Xiaohui KUANG[1], Gang ZHAO[1], Yong TANG, Jin LI[1]

Science and Technology Laboratory on Information System Security, Beijing Institute of System Engineering, Beijing, China (1)
College of Computer, National University of Defense Technology, Changsha, Hunan, 410073, China (2)

Structural Vulnerability Analysis for Large-Scale Distributed System based on Multi-layer Topology Model

Abstract. *Vulnerability analysis of LSDS (Large-Scale distributed systems) has become a growing focus nowadays. A new multi-layer topology model is proposed to describe complex relations between LSDS entities, redundancy mechanism and failure tolerant mechanism in LSDS. Based on this model, a structural vulnerability analysis algorithm based on weight is proposed, which can discover and validate structural vulnerabilities by computing weight of every entity and link with by graph pruning. After that, a new multi-topology model structural vulnerability analysis algorithm is proposed, which can detect the critical entities on different layers of LSDS according to the interdependencies between them. In the end, the efficiency of the algorithms is verified by experiment.*

Streszczenie. *Zaproponowano nowy wielowarstwowy model topologii systemów wielkiej skali rozproszenia LSDS uwzględniający mechanizm błędów i redundancji. Na podstawie tego modelu przeprowadzono analizę wrażliwości systemu. (**Analiza strukturalnej wrażliwości systemów o wielkiej skali rozproszenia LSDS)***

Keywords: large-scale distributed system (LSDS); structural vulnerability; topology module; fault tolerance; vulnerability analysis
Słowa kluczowe: układy wielkiej skali rozproszenia LSDS, topologia

1. Introduction

Along with the continuous development of information technology, Large-Scale Distributed Systems (LSDS), such as grid, P2P system, DNS, GIG and cloud facility[1], has been widely applied as key information infrastructures in the commercial, military and government fields. LSDS significantly enhances resource sharing and coordination capabilities by integrating high-speed network, computers, data bases and applications. However, because of some characteristics of LSDS, like wide area distribution, heterogeneity, dynamism and incapable centralized control, the security of LSDS encounters severe challenges. Among them, the structural vulnerability that results from the complex dependencies between entities and the complex dependencies between layers is a key issue demanding more research emphasis during the LSDS design, implementation and deployment process[1, 2].

Because LSDS commonly use redundant backup mechanism to improve the survivability and the dependencies between nodes in LSDS are more complex, structural vulnerability analysis is much more complex than other complex systems, like electricity system, transport system and network system. As a result, existing structural vulnerability analysis approaches, like the approaches[5, 6] used in structural vulnerability analysis for the power grid, are not appropriate for analyzing the structural vulnerability of LSDS.

Aiming at these limitations, a new topology model is proposed, and failure tolerant mechanism is put forward. Based on the model, a structural vulnerability analysis algorithm is proposed, which can discover the structural vulnerabilities by computing weight of every entity and link, and validates them by graph pruning. After that, a new multi-topology model structural vulnerability analysis algorithm is proposed, which can detect the critical entities on different layers of LSDS according to the interdependencies between them. The evaluation of our implemented prototype system demonstrates the effectiveness of our approach.

The rest of the paper is organized as follows: Section 2 presents the related work briefly. Section 3 describes the problem model and defines the structural vulnerability in LSDS. Section 4 proposes an entity topology model structural vulnerability analysis algorithm and Section 5 proposes a multi-layer topology model structural vulnerability analysis algorithm. Evaluation results and analysis are provided in Section 6. Finally, our conclusions are given in Section 7.

2. Related work

The concept of structural vulnerability was firstly defined by Lu[3], Agarwal[4], et al. in the research of the reliability of architecture engineering as the inherent defects in inter-dependences between the components of a complex system. The use of these defects will lead to unpredictable results or risk. Subsequently, people conduct an in-depth research on the structural vulnerability for electricity, transport and network and other complex systems. Re 'ka Albert, Mart' Irosas-Casals, et al. studied the structural vulnerability of the power grid of North America and Europe respectively[5, 6]. Jenelius E, et al. investigated deeply the structural vulnerability of Swedish road network[7]. Luca Dall'Asta analyzed the structural vulnerability of the global aviation network, and proposed a vulnerability analysis approach based on the weighted network topology[8]. In the area of network and information systems, a number of of work researched the cross-layer survivability of the single fiber failure[9-14]. Modiano et al. [9] firstly addressed the survivable lightpath routing problem with the focus on resolving the structural vulnerability of IP/WDM, and then they proposed to improve route survivability in topology by linear programing. The method proposed in [9] is further optimized in [10] and [14]. Based on the two-tier topology structure, M. Kurant constructed a heuristic algorithm, named with SMART-H, for analyzing and correcting the structural vulnerability of IP/WDM[15]. András Faragó proposed a minimum cut algorithm, called F-Cuts, to analyze the structural vulnerability of network topology [16].

