Modelling of Complex Adaptive System of Information Flow

W. Cao, X.H. He
Department of Information Engineering, Academy of Armored Forces Engineering
Beijing 100072, China

X.Q. Fan, G.F. Liu, D.S. Zhou, B. Luo
Beijing special vehicle institute
Beijng, China

***Abstract*—The traditional information systems analysis, basically from a technical point of view, cannot effectively describe the real object systems. From the hierarchical structure of complex adaptive systems point of view, proposed multi-level clustering algorithm based on the flow of information, used to design the structure of complex information systems, reduce the complexity of the system structure; optimize the structure of information systems, the complexity of solving complex information systems issues.**

Keywords-complex adaptive system; information flow modelling; multi-level clustering algorithm

I. INTRODUCTION

C4ISR systems with the development of new information technologies, gradually showing a large, networked, integrated and complex development trend, with a huge, complex interaction of many operational elements and characteristics, the emergence of system uncertainty characteristics of randomness, fuzziness and other complex systems[1][2]. To ensure the feasibility and correctness of the architecture design, reducing the complexity of the system architecture to solve the complex problems of complex information systems, a very important aspect is clearly describe the complex exchange of information between the main body of the military information systems conduct analysis of information flow.

II. C4ISR SYSTEM INFORMATION FLOW

A. Information Flow Connotation

From the perspective of information systems perspective, C4ISR system is a system for operational command information systems on the battlefield, it will be applied to all aspects of IT operations command on the battlefield, the collection, management, processing, dissemination and use of battlefield information for the troops combat Command to provide information support and decision support, or even directly used in combat. C4ISR systems to be able to achieve the necessary exchange of information, which should be able to request information between different C4ISR systems, can be a good transmission. The key to building C4ISR systems is how to achieve the required exchange of information at all levels of demand, one of the core of their concern is the exchange and use of information, namely the flow of information [3].

By modeling the flow of information describing the architecture, the identification C4ISR systems at all levels of information exchange requirements, information exchange between systems interface definition, in order to achieve a seamless transfer of information between the various C4ISR systems and C4ISR systems.

B. The Flow of Information Complexity Factors

C4ISR system as a complex system, the impact of which there are many complex factors, but the main factor is the function and structure of the system. In the C4ISR system, very close link between functions. Some functions need to perform other output functions as a prerequisite, but requires multiple functions to work together in some cases. Because of the close links between the different functions in the design process must select a reasonable logical structure, so that the system can smoothly flow of information, in order to achieve cooperation between functions, to achieve the desired goals.

Command and control system has obvious hierarchy. From the static structure of the system point of view, is the use of the individual soldier to classes, platoon, company, battalion hierarchy, regiment, division, army-based organizational system. Hierarchical command and control system is also under preparation system must be hierarchical in nature. From the dynamic structure of the system, the command and control system in accordance with the scope of the war occurred, can be divided into several levels of conflict, battle, battle, war, etc., and on the low-level structure is an integral part of the high-level structure.

Therefore, the use of multi-level theory to solve complex information systems to adapt to the structural complexity of the problem, and more to reduce the complexity of the system, improve the adaptability of the system.

III. THE INFORMATION FLOW MODELING METHOD

A. Multi-level Theory

Complex adaptive systems theory [4][5] put forward some concepts can reflect the system of multi-level structure, such as aggregation, within the model, building blocks and so on. One of the main causes of complexity is the level of the hierarchy. Generally, on a single level is unable to complete system integration tasks that require the integration step by step through the different levels to be able to eventually form the whole system, and therefore, there must be intermediate level between the elements of the system and the system as a whole. As can be seen, the multi-level organizational structure is an essential characteristic of complex adaptive system, which reflects the system on the complexity of the structure. From another perspective, the complex structure of the single-level multi-level structure can be used to improve the clarity of

structure and can improve system scalability by reducing the degree of coupling between the layers.

