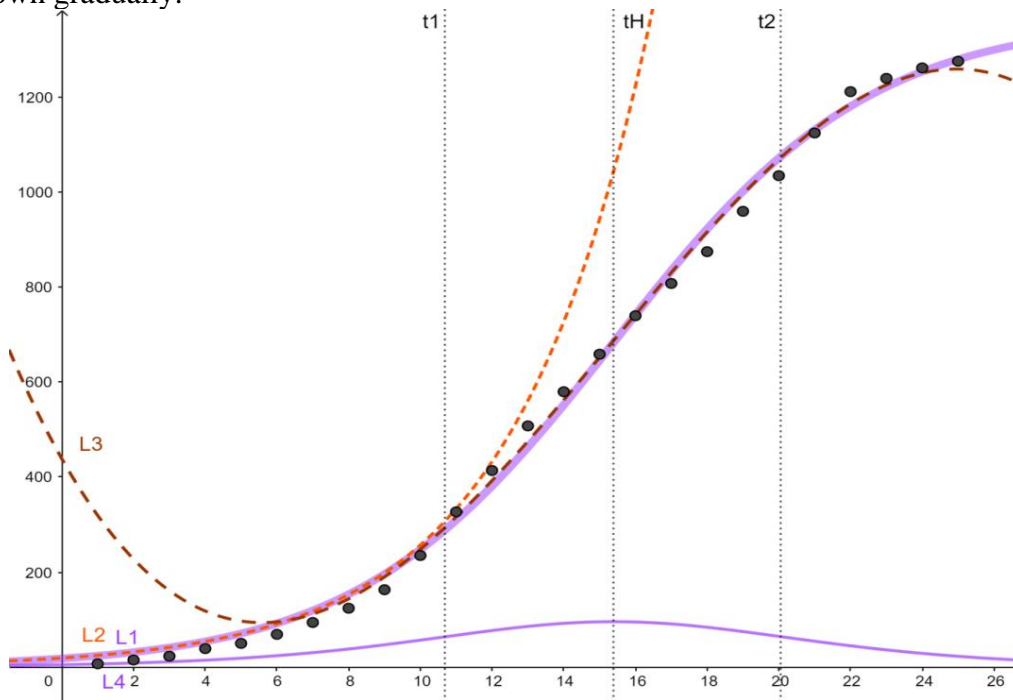

Fig. 1. Decomposition and fitting diagram of logical growth curve

L1: the overall fitting curve of the concept "library management", the result is the logic curve f (T) = (1366.66) / (1 + 74.32e ^ (- 0.28t)), (1 $\leqslant$ t $\leqslant$ 26);

L2: the data fitting curve of the first 11 years of the concept "Library management", the result is the exponential curve g (T) = 19.1e ^ 0.26t, (1 $\leqslant$ t $\leqslant$ 11);

L3: the data fitting curve of the 11th to 26th years of the concept "library management", the result is the cubic polynomial curve. Curve H (T) = -0.32t ^ 3 + 14.66t ^ 2-134.84t + 438.96, (11 $\leqslant$ t $\leqslant$ 26);

L4: velocity curve V (T) = (284 39.65e ^ (- 0.28t)) / (1 + 74.32e ^ (- 0.28t)) ^ 2, (1 $\leqslant$ t $\leqslant$ 26).

In the definition domain $[t_1, t_2]$, the corresponding value domain of function $f(t)$ is about$(0.21K, 0.79K)$, that is, in the rapid growth period of the whole logical growth, the cumulative proportion of frequency can account for about 58% of the expected scale. The time point t_1 can be roughly regarded as the watershed between the exponential growth pattern and the logical growth pattern. Once the frequency accumulation reaches 21% of the expected scale, the environmental resistance begins to play a significant role, which makes the frequency growth accumulation unable to maintain the exponential growth momentum, and gradually turns into the logical growth pattern.

For the concept of discipline, this kind of environmental resistance cannot be generally attributed to the material conditions, economic sources, authors' intelligence and other factors like the growth of scientific literature. The presentation of these hot concepts is summarized from the results of group academic research, and is not limited to the material and economic conditions and intention selection of a certain author or author group. The growth and accumulation trend of these concepts corresponds to the concentration trend of relevant knowledge in the corresponding time period, which directly reflects the development status of relevant disciplines and the main problems to be solved. However, every discipline has its own development track and life cycle. The problems faced by a discipline in different times are targeted and phased. In the course of discipline

development, many phased problems will be solved or replaced by more important problems. In the process of solving these problems, as well as the development and change of Science and technology and social environment In essence, the development of the whole discipline is the process of constantly putting forward, analyzing and solving problems, which makes the direct expression of the continuous scale and refinement of scientific research is the process of knowledge accumulation. Therefore, the direct reason that hinders the unlimited growth of these concepts is the decrease or transfer of the research heat of the corresponding problems. The factors leading to the decline or transfer of heat are more complex, mainly due to the development of scientific and technological environment and social environment.