3. Problem Formulation

3.1 LSDS Multi-layer Topology Model

In LSDS, the computers interconnected through network are called entities or nodes. The applications in entities are called components. In order to analyze the structural vulnerability of LSDS, we introduce the entity topology model and multi-topology model, defined as follows:

Definition 1. The complex dependencies between entities in function-layer of LSDS is represented by a simple loop-free directed graph $G^{\mu} = \langle V^{\mu}, E^{\mu} \rangle$, called LSDS entity topology model. The node set V^{μ} is the entities in LSDS and the edge set E^{μ} is the dependencies between entities.

If the functions of entity v_i rely on that of entity v_j, there exists an edge $\langle v_i, v_j \rangle$. The start point of a business process is denoted with node s.

Definition 2. In graph $G^{\mu} = \langle V^{\mu}, E^{\mu} \rangle$, if every path that starts from point u and end to any node which outdegree is zero passes through node v, defines that node u is fully dependent on the node v, denoted with $\xrightarrow{c}$; if all paths from s to v pass through node u, and there exists a path from s to a node which outdegree is zero, and this path contains node u but without node v, defines that node u is not fully dependent on the node v, denotes with $\xrightarrow{nc}$; if every path that starts from point s and includes node u while ending to any node which outdegree is zero passes through node v, and all the paths from s to v contain u, defines that the node u is equivalently dependent on the node v, represented with $\leftrightarrow$.

Definition 3. The basic environment of LSDS, such as the infrastructure-layer or networking-layer, can be represented by a simple undirected graph $G^{\lambda} = (V^{\lambda}, E^{\lambda})$, called the low-layer topology structure which is infrastructure such as DNS, network and transport.

Definition 4. The dependencies between entity topology and low-layer topology and the dependencies between low-layer topologies can be represented by the mapping relation M, denoted with $M(G^{\mu}) = G_R{}^{\lambda}$, which can be categorized into node-mapping and edge-mapping as follows:

a) node-mapping : $\forall v \in V^{\mu}, \exists v' \in V^{\lambda}$, satisfies $v' = m(v)$.

b) edge-mapping : $\forall e = \langle v_i, v_j \rangle \in E^{\mu}, \exists p(v_i', v_j') \in P^{\lambda}$, satisfies $p(v_i', v_j') = m(e)$, $v_i' = m(v_i)$ and $v_j' = m(v_j)$. Among them, the set P^{λ} is all the paths between nodes in the networking-layer, and $p(v_i', v_j')$ is the path between v_i' and v_j'.

The number of the topology model layer is determined by the specific operating environment of LSDS. If LSDS depends on the infrastructure such as DNS or network, the model can be divided into three layers. If IP/WDM is involved in the infrastructure, the number of layer is four. If only considering the routing and switching mechanism, the model will be divided into two.

In the mapping process, if several entities are mapped onto the same host, the rings, parallel edges and loops may appear in the LSDS entity topology. That the dependency between LSDS entities is represented by the simple loop-free directed graph is essentially the simplification for the complex relationship between entities. In the entity topology of LSDS, the ring considered as the internal dependency is of no concern. And the parallel edge considered as several dependencies between two entities can be represented by one edge. Therefore, the dependency relationship between LSDS entities can be considered as the simple directed graph. At the same time, the loop describes the fact that the implementation of an entity's function is eventually dependent on the availability of that entity's own function. Thus in the deployment of LSDS it should be to avoid loops in a single business process.

If one entity is mapped onto several hosts, increases fault tolerance models and accordingly adjust the entity topology of LSDS. Consequently, cases of the one-to-many and many-to-one in the node-mapping process only influence the LSDS entity topology. To simply the model description, only the case of one-to-one is considered when the node-mapping of multi-topology. And the case of many-to-one or one-to-many can be resolved by modifying the entity topology model.

3.2 Fault Tolerant Mechanism

In the entity topology model, various faults may be categorized into the following types.

- Node Fault and Edge Fault

In entity topology, all the nodes or edges may fall into fault. Edge fault will lead to the disruption of the dependencies between nodes, and may cause the starting node of this edge to fail. The node fault will incur all the edges connected to this node getting failed.

- Single point fault and combinational fault

Single point fault refers to that a node or edge falls into fault in an entity topology model. The combinational fault indicates that several single point faults occur concurrently or sequentially.

