Research on Engineering Software Data Formats Conversion Network

Wenbin Zhao
School of Instrument Science and Engineering, Southeast University, Nanjing, Jiangsu, 210096, China
E-mail: zhaowb19851015@gmail.com

Zhengxu Zhao
School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, Hebei, 050043, China
E-mail: zhaozx@stdu.edu.cn

Abstract—**Disperse engineering information constitutes a domain of complex relationship, into which researching with complex network could contribute significantly to its organization and preservation. A complex network which takes engineering software data formats as nodes is constructed and its statistical characteristics, such as directed edge or undirected edge, the node number, the edge number, average path length, clustering coefficient and node degree distribution, etc, are analyzed in this paper. A Voronoi diagram based complex network visualization and retrieval method is provided. According to statistical characteristics of constructed network, visualization method calculates network nodes' position in a two-dimensional plane and divides this plane into Voronoi diagram by feature nodes of network nodes' clustering and network nodes. Retrieval method restricts compared nodes to nodes in Voronoi domain of feature nodes which are more related to query node, the number of compared nodes is reduced in retrieval process. The experiment result indicates applying this method into data formats conversion path retrieval can ensure retrieval precision and improve retrieval efficiency, thus provide reliable basis for migrating numerous disperse engineering information.**

Index Terms—**information preservation, data migration, format conversion, complex network, Voronoi diagram**

I. Introduction

Along with the rapid development of computer and network technology, information increases more quickly than demand for information [1][2]. In engineering fields, such as automobile, shipbuilding, aerospace, military technology, etc., for maintenance, troubleshooting and retrofit in product lifetime, original design data and engineering information in production process must be persevered for a long periods of time. At present, according to STEP ISO 10303 Standard, the information of design and manufacture was stored in three-dimensional CAD model by some information systems, such as CAD, CAE, CAM, PDM, etc. [3]. But these systems' lifetime is generally shorter than their products' lifetime, some years later most of systems will no longer exist. So the use of information will encounter the compatibility problem between data and systems. On the other hand, take ASD-STAN's LOTAR project for example, it mainly considers how to cope with the increase of information quantity and disperstiveness in enterprise management, product design and production process, meanwhile in order to meet various requirements of different users in different periods, it mainly researches the generation and long-term preservation of engineering information. The manufacturing enterprises accumulate numerous engineering models of similar three-dimensional CAD model every year, these models are write-protected and stored with their own data formats. To ensure the long-term preservation and data reliability, these models must be regularly inspected, migrated and transformed. But these processes not only have compatibility problem, but also encounter efficiency problem of information process. For these two problems, there are no feasible solution and technology both at home and abroad. To realize compatibility of engineering information formats, engineering software data formats conversion should be considered. A complex network which takes engineering software data formats as nodes is constructed and its statistical characteristics which include directed edge or undirected edge, the node number, the edge number, average path length, clustering coefficient and node degree distribution, etc, are analyzed in this paper. A Voronoi diagram based complex network visualization and retrieval method is provided. According to statistical characteristics of constructed network, visualization method calculates network nodes' position in a two-dimensional plane and divides this plane into Voronoi diagram by feature nodes of network nodes' clustering and network nodes. Retrieval method restricts compared nodes to nodes in Voronoi domain of feature nodes which are more related to query node, the number of compared nodes is reduced in retrieval process. The experiment result indicates applying this method into data formats conversion path retrieval can ensure retrieval precision and improve retrieval efficiency, thus provide reliable basis for migrating numerous disperse engineering information.

Copyright credit: This work was supported by National Natural Science Foundation of China (No. 60873208), corresponding author ZHAO Zheng-xu.

doi:10.4304/jsw.7.11.2606-2613

II. Related Work

A. Engineering Software Data Formats Conversion

Engineering information have some characteristics such as complex form and scattered content, so structure relevance and data compatibility directly influence the management and use of digital engineering information resources. Especially in the long-term safety preservation and reliable invoke, there are some problems such as compatibility between information and system, efficiency of information inspection, migration and conversion. Using complex network research into the complicated relationship between engineering information could contribute significantly to its organization and preservation. Zhao studied complex relationship between data formats with small world model, established a data formats conversion network graph which takes software-supported data formats as nodes and software as edges, and analyzed statistical characteristics of complex network, such as average path length, clustering coefficient and node degree distribution [4-7].

B. Voronoi Diagram

Voronoi diagram is one of typical problems in computational geometry and one of basic data structures about space neighboring relation. It was put forward by G.F.Voronoi, Russian mathematician when he researched the neighboring problem. Voronoi diagram is one of data structures close to natural phenomenon, so it is applied to some fields which are related to geometrical information, such as GIS, geometrical reconstruction, image processing and pattern recognition, physics, chemistry, etc. [8][9][18]. But there are few researches in some fields which are unrelated to geometrical information, such as information space partition. René and Stanislav analyzed the research statue of information space partition, and pointed out two important evaluation criterions, including the position of nodes and the ratio of area weight [10]. Chen and Schuffels, etc provided an ET-map method which use different position and color to represent web pages [11]. Andrews and Kienreich, etc put forward an InfoSky method which takes news articles as 2-dimendional information space, and use weighted Voronoi diagram to divide this space [12]. An adaptive weighted Voronoi graph partition method is proposed and used to divide the web information space [10].