It is worth noting that the above is a quantitative expression of the basic laws of conceptual diachronic movement, mainly for the life cycle of a single concept, without considering the various forms of conceptual movement. The movement forms of concepts need to be combined with the diachronic combination of concepts and the directionality of the movement of concepts in different fields in the same period of time. In this paper, a hypothesis model of conceptual flow is proposed.

3. Concept flow

Then we should consider how concepts gathering into knowledge during the process of discipline development. In this theoretical model, we assume that knowledge will only spread from the highest potential Discipline U to the lowest potential Discipline L. In the first period, the knowledge expression of discipline M is KE_{P1}. After the Concept i and Concept j fused or migrated from Discipline U in the second period, KE_{P1} is becoming KE_{P2}. Then, if the Concept m contained in KE_{P2} transformed to a new concept m' during the third period, it will become KE_{P3}. And if a concept from KE_{P3} divided a part to Discipline L, it will also make difference to the knowledge expression of the target discipline.
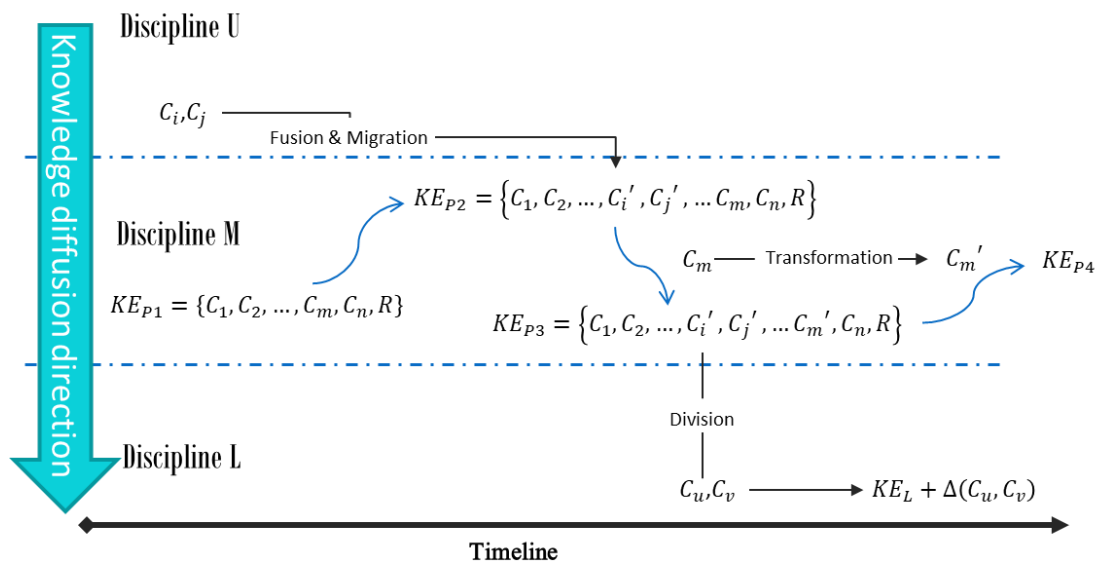

Fig. 2. Motion forms of concept and the knowledge elements they gathered

If we regard each form of a concept as a single node, take time as the X axis, and regard its displacement and transformation as the connection form, we will build a concept flow. However, it should be noted that the displacement and transformation of concepts are also a process, so there will also be a transition between different forms of the same concept that have changed. This transition can be counted by its own exposure frequency and the frequency of other linked concepts, and the frequency data will be expressing the height of the Z-axis in the three-dimensional system. According to the motion law of concept, the diachronic frequency data of a single concept can be identified by the speed curve of the logistic curve, so the frequency data displayed on the XZ plane should be roughly in a single peak normal distribution. The directional motion of related concepts