The entity topology model of LSDS commonly adopts various fault tolerance mechanisms to improve the robustness of the system. Generally, the fault tolerance mechanism of LSDS is achieved through the redundant backup technology. In order to characterize the fault tolerance mechanism of LSDS, we introduce a new fault tolerance model to the entity topology structure, defined as follows:

Definition 5. Fault tolerance model $F(n,k)$ means that if the number of entity without fault in n entities is not less than k, this entity set n could provide functionality.

Figure 1 shows an example to illustrate the fault tolerance model.

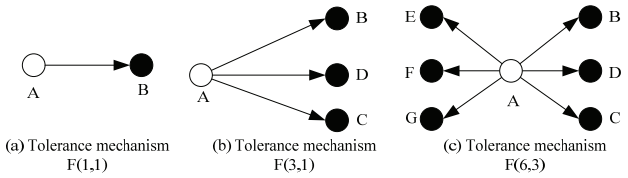

Fig.1. Fault tolerance mechanism

To improve the reusability of resources, the lower layer of LSDS multi-topology model generally adopts the redundancy and dynamic recovery technology. Taking into account the complexity of the fault tolerance model in lower layer topology, it isn't introduced in our multi-topology model.

3.3 Structural vulnerability

In the entity topology model of LSDS, if there is node or edge whose fault causes the starting node fail, we define that there exists some structural vulnerability in LSDS. If the structural vulnerability is caused by a single node or edge, we define that there exists a single point fault structural vulnerability in LSDS, and the node or edge is called as the structural vulnerability point; if structural vulnerability is incurred by multiple nodes or edges, we define that there exists combined fault structural vulnerability in LSDS, the collection of these nodes or edges is named with the set of structural vulnerability.

In the multi-topology model of LSDS, the core of structural vulnerability analysis aim to find the set of key nodes and edges in each layer of LSDS. If functions of elements in that set fall to fault, the starting node s in the entity topology will fail. That will cause that the business function become invalid.

4. Structural Vulnerability Analysis Algorithm for Entity Topology Layer

4.1 Problem Analysis

To analyze the structural vulnerability, we introduce weight to $G^{\mu} = \langle V^{\mu}, E^{\mu} \rangle$. The weight denoted as w is used

to characterize the importance degree of the nodes or edges for the business process, i.e., the influence degree of the node or edge fault on the root node's implementing the business processes. The weight value of the starting is 1.

Clearly, the weight will be spread with the dependent relations. If there does not exist fault tolerance models in the topology graph, all the influence degree of every node or edge on the root node should be 1. In other words, if there is any dependent node or edge getting failed, s should be failed. Because of the introduction of the fault tolerant model, the weight is divided. In order to characterize the fault tolerance model's impact on the weight, we firstly propose the following definition:

Definition 6. For the node v which makes use of the fault tolerant model F (n, k) to improve its robustness, let $\eta^{\mu}(v)$ be the weight partition coefficient of node v, which defines the fault tolerance model's impact on the weight. It's calculated with the following formula:

$$\eta^{\mu}(v)=\frac{1}{n-k+1} \tag{1}$$

To figure out the weights of nodes and edges in the entity topology, we introduce the following definitions:

Definition 7. In graph $G^{\mu}=\langle V^{\mu},E^{\mu}\rangle$, if the outdegree of node v^{μ} is 0, and there is a path from the node v to v^{μ}, represented with $p(v,v^{\mu})$, we define that node v is ending-dependent on node v^{μ}, denoted with $\mapsto$.

Definition 8. For the node v^{μ} in $G^{\mu}=\langle V^{\mu},E^{\mu}\rangle$, the node contained in every path, which starts from the point s and ends at point v^{μ}, is called as the associated node of the node v^{μ}.

With definition 8, it's apparent that s is the associated node of all nodes and the associated node of a node may be more than one.

Definition 9. In all the associated nodes of the node v^{μ}, the node nearest to v^{μ} is defined as the optimal associated node, represented with $A_{v^{\mu}}$.

Definition 10. If the outdegree of $A_{v^{\mu}}$ is equal to the indegree degree of v^{μ}, and in the sub-graph $G^{\mu}{}_{v^{\mu}}$ composed of all the path from $A_{v^{\mu}}$ to v^{μ}, in addition to the starting node and ending node, there is not any edge connected with the nodes outside this sub-graph, $A_{v^{\mu}}$ is denoted as the equivalent associated node of node v^{μ}, denoted with $Q_{v^{\mu}}$.

Deduced from definition, it's obvious $Q_{v^{\mu}}\leftrightarrow v^{\mu}$.