As can be seen from the multi-level concepts in the multi-level structure, the interlinkages between the layers, but also the mutual restraint. Therefore, how to divide the hierarchical structure of the system, to form a logical hierarchy is a problem. This paper presents a multi-level algorithm [6] based on weighted directed graph, single-level complex structure will be divided into a multi-information system hierarchy.

B. Multi-level Algorithm Description

The general procedure is agglomerative clustering method [7][8]: first the diagram are initialized each node as a sub-graph, the number of graph nodes, the number of sub-graph form, the design of a sub graph similarity or the distance between the calculation method to similarity or distance between the definition of sub-graph, and finally define the end of the cohesion process conditions, termination conditions may determine the final network hierarchy. After initialization, start calculating the similarity between each sub-graph, merging the highest similarity, the nearest sub-graph into a new, larger sub-graph, and then re-calculate the similarity or distance between all sub-diagram value, will this process is repeated until the end of the process conditions, we can draw the entire hierarchy diagram. Finally, you can use the hierarchy diagram or tree diagram shows these results in order to clearly see the division results obtained under a variety of different needs. The bottom of the hierarchy is the original map of the initial node, level up, indicating that the initial node functions condensed into larger virtual nodes.

IV. MULTI-LEVEL MODELING INFORMATION FLOW

A. Information System Topology Diagram Describes

Static information about the topology of complex systems is adapted to a network, however, the actual flow of information, the importance of each node in the input stream is not uniform on the nodes, and therefore, complex adaptive system topology information is actually a the weighted graph. This figure is abbreviated G (V, E, W, Pr), where V represents a set of graph nodes V = {v1, v2, ... vn,}, n is the number of nodes in FIG. E is a directed graph edge set E = {e11, e21, ... epq}, epq = (vp, vq) connected between said vp to vq side, $|E| = m$ showing the number of edges in the figure. In a directed graph, (vp, vt) said vp of the set, (vt, vp) indicates vp into the set. W = {w11, w21, ... wpq} have the right values to represent each side. In the information flow network, the greater the weight side, said how much weight the amount of information between two nodes, the amount of information, the greater the role in the network, the smaller the weight, the less amount of information. $P_r = \{pr_{11}, pr_{21}, \ldots \ldots pr_{pq}\}$ Represents the probability of the presence of the edge, wherein the$pr_{ij} \in (0,1]$. Uncertainty side view showing the structure of the characteristic information of the flow, the flow of information between two nodes in any case there is not, therefore, need to define the edges of the existence probability. Presence probability and is not relevant to the edge weights, weight does not indicate the presence of a large probability that large.

B. The Process of Multi-level Clustering Algorithm

According to the cluster centers algorithm can get to the lowest set of nodes, cluster centers as a virtual node to become the second level of the initial node and adjacent nodes connected to the gas level circulation cluster center, and the center of the largest cluster probability exists after a node is greater than the probability threshold θ_2 is set to become the second layer node, node set to get the second layer, the node reclassified, cohesion and high node classified as a class. After the second layer of the node classification, each class of the second layer of the node as a new node, the gas level of the node adjacent to the layer 3 will form the set of nodes, and so on, until all nodes are divided layer to form a hierarchical structure graph.

Before hierarchical clustering, you need to define the degree of a node to the cluster center cohesion, in order to summarize the node to the highest degree of cohesion cluster center. Most previous studies have set the minimum distance between nodes as a measure of the merger, the clustering process in this article, mainly in the formation of classes as a measure of the weight standard, which is also known as cohesion values. Cohesion value is to determine whether node quantitative indicators will be included in a class, the greater the degree of cohesion, the closer the relationship, the more likely fall into the class. When the node into a class, the class structure is changed, when the next node to be classified, need to rethink their degree of aggregation. Calculated as follows:

Definition: a set of nodes v_i to C_h exist between the probability of:

$$pr(v_i, C_h) = \max\left(pr(v_i, v_j)\right) \quad (1)$$

Among them $v_j \in H_h$.