on the XY axis plane can only be displayed through its two-dimensional mapping of the complex network, or learn from the relevant methods of knowledge flow to abstract classification in three connection forms of "sequential connection", "combined connection" and "shunt connection". A diachronic "ocean" composed of conceptual flows representing different concepts and conceptual systems will be generated by connecting each node on the XY plane according to the semantic relationship. Considering the huge amount of data and the difficulty of data arrangement required by this display method, and there is no suitable visual means to display it dynamically and diachronically at present, this paper only displays the theoretical model of conceptual flow with the three-dimensional surface graph of bivariate normal distribution, as shown in Figure 3.
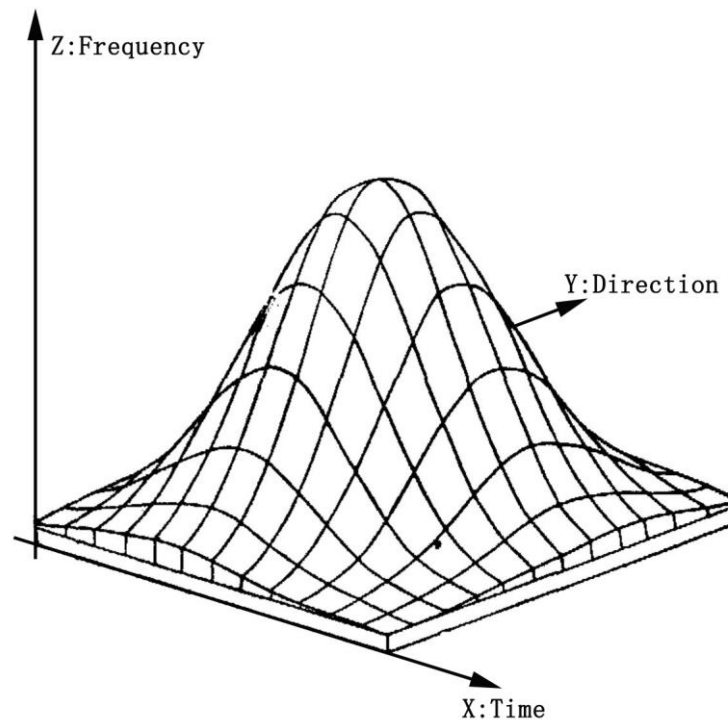

Fig. 3. Effect diagram of diachronic conceptual flow model of single concept

4. Diachronic evolution of subject knowledge based on concept flow

This paper takes "library" as the source, takes "computer" as the target, makes subject slices, analyzes the subject hot concepts extracted from the 1992-2016 papers of Library and information science collected by CNKI, finds the diachronic evolution law of the subject knowledge involved, and verifies the existence of the concept flow.

4.1 Data source and preprocessing

In the database of Chinese periodicals, the retrieval range is set as follows: literature classification catalogue information technology; literature source category is set as core periodical & CSSCI. Set the subject search term "library" and include keyword "computer" (fuzzy), respectively search the data of five groups of 25 years in 1992-1996, 1997-2001, 2002-2006, 2007-2011 and 2012-2016, respectively collect the description information and key words of 100 papers cited the most frequently in each group. BICOMB is used to extract the keywords of five groups of papers. After keyword cleaning, the keyword co word matrix of five groups of data is generated, corresponding to the above five time periods.

4.2 Experimental steps

Bicomb2 | Data collection and pretreatment | Keywords co-occurrence matrix

Ucinet6.0 | Network generation and structural hole detection | Medium nodes (bridge concepts) detective

Regression of logistic curve

GeoGebra | Speed curve generation | Diachronic transition of concepts
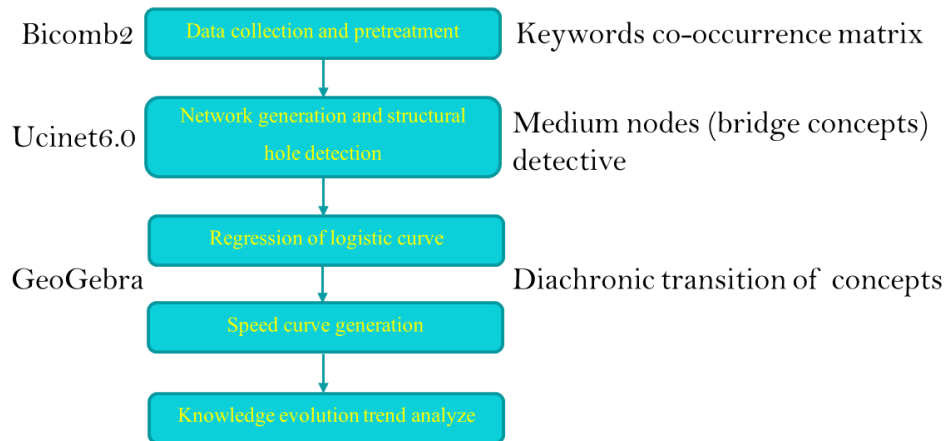
Knowledge evolution trend analyze