The weight of edge $e_{i,j}=\langle v_i,v_j\rangle$ is product of the weight of the edge's starting node and the weight partition coefficient:

$$w^{\mu}(e_{i,j})=\eta^{\mu}(v_i)*w(v_i) \tag{2}$$

According to the concept of weight, the following conclusion can be deduced obviously:

- If $u\xrightarrow{c}v$, $w(u)\le w(v)$
- If $u\xrightarrow{nc}v$, $w(u)>w(v)$
- If $u\leftrightarrow v$, $w(u)=w(v)$

Theorem 1. In $G^{\mu}=\langle V^{\mu},E^{\mu}\rangle$, each node has one and only one optimal associated node, and the weight of this node is equal or less than the weight of its optimal associated node.

Based on the above theorem, the weight $w(v_i)$ of node v_i can calculate as follows:

If the indegree of v_i is 0, v^{μ} is starting node, $w(v_i)$ = 1;

If the indegree of v_i is 1, the weight v_i is the weight of the edge ending at v_i;

If the indegree is greater than 1 and v_i has equivalent associated nodes, $w(v_i)=w(Q_{v_i})$. Otherwise, $w(v_i)$ is the minimum of the following three values: $w(A_{v_i})$; $\sum_{j\neq i}w(e_{j,i})$; In all the paths from A_{v_i} to v_i, the sum of the weights of all the edges starting from A_{v_i}.

According to the description of our model, the entity topology weighting algorithm is proposed as follows:

1) Using the path searching algorithm to work out the optimal associated nodes and equivalent associated nodes of all the nodes which indegree is greater than 1;

2) Starting with node s, calculate the completion time of each node $f[v]$ with the topological sorting algorithm;

3) After figuring out the completion time of a node, insert it into the list head;

4) In accordance with the order of the list, calculate the weights of all the nodes following the node weight calculating method.

Theorem 2. In the weighted entity topology graph $G^{\mu}=\langle V^{\mu},E^{\mu}\rangle$, the weights of all nodes are not larger than 1.

Theorem 3. In the weighted entity topology $G^{\mu}=\langle V^{\mu},E^{\mu}\rangle$, if the node is a single point fault structural vulnerability node, its weight is equal to 1.

Following the edge weight calculating method, the edge weight must be not greater than 1. If the edge weight is equal to 1, there must be a node whose weight is 1, and the node's fault tolerance model is F(1,1).

4.2 Algorithm Description

According to Theorem 3, the algorithm analyzing the single point fault structural vulnerability is represented as follows:

1) Calculate the weights of the nodes and edges in $G^{\mu}=\langle V^{\mu},E^{\mu}\rangle$ with the entity topology weighting algorithm;

2) Traverse the entity topology $G^{\mu}=\langle V^{\mu},E^{\mu}\rangle$, pick out all the nodes and edges whose weight is 1 except the starting node s, term them as $l_{\max}$ and append them to the set S;

3) Traverse each element $l_{\max}$ in the set S, utilize the pruning algorithm to remove $l_{\max}$ from $G^{\mu}=\langle V^{\mu},E^{\mu}\rangle$. The pruning algorithm described as follows:

a) If $l_{\max}\in E^{\mu}$, denotes as $\langle v_i,v_j\rangle$:

According to the distribution of the outdegree and indegree of v_i and v_j, cope with the graph in the following cases:

i). If $d_{out}(v_i)=1$, remove the edge $\langle v_i,v_j\rangle$ from G^{μ}, and delete the node v_i with the pruning algorithm;

ii). If $d_{out}(v_i)>1$, and $\langle v_i,v_j\rangle$ is an edge of the fault tolerance model F(n, k)

If $n==k$, remove the edge $\langle v_i,v_j\rangle$ from G^{μ} and delete the node v_i with the pruning algorithm;

If $n>k$, remove the edge $\langle v_i,v_j\rangle$ from G^{μ} and modify the

fault tolerance model F(n,k) to F(n-1,k);

iii). If $\mathrm{d}_{out}(v_i)>1$, and $\langle v_i, v_j\rangle$ is not one edge in the fault tolerance model F(n,k), remove the edge $\langle v_i, v_j\rangle$ from G^{μ} and delete the node v_i with the pruning algorithm;

b) If $l_{\max}\in V^{\mu}$, denotes as v

If the node v satisfies $\mathrm{d}_{out}(v)==\mathrm{d}_{in}(v)==0$, remove the node v from G^{μ}. Otherwise, remove the edges connected with v by way of the pruning algorithm.

Let $G'^{\mu}=\langle V'^{\mu}, E'^{\mu}\rangle$ be the sub-graph of G^{μ} worked out by the pruning algorithm. If the starting node s does not belong to V'^{μ}, the element $l_{\max}$ is the single point fault structural vulnerability of the entity topology $G^{\mu}=\langle V^{\mu}, E^{\mu}\rangle$.