III. Complex Network and Voronoi Diagram

A. Complex Network

1) Presentation of Complex Network

Network can be expressed as graph mathematically, so complex network also can be described with language and symbol in graph theory. It can be defined as a graph G= (V, E) which is compose of the node set V and the edge set E, where the number of nodes is $N=|V|$, the number of edge is $M=|E|$. Each edge in E is corresponded to a pair of nodes in V[13].

According to the difference of network edge definition, network can be defined as directed network graph and undirected network graph, weighted network graph and un-weighted network graph, as well as simple graph and complete graph in graph theory. Complex network mainly researches the relationship between microscopic characteristics of nodes and edges in the network and macroscopic properties.

2) Average Path Length

The distance between two nodes in network is the edge number of the shortest path which connects them. The diameter of network is the maximum of distance between any two nodes. Average path length is defined as the average of distances between all nodes, also called characteristic path length.

$$L = \frac{1}{\frac{1}{2}n(n-1)} \sum_{i<j} d_{ij} \tag{1}$$

Average path length describes the separation degree between nodes in network. In complex network, most real networks own smaller average path length, it is small-world property.

3) Clustering Coefficient

Clustering coefficient is used to measure the aggregation of nodes in network. Suppose node vi is connected with ki nodes by ki edges. There can be maximum ki (ki-1)/2 edges between ki nodes, and the actual number of edges is EI. Clustering coefficient Ci of node vi is defined as the ratio of network edges' actual number to its maximum.

$$C_i = \frac{E}{k_i(k_i-1)/2} \tag{2}$$

Clustering coefficient C of the whole network is the average of all nodes' clustering coefficients, and $0 \le C \le 1$. Clustering coefficient represents distance relationship between all nodes in the whole network.

4) Degree and Degree Distribution

The degree of node is defined as the number of edges which connect this node in graph theory. In directed network the degree of node is divided into out-degree and in-degree. Out-degree is the number of edges which is directed from this node to other nodes, in-degree is the number of edges which is directed from other nodes to this node. In directed network the degree of node equal to the sum of out-degree and in-degree. The average of all nodes' degree is defined as the average degree of network, recorded as $<k>$. The distribution of nodes' degree is described with function p (k), its meaning is the probability that an arbitrary node is connected with k edges, and equal to the ratio of nodes' number which is connected by k edges to nodes' total number.

In order to avoid error caused by smaller network, P (k) is used to represent the cumulative distribution function of p (k).

$$P(K) = \sum_{k'=k}^{\infty} p(k') \tag{3}$$

P (k) represents the probability distribution of nodes whose degree are not less than k. Degree is a basic parameter which describe structural characteristic of network nodes, and reflects macroscopic statistical property of network system.

There are some statistical characteristics in complex networks, such as the connectivity of network, the relevance of nodes, modularity, etc.

B. Voronoi Diagram

Let p and q are two discrete points in plane, L is the perpendicular bisector of line pq and divides the plane into two parts L_p and L_q. Points in the L_p satisfy $d\ (x, p)< d\ (x, q)$, where d is Euclidean distance. L_p is expressed as $dom\ (p, q)=\{x \in R^2 | d\ (x, p)<d\ (x, q)\}$, and L_q is expressed as $dom\ (q, p)$, as shown in Figure 1.

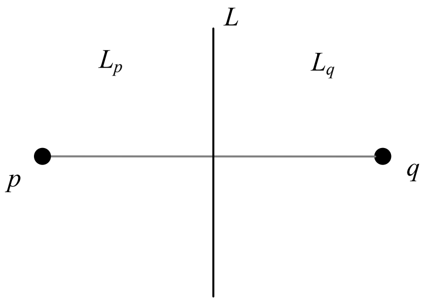


Figure 1. Diagram of L_p and L_q.

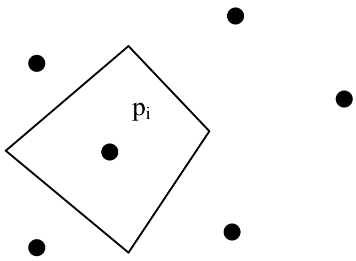


Figure 2. Voronoi polygon of p_i.

Let S be a set of n points in plane, $S = \{ p_1, p_2, ..., p_n \}$. $reg\ (p_i)=\cap dom\ (p_i\ , p_j)$, where $p_j \in S - \{ p_i \}$, $reg\ (p_i)$ is more close to p_i than other points, and is the intersection of n-1 half plane. It is a convex polygon that has no more than n-1 edge, which is called p_i's Voronoi domain or Voronoi cell[8], As shown in Figure 2.