Fig. 4. Experimental steps

In complex networks, the existence of structural holes makes intermediary nodes occupy an important position of contact, which can control the flow of network resources to a large extent [15]. Intermediate nodes play an important role in maintaining the existing topological structure of the network. With the increasingly rich connections in the network and the increase of the overall figure density, the number of intermediate nodes that can lead to structural collapse will gradually decrease.

The co-word matrix of the above five time periods is imported into the social network analysis software UCINET, which is transformed into correlation matrix by Spearman coefficient, and the subject concept network of the five time periods is generated. Using the structural hole detection algorithm proposed by Burt in UCINET, we detect each node in five networks respectively, extract the relevant data of eff size, efficie, constraint and hierarc of each node in different stages, and refer to the egonet betweenness proposed by Borgatti. After row comprehensive sorting, the mediation nodes of 5 stages are obtained.

Because this paper is to investigate the diachronic evolution of subject knowledge, we extract 11 concepts that have only appeared once in the intermediary nodes of the above stages, collect their frequency data in the information technology category in a total of 25 years from 1992 to 2016, and calculate the cumulative growth. By using the data analysis software geogebra to fit the cumulative value of each concept's growth over the years, generate the speed curve of each logic curve, and display it in the same rectangular coordinate system, we can get the diachronic change chart of hot concepts in the interdisciplinary field of "library & computer" in the past 25 years.

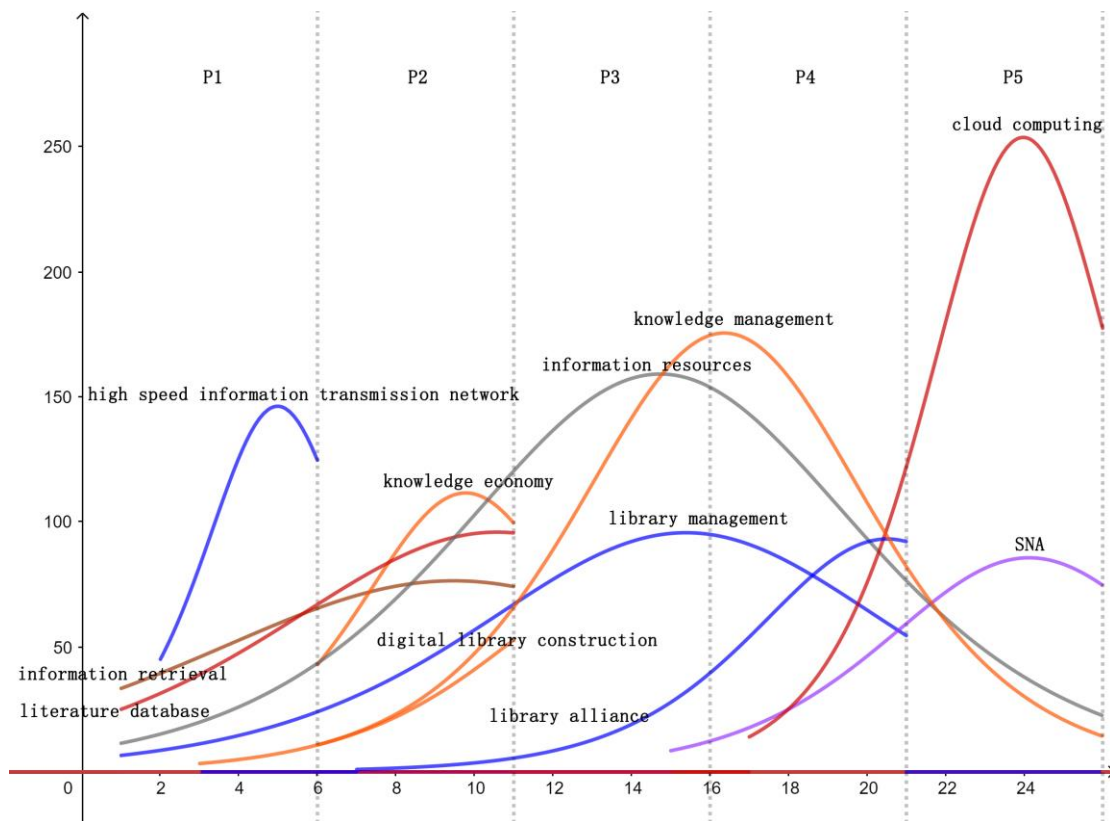## 4.3 Analysis of experimental results

Fig. 5. Diachronic change of the hot concept of "library & computer"