When there is no node or edge whose weight is 1 in the weighted entity topology graph, we should analyze the combined fault structural vulnerability. Such structural vulnerability is closely related to the number of elements in the structural vulnerability set. To characterize the combined fault tolerance ability of LSDS, we introduce R to denote the combined fault tolerance strength of LSDS. It refers to that only if there are at least (R+1) elements, except the starting node s, are failed, s will fall into fault.

The algorithm analyzing the combined fault structural vulnerability is represented as follows:

$Iterator=1$ denotes the iteration count.

(1) Assign the weights of the nodes and edges in $G^{\mu}=\langle V^{\mu}, E^{\mu}\rangle$ with the entity topology weighting algorithm;

(2) Traverse the entity topology $G^{\mu}=\langle V^{\mu}, E^{\mu}\rangle$, pick out the node or edge whose weight is maximal except the starting node s, term it as $l_{\max}$ and append $l_{\max}$ to the set S. If there are more than one node or edge with the same maximal weight, select the node or edge nearest to the starting node s;

(3) Remove the element $l_{\max}$ from $G^{\mu}=\langle V^{\mu}, E^{\mu}\rangle$ with the pruning algorithm which is same as the one adopted to analyze the single point fault structural vulnerability. The worked out sub-graph of G^{μ} is expressed with $G'^{\mu}=\langle V'^{\mu}, E'^{\mu}\rangle$;

4) If the starting node s is not in V'^{μ}, the set S is the structural vulnerability node of the entity topology $G^{\mu}=\langle V^{\mu}, E^{\mu}\rangle$, and the combined fault tolerance strength of this entity topology is $Iterator-1$. Otherwise, replace G^{μ} with G'^{μ}, $Iterator++$.

If $Iterator<R$, repeat the step 1; Otherwise, complete the algorithm, and the combined fault tolerance strength of this entity topology is greater than R.

5. Structural Vulnerability Analysis Algorithm for Multi-Layer Topology

Due to differences in the environment, deployment and layer-to-layer dependency, it is very difficult to analyze the structural vulnerability of the low-layer topology in LSDS. On the basis of the entity topology model structural vulnerability analysis algorithm, we propose the multi-topology model structural vulnerability analysis algorithm by the mapping-backup thought to solve that problem. The core idea of this algorithm is to calculate weights of nodes and edges in the lower layer topology according to the layer-to-layer mapping relation and weights of nodes and edges in the entity topology; then based on the mapping relation, analyze the influence on the upper layer topology when the maximal weight node selected in the lower layer topology fails. After that, we can figure out structural vulnerabilities of the lower layer topology.

Because of the complex dependency between layers, structural vulnerability analysis algorithms vary with the dependency relation. Considering that most LSDSes are dependent on the networking-layer infrastructure, we take the case of the double-layer topology model comprised of the entity layer and networking layer to describe the LSDS multi-topology model structural vulnerability analysis algorithm.

In order to calculate weights of nodes and edges in the networking-layer topology, we propose the following definitions on the basis of the multi-topology model:

Definition 11. In entity topology G^{μ}, if exits a path containing the edge e_1 and e_2, and this path starts from the point s, defines that e_1 and e_2 are the associated edges each other.

Definition 12. In entity topology G^{μ}, if each element of the edge set $\{e_1, e_2\cdots, e_n\}$ is one of the n edges in the fault tolerance model F(n, k), defines that the edge in the set $\{e_1, e_2\cdots, e_n\}$ is the k-order equivalent edge.

If the networking-layer node is directly mapped from the entity topology, its weight is equal to the weight of the corresponding node in the entity topology. Otherwise, figuring out the entity-topology edge set E mapping onto that networking-layer node, and removing the associated edges and equivalent edges from E, the weight of that networking-layer node is the sum of the weights of all the edges in E.

To calculate the weight of the networking-layer edge, similarly as above, we firstly figure out the entity-topology edge set E' mapping onto that networking-layer edge and then remove the associated edges and equivalent edges from E'. After that, the weight of that networking-layer edge is the sum of the weights of all the edges in E'.

The LSDS multi-topology model structural vulnerability analysis algorithm is described as follows:

S represents the edge set of G^{λ}.

T represents the node set of G^{λ}.

$p_i(v'_i, v'_j)$ represents the networking-layer path mapped from the entity topology edge $e_i=\langle v^{\mu}{}_i, v^{\mu}{}_j\rangle$.

$M(e^{\lambda}{}_{i,j})$ represents the entity-topology edge set, which satisfies that the networking-layer path mapped from each element of this set contains the networking-layer edge $e^{\lambda}{}_{i,j}=(v^{\lambda}{}_i, v^{\lambda}{}_j)$.