Each point in set S may be corresponding to a Voronoi domain, the diagram of n Voronoi domains is called Voronoi diagram, $V\ (S)$, as shown in Figure 3. The vertex and edges are respectively called Voronoi vertex and Voronoi edge.

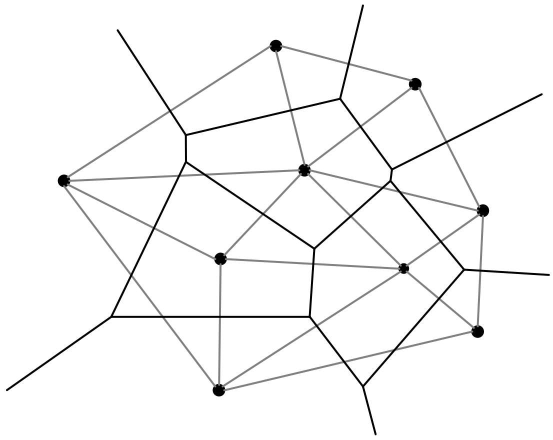


Figure 3. Voronoi Diagram.

IV. ENGINEERING SOFTWARE DATA FORMATS CONVERSION NETWORK

Engineering software data formats conversion network is a complex network which takes engineering software data formats as nodes and is based on the conversion relationship between data formats. It is defined as G= (V, E), which expresses the topological relationship between data formats. $V=\{v_i | i=1, 2, …, n\}$ is the node set of engineering software data formats, If data format $v_i \in V$ can be converted to data format $v_j \in V$, a directed edge e (v_i, v_j) is created in network graph. Matrix W is the adjacency matrix of network graph G. If a directed edge exists between v_i and v_j, w_{ij}=1, otherwise w_{ij}=0.

TABLE I.
ENGINEERING SOFTWARE AND THE NUMBER OF DATA FORMATS

Software	Data Formats	Software	Data Formats
AdobePhotoshop	124	MATLAB	44
CorelDRAW	122	MicrosoftVisio	42
Pro/E	106	MicrosoftOffice Access	40
Macromedia Dreamweaver	93	PLSQL Developer	32
MicrosoftOffice PowerPoint	85	Microsoft VisualBasic	25
SolidWorks	75	MIS	21
CATIA	68	WinRAR	15
ACDSee	68	ANSYS	13
Geomagic	67	VRPlatform	13
Nero	66	Virtools	12
Microsoft VisualC++	65	Queat3D	12
3dsMax	62	CAJViewer	12
UGSNX	61	LabVIEW	10
MicrosoftOffice Excel	60	Microsoft SQLServer	10
Macromedia Flash	60	EndNote	10
AdobeFlash	59	AdobeReader	8
MicrosoftOffice Word	53	FoxitReader	3
MacromediA Fireworks	53	Notepad	2
AutoCAD	48		

This paper selects two-dimensional and three-dimensional graphic design software, engineering simulation software and data document processing software and statistics their input and output formats. Table 1 shows engineering software and the number of their data formats.

According to statistical engineering software and data format conversion relationship, engineering software data format conversion network is constructed, as shown in Figure 4. In the network constructing process, only software and data formats are considered, data storage duration, software lifetime and data format term are considered as ideal condition, namely at the same time.

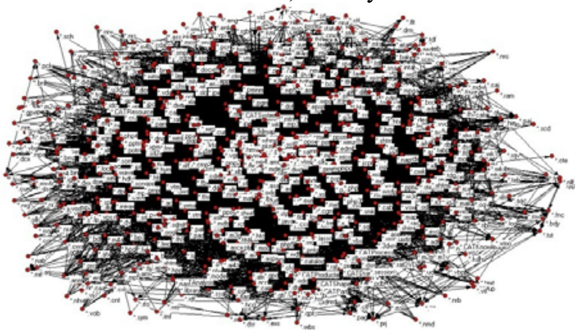


Figure 4. Engineering Software Data Formats Conversion Network

This paper analyzes engineering software data format conversion network by the statistical characteristics of complex network, such as average path length, clustering coefficient and degree distribution. Table 2 show the number of nodes, average path length, clustering coefficient and network diameter of engineering software data format conversion network.

TABLE II.
STATISTICAL CHARACTERISTICS OF ENGINEERING SOFTWARE DATA FORMATS CONVERSION NETWORK

The number of nodes	Average Path Length	Clustering Coefficient	Diameter
54	1.48186	0.49968	2
82	1.55829	0.57296	2
106	1.73248	0.54861	3
131	1.81549	0.54213	3
157	1.85644	0.52463	3
186	1.87075	0.57556	4
223	1.88114	0.63686	4
268	1.88823	0.62332	4
306	1.93619	0.58704	4
346	1.97497	0.57292	4
378	2.03986	0.56924	4
397	2.06192	0.56428	4
416	2.07654	0.57589	4
431	2.12055	0.58968	4
456	2.13339	0.60699	4
473	2.13819	0.60957	4
488	2.13865	0.61743	4

Figure 5 show the relationship between average path length and the number of nodes. The increase of become gradually slow, it indicates engineering software data format conversion network follows small world property, average path length grows at logarithm type or lower logarithm type with network size or number of network nodes.

