Conference Paper for MAFIOK 2016.
Óbuda University Alba Regia Technical Faculty, Hungary, Székesfehérvár

P-graph methodology - an efficient approach for optimizing computer systems

Nemes Teréz*, Süle Zoltán**, Dávid Ákos**

**Department of Methodology of Budapest Business School Budapest*
***Faculty of Information Technology of University of Pannonia, Veszprém*
Nemes.Terez@uni-bge.hu, davida@almos.uni-pannon.hu, sule@dcs.uni-pannon.hu

Keywords: P-graph, security level, categorization method

Abstract

In recent years IT experts have encountered a continuously rising number of security events in computer systems. Due to the evolution of various types of cyber-attacks, designing an adaptive defense becomes challenging.

The two fundamental elements of computer networks are robustness and the number of known vulnerabilities in the system. Measuring the level of security becomes even more demanding, but currently there is no objective method to do this automatically. The available measuring methods are usually subjective and are based on the surveying of system operators.

Our aim is to design an objective definition for security level that is based on the topology provided by the operators. We provide an algorithmic method to calculate this value. This gives the optimization of the current architecture complying financial constraints. This method is used for measuring the robustness and vulnerability of computer systems.

The mathematical background of our method is based on the P-graph methodology that has been used for modelling various engineering systems, as well as business processes and supply-chains [1].

In this paper, we examine the security of computer systems with P-graph by transforming the current architecture topology given by a typical diagram representation into a P-graph model.

[1] Botond Bertok, Karoly Kalauz, Zoltan Sule, Ferenc Friedler: Combinatorial Algorithm for Synthesizing Redundant Structures to Increase Reliability of Supply Chains: Application to Biodiesel Supply, *Industrial & Engineering Chemistry Research*, Vol. 52 Issue 1, pp. 181-186, 2013.

Conference Paper for MAFIOK 2016.
Óbuda University Alba Regia Technical Faculty, Hungary, Székesfehérvár

P-gráf módszertan hatékony alkalmazása informatikai rendszerek optimalizálására

Nemes Teréz, Süle Zoltán, Dávid Ákos

Kulcsszavak: P-gráf, biztonság szintjei, kategorizálási módszer

Kivonat

Az utóbbi években a számítógépes rendszerekben egyre növekvő számú biztonsági eseménnyel kell számolnia a szakértőknek. A fenyegetések különböző fajtáinak fejlődésével az adaptív védekezés kialakítása napjainkra egyre nagyobb kihívást jelentő feladattá vált.

A számítógépes hálózatok biztonságának két meghatározó eleme a robusztusság és az előforduló sérülékenységek száma. A kialakított rendszer biztonság szintjét fontos minél pontosabban felmérni, de erre csak szubjektív, kikérdezésen és a szakértők tudásbázisán alapuló mérési módszerek léteznek.

Munkánk célja egy olyan módszertan kidolgozása, amely egy az üzemeltetők vagy fejlesztők által megadott topológiából kiindulva megadja a rendszer biztonsági szintjét és megadott költségkorlát és a meglévő informatikai eszközök mellett optimalizálási lehetőségeket is felkínál. Ezzel egy objektív mérési metodikát alakítunk ki az informatikai rendszer robusztusságának és sérülékenységének mérésére.

Módszerünk működésének matematikai hátterét a P-gráf módszertan biztosítja, amely hatékony eszköznek bizonyult különböző műszaki rendszerek vizsgálatára, ugyanakkor számos gazdasági és logisztikai alkalmazása is ismert, így például a bonyolult üzleti folyamatok és ellátó rendszerek optimalizálásában is hasznos eszköztárat nyújt [1].

Jelen cikkben bemutatjuk a P-gráf módszertan adaptációját számítógépes rendszerek biztonságának vizsgálatához, amelynek egy fontos lépése az analógia formális definiálása a hálózati topológiát reprezentáló elemek és a P-gráf csomópontjai között.