1) Firstly, remove the redundant associated edges from $M(e^{\lambda}{}_{i,j})$.

for each set $M(e^{\lambda}{}_{i,j})$

for k = 1 to $|M(e^{\lambda}{}_{i,j})|$

for n=k+1 to $|M(e^{\lambda}{}_{i,j})|$

if (both m_k and m_n are associated edges) && ($w^{\mu}(m_k)\geq w^{\mu}(m_n)$)

remove m_n from $M(e^{\lambda}{}_{i,j})$

else

remove m_k from $M(e^{\lambda}{}_{i,j})$

After that, the set $\overline{M}\left(e^{\lambda}_{i,j}\right)$ without redundant associated edges is worked out.

2) Secondly, remove the redundant equivalent edges from $\overline{M}\left(e^{\lambda}_{i,j}\right)$. For each set $\overline{M}\left(e^{\lambda}_{i,j}\right)$, traverse all the fault tolerance models of the entity topology;

3) If $\overline{M}\left(e^{\lambda}_{i,j}\right)$ has the subset, i.e., $\overline{N}\left(e^{\lambda}_{i,j}\right)\subseteq\overline{M}\left(e^{\lambda}_{i,j}\right)$. The edges of $\overline{N}\left(e^{\lambda}_{i,j}\right)$ are k-order equivalent edges of the fault tolerance model F(n,k). And satisfy $\left|\overline{N}\left(e^{\lambda}_{i,j}\right)\right|>n-k+1$. Then randomly select $\left|\overline{N}\left(e^{\lambda}_{i,j}\right)\right|-(n-k+1)$ edges and remove them from $\overline{M}\left(e^{\lambda}_{i,j}\right)$.

The worked out set without redundant equivalent edges is represented by $\overline{\overline{M}}\left(e^{\lambda}_{i,j}\right)$.

4) Calculate weights of nodes and edges in the networking-layer topology. The weight of $e^{\lambda}_{i,j}$ is the sum of the weights of all the edges in the set $\overline{\overline{M}}\left(e^{\lambda}_{i,j}\right)$, i.e., $w^{\lambda}\left(e_{i,j}\right)=\sum w^{\mu}(e)$, $e\in\overline{\overline{M}}\left(e^{\lambda}_{i,j}\right)$. The weight of v^{λ}_{i} is one-half of the sum of the weights of the edges connected with v^{λ}_{i}.

5) Traverse the networking-layer topology $G^{\lambda}=\left(V^{\lambda},E^{\lambda}\right)$, pick out the node whose weight is maximal, term it as $v^{\lambda}_{\max}$ and append it to the set T. Figure out the $\overline{\overline{M}}\left(e^{\lambda}_{i,j}\right)$ elements which map onto the networking-layer edges connected with $v^{\lambda}_{\max}$ and append them to the set S. Remove $v^{\lambda}_{\max}$ and the edges connected with $v^{\lambda}_{\max}$ from $G^{\lambda}=\left(V^{\lambda},E^{\lambda}\right)$.

6) Remove each element of the set S from $G^{\mu}=\left\langle V^{\mu},E^{\mu}\right\rangle$ with the pruning algorithm and the worked out sub-gruph is expressed with $G'^{\mu}=\left\langle V'^{\mu},E'^{\mu}\right\rangle$. If the starting node s is not in V'^{μ}, the set $T\subseteq V^{\lambda}$ is the structural vulnerability point of the networking-layer topology; otherwise, repeat the step 4.

In the step 4 of the multi-topology model structural vulnerability analysis algorithm, change the maximal weight node with the maximal weight edge, we can also get the vulnerability analysis result. Now, $T\subseteq E^{\lambda}$.

The multi-topology model doesn't introduce the networking-layer fault tolerance mechanism. Consequently, the result of the networking-layer topology structural vulnerability analysis algorithm is true for the current routing mode, but these vulnerabilities may be fixed through the fault tolerance mechanism. Therefore, this result should be further analyzed and verified. The vulnerability analysis environment based on the fault-injection is effective to do this.

6. Simulation Analysis and Validation

We implement, analyze and validate the performance of our algorithms with a PC which has 2.33GHz CPU and 1.5GB RAM.

6.1 Analysis of the entity topology model structural vulnerability

In the experiment, the entity topology graphs are constructed by removing the loops from the randomly generated directed graphs. The fault tolerance models in these entity topology graphs are built by randomly selecting the dynamically generated nodes whose outdegree is larger than 1.