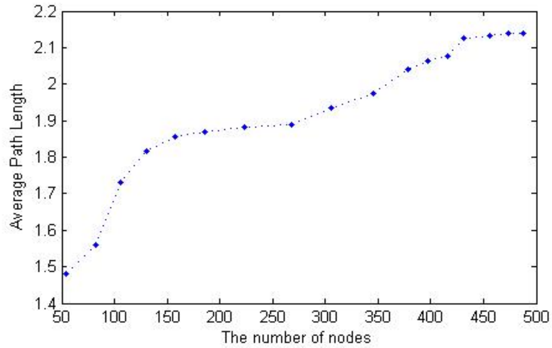


Figure 5. Relationship between Average Path Length and Nodes' Number

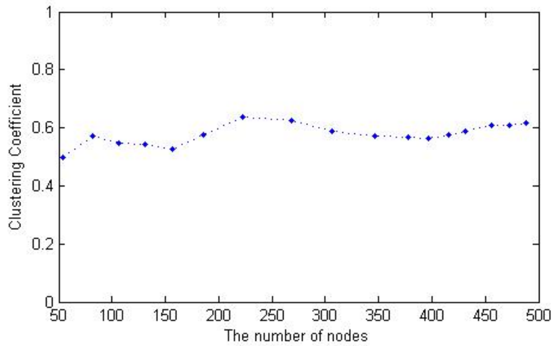


Figure 6. Relationship between Clustering Coefficient and Nodes' Number

Figure 6 show the relationship between clustering coefficient and nodes' number. When the number of nodes increases gradually, clustering coefficient of network fluctuates in small-scale. Therefore, clustering coefficient is irrelevant to network scale, and it is consistent with WS model and Newman model.

Figure 7 show the probability distribution of nodes' degree. Nodes, whose degree is less than or equal to 180, are account for about 90% of the total number of nodes. The average of nodes' degree is 63.98. The degree of nodes can express the importance of nodes in a sense. In this experiment, bmp, jpg, gif, pcx and tif have higher importance. In order to take into account the importance of these nodes, the cumulative probability distribution is used to describe nodes' degree, as shown in the Figure 8. Nodes which have higher degree are less, but they are the key node of network.

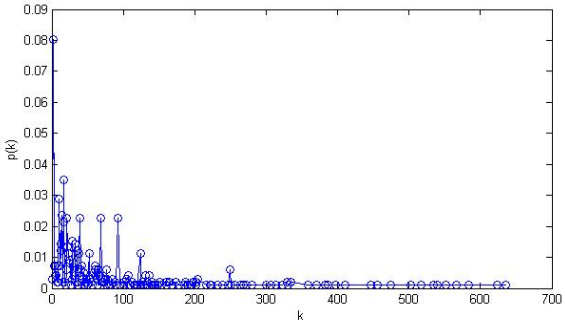


Figure 7. Distribution Curve of Nodes' Degree

In Figure 8, the cumulative degree of nodes in engineering software data format conversion network follows exponential distribution. Through curve fitting, the result is as shown in Figure 9. Nodes' cumulative degree follows exponential distribution curve P (k)=0.7502e$^{-0.0199k}$.

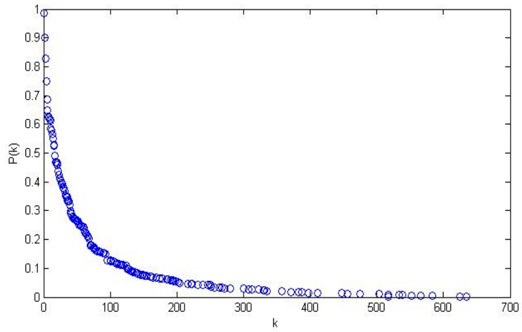


Figure 8. Distribution Curve of Nodes' Cumulative Degree

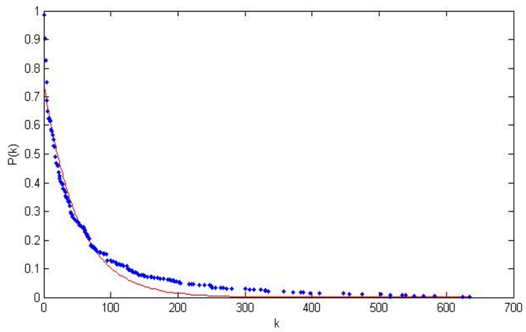


Figure 9. Fitting Curve of Nodes' Cumulative Degree

In semi-logarithmic coordinate, the distribution of these nodes' cumulative degree is approximately considered as a straight line, but in double-logarithmic coordinate, it is not considered as a straight line. So engineering software data format conversion network follows exponential distribution, but does not follow power-law distribution, so it does not belong to the scale-free network.