A set of probabilities and the calculation method of the same set of probability, that is, the presence of the selected node and the child node in the class the highest probability as the probability of the existence probability of an adjacent node dielectric.

Definition: node v, to the centralization between Class C scale:

$$w(v_i, C_h) = \frac{\sum w(v_i, v_j)}{\sum e_{ij}} \quad (2)$$

Among them $v_j \in C_h$.

The centralization weight calculation method and the method of calculating the same heavy centralization are taking classes in nodes and sub-nodes of the right set of weight average.

The following shows the process of multi-level algorithm algorithm . Given a graph G = (V, E, W, Pr),

$$V = \{v_i\}, i = 1 \ldots n, \quad (3)$$

$$E = \{e_{ij}\}, \ i \neq j, \ e_{ij} \neq e_{ji} \quad (4)$$

$$W = \{w_{ij}\} \quad (5)$$

$$Pr = \{pr_{ij}\} \quad (6)$$

$$C_1 = \{C_1^1, C_1^2 \ldots . C_1^q\} \quad (7)$$

$C_1^i \subset G$ as the initial class,

$$\forall C_1^i, C_1^j \subset C_1, C_1^i \cap C_1^j = \phi \quad (8)$$

Suppose the function F is a set of layered, CF classification function for a set of graphs, Gi of the i-layer is the set of nodes. k is the number of final stratification, qi classification number for each layer , $\theta_i^j$is the first i layer j class existence probability threshold, w_{ij} pr_{ij}respectively from the right side of the node vi to vj of scale and presence probabilities. $\beta > 1$ is the node out of the impact factor.

Each layer has a ki categories

$$H_i = \left\{C_i^1, C_i^2, \ldots \ldots C_i^{k_i}\right\} \quad (9)$$

when i = 1,

$$C_1 = \left\{C_1^1, C_1^2 \ldots . C_1^q\right\} \quad (10)$$

is a cluster center.

The function F i-th hierarchical layer is as follows:

$$F_i = (G, H_{i=1}) = H_i = \{v_i | v_i \subset G/(H_1 \cup H_2 \cup \ldots \cup H_{i=1})\} \quad (11)$$

$$v_i \subset \mathrm{Nei}_{H_{i=1}}^{h_i}$$

$$\max\left(\mathrm{pr}\left(v_i, C_i^j\right), \mathrm{pr}\left(C_i^j, v_i\right)\right) \geq \theta \quad (12)$$

CF j-classification function of the i layer is as follows:

$$CF_i^j(H_i) = \left\{v_i \middle| v_i \subset H_i, \forall v_m \subset H_i, \mathrm{coh}\left(v_i, C_i^j\right) \geq \mathrm{coh}(v_i, v_m)\right\} \quad (13)$$

Wherein,

$$\mathrm{coh}\left(v_i, C_i^j\right) = \sum_{v_j \in C_i^j}\left(\beta w_{ij} + w_{ji}\right), w_{ij} \in \{\alpha_1, \alpha_2 \ldots \ldots \alpha_n\} \quad (14)$$

In summary

$$G = H_1 \cup H_2 \cup \ldots \cup H_k = H_1 \cup F_i(G, H_{i-1}) = H_1 \cup \left(\cup F_2^j(H_1)\right) \cup \left(\cup CF_3^j(H_2)\right) \cup \ldots \cup \left(\cup CF_k^j(H_{k-1})\right) \quad (15)$$

Wherein $i = 2,3, \ldots k$, $j = 1,2, \ldots q^i$,the final level $H = \{H_1, H_2 \ldots \ldots H_k\}$is a set of linear structure, and each layer is only the upper and lower levels of the hierarchy have a connected relationship, that is:

$$\forall H_i, H_j, H_m \subset H, e\left(H_1, H_j\right) = 1, \quad (16)$$

$$e(H_i, H_m) = 1 \rightarrow e\left(H_j, H_m\right) = 0 \quad (17)$$

Figure layered process involves a degree of aggregation node and the like, this level is relative, and the other is to compare the cluster center. value coh (vi, Cj,) the greater the higher the degree of aggregation node and the cluster center. If the node to the cluster center value of the two coagulation is the same, the random selection of one neighbor node. Parameter is the degree of aggregation of the decision weights of the edges; if there is no edge between two nodes, and the existence probability of the weights is zero.