Figure 2 shows that the runtime of algorithm changes with the number of nodes nonlinearly. However, when the number of nodes is 2000, the algorithm only costs 500 seconds. From this figure 3, it can be seen that, in the randomly generated entity topologies, the number of single point fault increases rapidly with the number of nodes. This indicates that the more complex is the dependencies between LSDS entities, the higher is the occurrence probability of the structural vulnerability. It can be seen from figure 4 that the more are fault tolerance models, the more obvious is the decreasing trend of the vulnerability count.

The implementation and analysis of our algorithm demonstrate that the entity topology model structural vulnerability analysis algorithm costs less and the runtime of that algorithm is insensitive to the number of fault tolerance model. Meanwhile, in the randomly generated entity topologies, the number of single point faults increases significantly with the number of nodes and decreases rapidly with the number of fault tolerance models. Therefore, the more complex is the LSDS structure, the greater is the occurrence possibility of the structural vulnerability. And the redundant backup mechanism is an effective way to reduce the structural vulnerability.

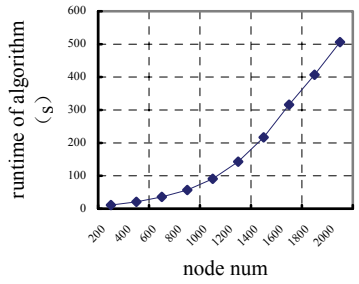

Fig.2. The runtime of algorithm for different node number

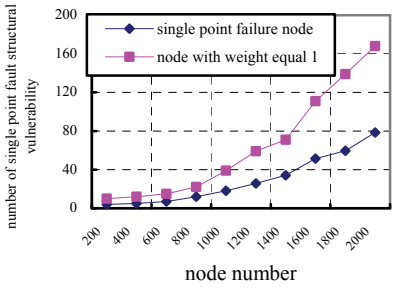

Fig.3. Structural vulnerability for different node number

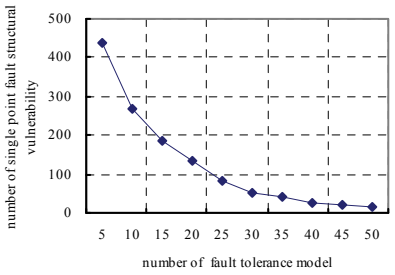

Fig.4. Structural vulnerability for different number of fault tolerance model

6.2 Analysis of the multi-topology model structural vulnerability

In the experiment, there are 500 nodes in the entity topology whose average shortest path length is 6. And this topology has 36 structural vulnerability nodes and 60 structural vulnerability edges. The networking-layer

topology generated through the large-scale topology generation framework [17] implemented by us is divided into AS-tier, router-tier and terminal-tier. And the ratio of terminals to routers is 4:1.

In Figure 5, it can be seen that the runtime of algorithm changes with the number of nodes linearly. According to the multi-topology algorithm, the runtime is related to the path length; and the number of router nodes changes with the topology scale slowly. Consequently, the time cost changes not significantly with the networking-layer topology scale. Knowing weights of the nodes and edges in the entity topology, multi-topology algorithm costs less and that cost changes with the node number linearly.

In Figure 6, it can be seen that the number of single point fault structural vulnerabilities in the networking-layer topology is much greater than that in the corresponding entity topology. At the same time, in the networking-layer topology, because single point fault structural vulnerabilities are mainly caused by the mapping from the vulnerable edges in entity topology, the number of single point fault structural vulnerabilities increases with the node number slowly. Because the average path length increases with the node number, the number of vulnerable edges, mapped from vulnerable edges in the entity topology, will consequently increase. Thus the numbers of single point fault structural vulnerabilities in the networking-layer increases slowly.

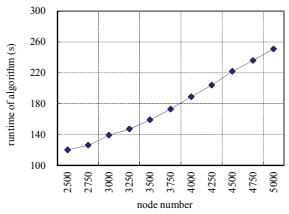

Fig.5. The runtime of algorithm for different node number

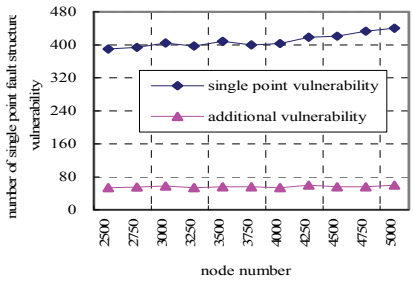

Fig.6. Structural vulnerability for different node number of networking-layer topology

7. Conclusion

Structural vulnerability is one of the typical vulnerability of complex systems. The methods to analyze it vary significantly in different complex systems. Focusing on the entity dependency and layer-to-layer dependency of LSDS, the entity topology model structural vulnerability analysis algorithm and the multi-topology model structural vulnerability analysis algorithm were proposed respectively. And the evaluation results demonstrate the effectiveness of the algorithms. This work contributes to solve the structural vulnerability problem of LSDS which has the property of redundant backup and layer-to-layer dependency.