V. Voronoi Diagram based Complex Network Visualization and Retrieval

A. Voronoi Diagram based Complex Network Visualization Method

1) SPH Layout Algorithm

Voronoi diagram based complex network visualization needs to determine the position of network nodes in the two-dimensional plane, and divide this plane into Voronoi diagram. With SPH layout algorithm, visualization method arranges network nodes by the statistical characteristics of complex network and determines their position in the two-dimensional plane with point-by-point insertion method.

SPH algorithm is a widely used mesh-free method. It has been used in many fields of research, including astrophysics, ballistics, volcanology, and oceanography. Its basic idea is that the fluid is divided into a set of discrete particles. These particles have a spatial distance, over which their properties are "smoothed" by a kernel function. This means that the physical quantity of any particle can be obtained by summing the relevant properties of all the particles which lie within the range of the kernel[14]. In this paper, the nodes of complex network are considered as a series of particles which own quality, speed and energy, their distribution is calculated in a two-dimensional plane. Discrete SPH function is as follow:

$$u(x) = \frac{1}{N}\sum_{i=1}^{n}[u(x_i) + w(x - x_i, h)] \tag{4}$$

Where, w (x-x_i, h) is the interpolation kernel function of the relevance between node x and xi as well as control factor h, h determines the size of the kernel function's influence domain. *B*-splint function is adopted as the kernel function.

$$w(\bar{d}, h) = \frac{1}{\pi h^3}\begin{cases} (1 - 6\bar{d}^2 + 6\bar{d}^3), 0 \le \bar{d} \le 1/2 \\ 2(1 - \bar{d})^3, 1/2 \le \bar{d} \le 1 \\ 0, \bar{d} \ge 1 \end{cases} \tag{5}$$

where, $\bar{d} = \|x - x_i\| / 2$.

After the position coordinates of network nodes are worked out, there may be one problem which these coordinates are too large or too small, so these coordinates can be transformed to an appropriate range by display effect.

2) Feature node

Feature node is a network node that can represent network nodes' clustering in Voronoi diagram based complex network visualization, and it is key node in Voronoi diagram based node-space model retrieval.

Let S is a set of n network nodes, $S=\{o_1, o_2, ..., o_n\}$. V is a set of m network nodes' clustering, $V=\{V_1, V_2, ..., V_m\}$, V_j is a clustering in clustering set, the number of network nodes in V_j clustering is k_j, network nodes in V_j clustering is $V_j=\{o_{1j}, o_{2j}, ..., o_{kj}\}$. If $1 \le j \ne j' \le$ C, $V_j \cap V_{j'} = \varnothing$ and $\sum_j^m k_j = n$ · x_{gm}^j is the mean-value point of network nodes' distribution, feature node of clustering V_j is defined as $CR_j = argmin_{i=1,...,k_j}(Dist(o_{ij}, x_{gm}^j))$, namely clustering's feature node is network node which is nearest to its mean-value point.

For a set with m points $\{x_1, x_2, ..., x_m\}$, $x_i \in R_n$, the mean-value point is defined as:

$$x_{gm} = argmin_{y \in R^n}\sum_{i=1}^{m}\|x_i - y\|_2 \tag{6}$$

The calculation of the mean-value point has no explicit formula, so this paper takes it as optimization problem and solve it by particle swarm optimization algorithm, PSO[15]. PSO algorithm's basic idea is that a group of particles, which have no volume and quality, are randomly initialized, each particle is considered as a feasible solution of the optimization problem, and is controlled by a fitness function which has been set in advance. The fitness function which solves the mean-value point is defined as:

$$f(y) = -\sum_{i=1}^{m} \left\| x_i - y \right\|_2 \tag{7}$$

Each particle moves in feasible solution space, and its direction and distance are determined by a speed variable. Generally these particles follow currently optimal particle and the optimal solution can be achieved by searching. With PSO algorithm the mean-value point of clustering is calculated, thereby the feature node of clustering is determined.