Figure 5 shows the trend of subject knowledge evolution made by taking "library & computer" as an example. The section thickness between "library" and "computer" is relatively low, so the directional motion in the XY plane may not be considered. From the perspective of the display effect of XZ plane, the frequency of concepts are generally similar with the single peak normal distribution, overlapping and changing in time order, and the overall trend is wavy like, which is basically consistent with the effect of the concept flow theoretical model expected above.

This is the time sequence transition diagram of 11 bridge concepts generated in this experiment. The "High speed information transmission network", "knowledge economy", "library management" and "cloud computing" have been migrated or transferred during the developing of the discipline. The "information retrieval", "literature database", "digital library construction", "library alliance" and "SNA" used to flow across disciplines by concept division and concept fusion.

In conclusion, this experiment has successfully verified the real existence of temporal conceptual flow. The experimental results better reflect the diachronic evolution law of the knowledge of the target discipline. The relevant methods help to grasp the development context of the target discipline intuitively. When the number and representation of samples are sufficient, it can further grasp the knowledge diffusion and knowledge absorption of the target discipline, and may help to complete the hot spot detection work at the front of the discipline.

5. Summary

After tell the meaning of knowledge and concept, this paper focuses on the movement form and law of objective knowledge world concept. On this basis, a theoretical model of diachronic conceptual flow is proposed, and its practical significance is briefly described. Finally, taking the interdisciplinary field of "library & computer" as an example, the diachronic evolution of subject knowledge is visualized and analyzed based on the diachronic frequency data of sparse complex network intermediary concept. The experimental results confirm the authenticity of the temporal

concept flow following the logic growth law, which is the future research of concept movement direction and quantitative research of knowledge evolution.

References

[1] W. C. Wu, Theroretical Research of Library Science, *Beijing: Higher Education Press*, 2004.

[2] Y. K. Wen, The Knowledge Organization and Searches with the Knowledge Element, *Computer Engineering and Applications*, pp. 55-57, 91, 2005, (1).

[3] K. R. Popper, The logic of scientific discovery. Прогресс, 1983.

[4] B. C. Brookes, The foundations of information science. Part I : Philosophical aspects, *Journal of Information Science*, pp. 125-133, 1980, 2(2).

[5] F. C. Ma, On the quantitative method of Brooks' Informatics, *Information Science*, pp. 1-9, 1983, (04).

[6] J. Farradane, The Nature of Information, *Journal of Information Science*, pp. 13-17, 1979, (1).

[7] C. Cole, Operationalizing the notion of information as a subjective construct, *JASIS*, pp. 465-476, 1994, (7).

[8] R. J. Todd, Back to our beginnings: Information utilization，Bertram Brookes and the fundamental equation of information science, *Information Processing & Management*, pp. 851-870, 1999, 35.

[9] I. Nonaka, T. Yamanouchi. Managing innovation as a self-renewing process, *Journal of Business Venturing,* vol.4. pp. 299-315, 1989.

[10] Z. G. Hai, A knowledge flow model for peer-to-peer team knowledge sharing and management. *Expert Systems with Applications*, vol.23, pp. 23-30, 2002(1).

[11] M. Z. Han, Research on Knowledge Diffusion Based on the Citation Network, *Rizhao: Qufu Normal University*, 2016.

[12] P. Ke, Knowledge Management Science, *Beijing: Science Press*, pp. 3-14. 2007.

[13] J. S. Gao, M. Z. Han. Validation on the Growth Law and Attribute Sorting of Field Hot-concepts: Taking Domestic Library and Information Science as an Example, *Library and Information Service*, pp. 44-54. 2019(10).

[14] P. F. Verhulst. Mathematical investigations on the law of population growth in Annotated Readings in the History of Statistics. *Berlin: Springer*, pp. 69-75, 2001.

[15] R. S. Burt. Structural holes and good ideas1. American Journal of Sociology, vol.110, pp.349-399. 2004(2).