[1] Botond Bertok, Karoly Kalauz, Zoltan Sule, Ferenc Friedler: Combinatorial Algorithm for Synthesizing Redundant Structures to Increase Reliability of Supply Chains: Application to Biodiesel Supply, *Industrial & Engineering Chemistry Research*, Vol. 52 Issue 1, pp. 181-186, 2013.

INTRODUCTION

In recent years, IT experts have encountered a continuously rising number of security events in computer systems. Due to the evolution of various types of cyber-attacks, designing an adaptive defense becomes challenging.

The two fundamental elements of computer networks are robustness and the number of known vulnerabilities in the system. Computer security and robustness cannot be analyzed without each other. While examining a reliable operation of either bare hardware or cloud-based systems, the robustness originates practically from structure, bounded by topological and physical dependencies.

Computer networks are supposed to meet predefined security levels. Both the definition and the measurement of such security levels are usually based on subjective methods. In the research carried out by Kresimir Solic and others, network admins, normal users and security experts were asked to evaluate their systems in terms of security [4]. This solution obviously depends on the competence of the participants of the survey, and is very subjective. In other cases, I found similar solutions. For example in Hungary, according to the „Information Security Law" on the electronic security of state, government and self-government institutions, which defines five available security levels, and the categorization can be executed using a questionnaire. In this case, a miscategorization can lead to severe legal and financial implications. So measuring the level of security becomes even more demanding, but currently there is no objective method to do this automatically.

MY TASKS

Our aim was to design an objective definition for security level that is based on the topology provided by the operators. We provide an algorithmic method to calculate this value. This gives the optimization of the current architecture complying financial constraints. This method is used for measuring the robustness and vulnerability of computer systems.

The mathematical background of our method is based on the P-graph methodology, which has been used for modeling various engineering systems, as well as business processes and supply-chains.

Administrators and architects can measure or improve the security level of their computer networks based on objective parameters using our method. This method does not only maximize reliability based on the current architecture, but also satisfies financial constraints.

This method is based on the methodology of P-graphs. The topology given by the operator or architect can be modeled by P-graphs, and by performing optimization on this P-graph, robustness can be measured. The result will be reflected in the original topology, therefore we can make recommendations on the enhancement of the network's robustness

P-GRAPH METHODOLOGY

Professor Fan and Friedler developed a model to examine special optimization problems in 1970s [2]. In this model, there are different raw materials and operating units. Operating units can create other materials from raw materials or other intermediate materials. The goal is to produce specific materials (product materials) by connecting the operating units appropriately. The mathematical representation is a directed bipartite graph. The first vertices set represents the materials, the other represents the operating units. There is an arc from a material to an operating unit, if the material is processed by the operating unit. There is an arc from an operating unit to a material, if the operating unit produces that material.

Using this model it is possible to exploit combinatorial properties of solution structures, thus optimizing large combinatorial problems [1].

This model is applied in the following areas amongst many:

- Multi-processor systems
- Security systems
- Financial processes
- Logistic systems
- Optimization of infrastructure services
- Supplier scheduling
- Pipe systems
- Transport optimization
- Financial transactions

MATHEMATICAL DEFINITION OF P-GRAPH:

Let M be a set of objects is finite and not empty

$O \subseteq \wp(M) \times \wp(M)$, where $\wp(M)$ denotes powerset of M

(M, O) pair called P-graph

Where

The set of vertices: $M \cup O$

Set of edges: $A = A1 \cup A2$,

where
$A1 = \{(x, Y) : Y = (\alpha, \beta) \in O, x \in \alpha\}$
$A2 = \{(Y , z) : Y = (\alpha, \beta) \in O, z \in \beta\}$.
If $(\alpha, \beta) \in O$, then α is an input, β is an output material set [3]

P-GRAPH REPRESENTATION

Three cornerstones of the P-graph framework

- Structural representation: P-graph
- Axioms of the combinatorically feasible process networks
- Algorithms for
    - Generating the rigorous superstructure
    - Generating the combinatorically feasible process structures
    - Generating optimal or n-best networks

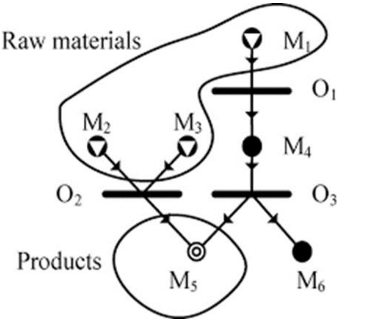

Figure 1. an example of P-graph

The optimal solution structure generated by process-network synthesis must have several basic features that are taken for granted as axioms, and the introduction of which improves the efficiency of the combinatorial search during the process.