3) Visualization Method

Voronoi diagram based complex network visualization method includes the following steps:

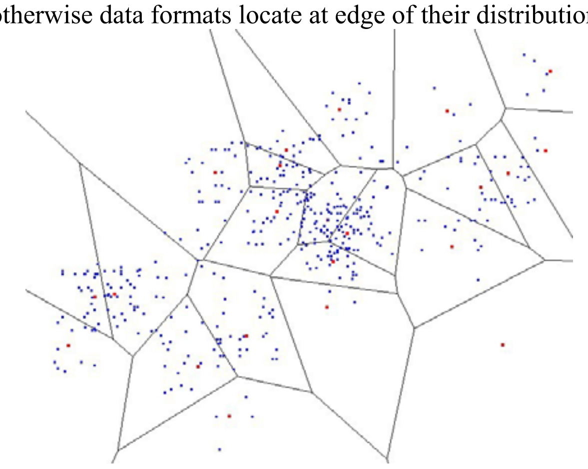
(1) According to engineering software data format conversion network, its statistical characteristics are calculated, such as directed or undirected, the number of nodes, the number of edges, average degree, average path length, clustering coefficient and so on;

(2) Network nodes are ranged by node degree in descending order and their positions are determined with SPH layout algorithm in the two-dimensional plane;

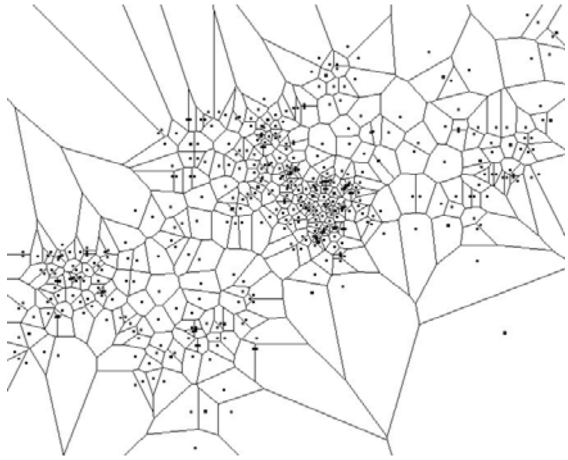
(3) According to the distribution of network nodes, the mean-value points and feature nodes of network nodes' clustering are determined;

(4) The two-dimensional plane is divided into Voronoi diagram by Voronoi nodes as which feature nodes are taken.

According to Voronoi diagram, engineering softwate data format conversion network can be analyzed from several angles. Figure 10 show Voronoi diagram of engineering software data format conversion network. In figure 10, data formats are divided into several clustering, this shows that the proposed visualization method is effective. The distribution of some clustering is denser, and the distribution of other clustering is sparser. This is because sparse clustering contains more data formats which can be interconverted, and conversely data formats are less. In Voronoi diagram the neighboring relationship between data formats can directly reflect their conversion relationship. Data formats which can be converted into more data formats locate at center of their distribution, otherwise data formats locate at edge of their distribution.



a) Feature Nodes based Voronoi Diagram



b) Network Nodes based Voronoi Diagram

Figure 10. Voronoi Diagram of Engineering Software Data Formats Conversion Network

B. Voronoi Diagram based Complex Network Retrieval Method

1) Path Retrieval

For the long-term safety preservation of engineering information, a problem how to realize the conversion between data formats in information migration, namely format conversion path retrieval, must be considered. At present the presented path retrieval algorithm has various types, for example exhaustion search strategy algorithm, heuristic search strategies based branch definition method, the greedy algorithm, climbing method and simulated annealing method, etc. Dijkstra algorithm, which is in 1959 by E.D. Dijkstra proposed [16], is a shortest path algorithm based on greed strategy. Its basic idea is that an path tree is constructed by increasing path length, thus the shortest path from root node to all other nodes is achieved.

Engineering software data format conversion network is a graph *G (V, E)* which has *n* nodes and *m* edge, the starting node is *s*, the weight of edge *w (x, y)* is positive real number. If node *x* is an arbitrary node of network graph, and *P* is a path from *s* to *x*, the path *P*'s weight *w (P)* is defined as the sum of all edge's weight in path *P*. The shortest path problem is that path P_0 of minimum weight is achieved from all paths from *s* to *x*, namely $w(P_0)=min\{w(P)\}$. Each node is set a label (d_j, p_j), where d_j is the sum of shortest path's weight from *s* to *j*, and p_j is the former node of the shortest path from *s* to *j*. the terminal node is *e*, the steps of achieving shortest path from *s* to *e* are as follows.

1) The label of the starting node *s* is set as $d_j=0$, p_j is empty, and the labels of other nodes are set as $d_i=\infty$, p_j is empty. The starting node is marked as $k=s$ and other nodes are unmarked.

2) If unmark node *j* is directly connected to marked node *k*, the distance from node *k* to node *j* is checking, and $d_j=min\{d_j, d_j+l_{kj}\}$, where, l_{kj} is the direct distance form node *k* to node *j*.

3) The minimum distance d_i is selected from d_j, $d_i=min\{d_j|unmarked\ node\ j\}$, and node *i* is chosen as a node of the shortest path and is marked.

4) Check whether all nodes are marked. If marked, then exit, otherwise $n=i$, turn to step 2.

The above Dijkstra algorithm is one of the classic algorithms which solve the shortest path problem between two nodes of network in graph theory and it is adopted to retrieve the shortest path of data format conversion in engineering software data format conversion network. Fig. 11 shows the retrieval result of shortest path in data format conversion.