Affect the degree and the degree of a node is not the same degree of cohesion, the impact of the degree will be greater than the penetration. For a node, the call information to the other nodes is called the degree of coupling than other nodes will be higher, so the added cohesion parameter β value calculation.

In the hierarchical clustering process, the sub-graph selection as the cluster center full connectivity with high accuracy. Fully connected subgraph represents the region with the highest degree of polymerization of the network and do not consider the edges in the region has anisotropy, which is connected only to the extent of considering, since the average weight and the existence probability of existence of the degree of aggregation is not vulnerable identical the cluster center, the polymerization process more accurate. In the hierarchical clustering process, consider a set of nodes and the set of weights as well as the existence of probability, the combined weight of each node, class of weight and presence probability will change, the degree of aggregation of each node to the cluster center will also changes. Therefore, for exactly the same with other nodes in the cluster center of the agglomeration, there will be more opportunities to correct division.

In the multi-level algorithm, the high-level node is h layer neighbor node low level node set, a set of nodes in each layer clustering results are not affected by the low level of clustering results. If a layer by adding a new node and re-set the layer node cluster, and will not affect the collection of nodes and clustering results of the layer of the upper or lower. Therefore, the multi-level structure algorithm makes high adaptability, which can better support the dynamic clustering layer.

V. SUMMARY

In complex adaptive systems theory as the theoretical basis, from the hierarchical structure of complex adaptive systems point of view, put forward the information flow modeling approach, taking into account the information flow diagram of information uncertainty, and strong connectivity information nodes proposed multi-level clustering algorithm first obtain figure based on connectivity between the nodes of the cluster center, the cluster centers as the initial class of the second layer, the depth of the initial class of gas-class collection of nodes adjacent nodes form a second layer, and then according to the degree of aggregation of the cluster nodes, the class formed by the initial node is the higher level. So, eventually forming a multi-level structure, used to design the structure of complex information systems, reduce the complexity of the system structure, optimize the structure of information systems to solve complex problems of complex information systems.

REFERENCES

[1] Xu.Dong.Wang, Xue.Wang, Ye.Guan. Strategic delivery characteristics of complex systems [J]. Military Jiaotong University, 14 (11) , 2012.

[2] He.Yin.Hou. Characteristics of complex adaptive systems and sustainable development issues [J] Systems Science PKU CSSCI, 16 (4) DOI: 10.3969 / j.issn.1005-6408.2008.04.014, 2008.

[3] Yan. Long, Su.Yi.Huang, Ke.Liu. The basic features of the large system of material flow, energy flow and information flow [J] Huazhong University of Science and Technology: Natural Science Edition, (12): 87-90, 2008.

[4] Zhou.Yan.Wang, Huan.Chen.Wang. Complex adaptive system [J] Science and Management, (11): 50-52, 2000.

[5] Wei.Ou.Complex adaptive systems theory in combat simulation [J] Journal of Sichuan Ordnance, (8): 129-131, 2011.

[6] Sheng.Jun.Qin, Yan.Chen, Ming.Yang. Research and application of multi-level modeling of complex information systems [J]. Dalian

Maritime University, (4) , 2012.

[7] Jian.Fei.Shi, Huai.Zhi.Yan, Zhan.Yun.Niu. Improved cohesion based hierarchical clustering algorithm [J]Beijing Institute of Technology, (1) , 2008.

[8] Xing.Bo. Liu. Research condensed type hierarchical clustering algorithm [J]. Technological information (science and research), 2008.