The dynamic reconfiguration and redundant backup of LSDS low-layer topology make the fault tolerance model more complex. The algorithm proposed in this paper is concerned mainly with the fault tolerance mechanism of LSDS entity topology. The influence of the fault tolerance mechanism in the low-layer topology model on structural vulnerability is studied mainly through experiments. That has a lot of limitations. In the future, we will work chiefly on the fault tolerance mechanism of the low-layer topology and optimize the structural vulnerability analysis algorithm.

REFERENCES

[1] Helsinger, A.; Ferguson, W.; Lazarus, R . Exploring large-scale, distributed system behavior with a focus on information assurance[C]. In: DARPA Information Survivability Conference & Exposition II, 2001. DISCEX apos;01. Proceedings Volume 2, Issue , 2001 p273~286 vol.2

[2] Qiang Weizhong, Zou Deqing, and Jin Hai. Research on privacy preservation mechanism for credentials and policies in Grid computing environment [J]. Journal of Computer Research and Development, 44(1), p11~19，2007

[3] Lu Z, Yu Y, Woodman NJ & Blockley DI. A theory of structural vulnerability [J]. The Structural Engineer, 77(18):p17~24，1999

[4] Agarwal J, Blockley DI & Woodman NJ. Vulnerability of 3D trusses [J]. Structural Safety, 23(3):203-220，2001.

[5] R. Albert, I. Albert, G.L. Nakarado. Structural vulnerability of the North American power grid [J], Phys. Rev. E 69 (2004) 025103(R).

[6] Rosas-Casals, M., S. Valverde, and R. Solé. Topological vulnerability of the European power grid under errors and attacks[J]. International Journal of Bifurcations and Chaos, 2007.17(7).

[7] Jenelius, E., Petersen, T. and Mattson, L. Importance and exposure in road network vulnerability analysis [J]. Transportation Research vol.40, p537~560，2008

[8] L. DallAsta, A. Barrat, M. Barth´elemy, and A. Vespignani. Vulnerability of weighted networks [J], physics/0603163, 2006.

[9] Eytan Modiano and Aradhana Narula-Tam. Survivable lightpath routing: A new approach to the design of WDM-based networks. IEEE Journal on Selected Areas in Communications, vol. 20, no. 4, pp. 800-809, May 2002

[10] Q. Deng, G. Sasaki, and C.-F. Su. Survivable IP over WDM: a mathematical programming problem formulation. In Proc 40th Allerton Conference on Communication, Control and Computing, Monticello, IL, October 2002.

[11] Maciej Kurant and Patrick Thiran. Survivable Routing of Mesh Topologies in IP-over-WDM Networks by Recursive Graph Contraction. IEEE Journal on Selected Areas in Communications, 25(5):922 – 933, 2007.

[12] Arunabha Sen, Bin Hao, and Bao Hong Shen. Minimum cost ring survivability in WDM networks. In Workshop on High Performance Switching and Routing, 2003, pages 183– 188, 2003.

[13] Kayi Lee, Eytan Modiano. Cross-Layer Survivability in WDM-Based Networks.

[14] Ajay Todimala and Byrav Ramamurthy. Survivable virtual topology routing under Shared Risk Link Groups in WDM networks. In BROADNETS '04: Proceedings of the First International Conference on Broadband Networks, pages 130–139, Washington, DC, USA, 2004. IEEE Computer Society

[15] M. Kurant and P. Thiran. Survivable Mapping Algorithm by Ring Trimming (SMART) for large IP-over-WDM networks[C]. In: Proceedings of BroadNets 2004, p25~29, San Jose, California, USA.

[16] A.Farago. A graph theoretic model for complex network failure scenarios[C], In: INFORMS 2006

[17] Kuang Xiaohui, Huang Minhuan, Li Jin. Large-scale network topology generation for emulation environment[C]. Proceedings of the 1st International Conference on Networking and Distributed Computing, p392-396,Hangzou,China, 2010

***Authors**: associate professor, doctor, Kuang Xiaohui, Section 4, P. O. Box 9702-19, Beijing 100101, P.R. China, E-mail: xiaohui_kuang@163.com; professor, doctor, Zhao Gang, Section 4, P. O. Box 9702-19, Beijing 100101, P.R. China, E-mail: zhaog03@mails.tsinghua.edu.cn; associate professor, doctor, Yong Tang, Section 4, P. O. Box 9702-19, Beijing 100101, P.R. China, E-mail: xiaohui_kuang@163.com; lecturer, master, Li Jin, Section 4, P. O. Box 9702-19, Beijing 100101, P.R. China, E-mail: zyan1981@gmail.com.*