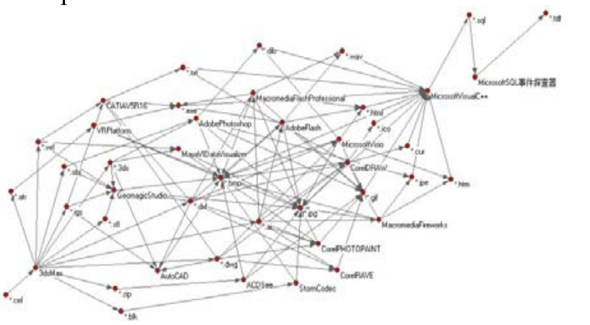


Figure 11. Shortest Path of Data Format Conversion between *.cel and *.tdf

2) Retrieval Method

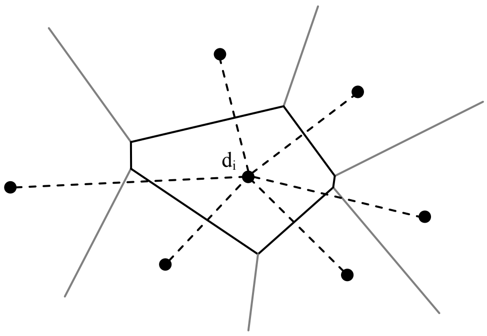
a Voronoi diagram based complex network retrieval method is presented, which can improve path retrieval efficiency by reducing the number of compared nodes. This method is as follows:

(1) Feature nodes are selected from network nodes' clustering, with PSO algorithm the mean-value points are calculated and network nodes which are the nearest to the mean-value point are taken as feature nodes;

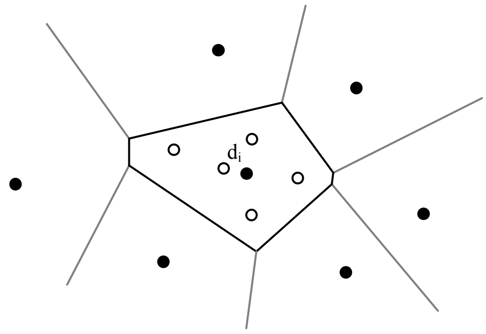
(2) Query node is compared with feature nodes. According to distance between them, feature nodes are ranged in descending order, the former h feature nodes are selected as alternative nodes;

(3) With Voronoi diagram based RNN algorithm[17], network nodes in alternative nodes' Voronoi domain are determined. Figure 12 show alternative node' Voronoi domain and its network nodes;

(4) Query node is compared with network nodes in alternative nodes' Voronoi domain, the most related network nodes are taken as retrieval result.



(a) Alternative Node d_i 's Voronoi Domain



(b) Network Nodes of Alternative Node d_i's Voronoi Domain

Figure 12. Alternative Nodes' Voronoi Domain and Network Nodes

Voronoi diagram based complex network retrieval method need not compare query node with all network nodes, so retrieval task can be accomplished in smaller compare scope. Compared with linear scanning retrieval method, this method can save time that query node is compared with network nodes, thus it is more efficient retrieval method.

At present, the Precision-Recall ratio is commonly used to measure the effectiveness of retrieval method in retrieval fields. In this paper the Precision-Recall ratio is used to compare Voronoi diagram based node-space model retrieval method with linear scanning retrieval method. Figure 13 shows the Precision-Recall ratio of two methods, two curves are almost coincided, so their retrieval performances are almost same.

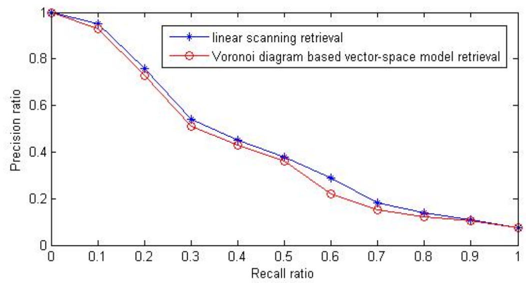


Figure 13. Comparison of Two Methods' Precision-Recall Ratio Curves

The retrieval efficiency of Voronoi diagram based complex network retrieval method and linear scanning retrieval method is comparatively experimented, the former increases 46.23% than the latter.

VI. Conclusion

Engineering information have some characteristics such as complex form and scattered content, so structure relevance and data compatibility directly influence the management and use of digital engineering information resources. Especially in the long-term safety preservation and reliable invoke, there are some problems such as compatibility between information and system, efficiency of information inspection, migration and conversion. Using complex network research into the complicated relationship between engineering information could contribute significantly to its organization and preservation. A complex network which takes engineering software data formats as nodes is constructed and its statistical characteristics which include directed edge or undirected edge, the node number, the edge number, average path length, clustering coefficient and node degree distribution, etc., are analysed in this paper. A Voronoi diagram based complex network visualization and retrieval method is provided. According to statistical characteristics of constructed network, visualization method calculates network nodes' positions in a two-dimensional plane and divides this plane into Voronoi diagram by feature nodes of network nodes' clustering and network nodes. Retrieval method restricts compared nodes to nodes in Voronoi domain of feature nodes which are more related to query node, the number of compared nodes is reduced in retrieval process. The experiment result indicates applying this method into data formats conversion path retrieval can ensure retrieval precision and improve retrieval efficiency, thus provide reliable basis for migrating numerous disperse engineering information.

ACKNOWLEDGMENT

This work was funded by the National Natural Science Foundation of China under Grant No.60873208.

REFERENCES

[1] National Institute of Standards and Technology, "Long term knowledge retention (LTKR): Archival and representation standards", *Gaithersburg: NIST,* http: // edge.cs.drexel.edu/LTKR/, 2006.

[2] Ma Zhanghua, Information Organization, *Beijing: Tsinghai University Press*, 2003, pp.69-77.

[3] National Institute of Standards and Technology, "The Role of ISO 10303 (STEP) in long term data retention: Long Term Knowledge Retention Workshop", *Gaithersburg: NIST*, 2006.

[4] Zhengxu Zhao, Lee Zhuo Zhao, "Small-world phenomenon: toward an analytical model for data exchange in Product Lifecycle Management", *International Journal of Internet Manufacturing and Serv*ices, 2008, 1 (3): 213-230.

[5] Zhengxu Zhao, Wenbin Zhao, "Engineering data formats: Visualization, conversion and migration", *2nd International Conference Intelligent Control and Information Processing (ICICIP)*, 2011, 295-300.

[6] Zhengxu Zhao, Jun Feng, Zhihua Zhang, "Complexity Analysis on Engineering Software Data Format Conversion Networks", *Complex Systems and Complexity Science*, 2010, 7 (1): 75-81.

[7] Zhao Zhengxu, Long Rui, Guo Yang, Liu Jiajia, "Research into Small World Effect in Engineering Software", *Journal of Shijiazhuang Tiedao University (Natural Science)*, 2010, 23 (3): 1-6.

[8] Franz Aurenhammer, "Voronoi Diagrams-A Survey of a Fundamental Geometric, Data Structure", *ACM Computing Surveys*, vol. 23, no. 3, 1991.

[9] R. Zhu, "Intelligent Rate Control for Supporting Real-time Traffic in WLAN Mesh Networks," *Journal of Network and Computer Applications*, vol. 34, no. 5, pp. 1449-1458, 2011.

[10] René Reitsma, Stanislav Trubin, "Information space partitioning using adaptive Voronoi diagrams", *Information Visualization*, 2007, 6 (2): 123-138.

[11] Chen H, Schuffels C, Orwig R, "Internet categorization and search: a self-organizing approach", *Journal of Visual Communication and Image Representation*, 1996, 7: 88-102.

[12] Andrews K, Kienreich W, Sabol V, Becker J, Droschl G, Kappe F, Granitzer M, Auer P, Tochtermann K, "The InfoSky Visual Explorer: exploiting hierarchical structure and document similarities", *Information Visualization*, 2002, 1: 166-181.

[13] Wang Xiaofan, Li Xiang, Chen Guanrong, "Complex Network Theory and its Application", *Beijing: Tsinghai University Press*, 2006.

[14] Liu G. R, Liu M.B, "Smoothed Particle Hydrodynamics: a meshfree particle method", *Singapore: World Scientific*, 2003.

[15] Kennedy J, Eberhart R, "Particle swarm optimization", *Proceedings of IEEE International Conference on Neural Networks, Washington DC: IEEE Computer Society Press*, pp. 1942-1948, 1995.

[16] E.W. Dijkstra, "A note on two problems in connexion with graphs", *Numerische Mathematik*, 1959, 1: 269-271.

[17] Ioana Stanoi, Mirek Riedewald, Divyakant Agrawal, Amr E Abbadi, "Discovery of influence sets in frequently updated databases", *In Proc. Int. Conf. on Very Large Databases (VLDB)*, pp.99-108, 2001.

[18] G.F. Voronoi, "Nouvelles applications des parameters continus à la théorie de forms quadratiques", *Journal für die reine und angewandte Mathematik*, 1908, 134: 198-287.

Wen-bin Zhao received his B.Sc. and M.Sc. degrees in Computer Science from the Shijiazhuang Tiedao University, China, in 2007 and 2010, respectively. His research interests include virtual reality and visualization technology.
E-mail: zhaowb19851015@gmail.com

Zheng-xu Zhao received his PhD in Applied Computing from Staffordshire University, CNAA, UK, in 1992. He had been Professor and Chair in Computer Integrated Manufacturing Systems in School of Computing at the University of Derby, UK, since 1998. His current research interests include computer graphics, virtual environment, software engineering and knowledge retention.
E-mail: zhaozx@stdu.edu.cn