Axioms of a feasible solution:
Every feasible solution should have the following properties (S1) - (S5).
These properties are used in the P-graph algorithms.

(S1) Each final product is represented in the graph.

(S2) An M-type vertex does not have an input if and only if it represents a raw-material.

(S3) Each O-type vertex representing an operating unit is defined in the network synthesis problem.

(S4) There must be at least one route from each O-type (operating unit) vertex represented in the structure leading to an M-type vertex representing a product.

(S5) If an M-type vertex belongs to the graph, then there must be at least one route leading to an O-type vertex or a route from an O-type vertex to the given M-type vertex

There are three known algorithms to calculate the optimal solutions: MSG algorithm, SSG algorithm, ABB algorithm.

The MSG algorithm generates the maximal structure in polynomial time. The maximal structure contains all combinatorically possible solutions, thus the optimal one too.

The SSG algorithm enumerates all possible solutions exactly once, including the optimal one.

The ABB algorithm generates the optimal and near-optimal solutions. It is based on the SSG algorithm, and tries to find the optimal solution based on a cost function, using branch and bound method.

This model is used for the problem mentioned in the introduction: developing an objective method for defining and calculating robustness of computer networks.

First, we need a formal definition of computer networks.

The definition of network topology is a directed graph D=(V,A)

A represents the nodes:

$V = V\_nh \cup V\_sp \cup V\_dp \cup V\_c$, where

V_nh denotes the network hardware: routers, switches, multilayer switches

V_sp denotes the service provides: webservers, mail servers …

V_dp denotes the data providers: databases

V_c denotes the clients

and A denotes the arcs of the graph: network connections.

There are two nodemaps for robustness and vulnerability:

$$f_r{:}V \rightarrow \mathbb{R}^+$$

$$f_v{:}V \rightarrow \mathbb{R}^+$$

Our method utilizes an objective measure of vulnerability of the nodes during calculating robustness: the vulnerability score. We define the vulnerability score using the official CVSS scores, which is publicly available in the NVD database. We consider the set of security issues happened in the past, and thus we estimate the probability of issues in the future. This means this method has a statistical approach, which is based on the already known vulnerabilities. I developed a script tool that is able to parse and merge many NVD database snapshots, thus collect the vulnerability data. I developed a graphical frontend as well to select the software used and to calculate the accumulated CVSS score, from which we calculate the vulnerability score.

Our goal is to develop an automatic system that optimizes this reliability level by changing topology, network elements or software.

We have specified an XML schema, which can represent the given topologies for further processing. The XML schema will contain nodes, which can be either hosts, servers or network devices, and the connections between them. The XML file will contain the reliability metrics of the nodes. We have developed a method that generates a valid XML file from a given network topology. We have elaborated a system between the network topology and the P-graph methodology.

We consider the following thesis: each entity in the original topology can be represented as a service.

A client is modelled as an operating unit that generates a request from user event. Thus in the P-graph we replace all clients with an operating unit, and add a new node for the user event as a raw material, and a new arc is added from the event to the operating unit.

The same thesis is applied to data providers. Data providers (e.g a databases) are considered as an operating unit that constantly provides input for the other network elements. Thus we replace the data provider with an operating unit and attach a new raw material to it as an input, similarly to the clients.

For all other equipment: we consider they provide some service (e.g login user) that is triggered from some request and using some data provider information. Thus these nodes are replaced by an operating unit.

We add a new node as intermediate product to every arcs of the original graph. The nodes on an outgoing arc will be the provided service that the operating unit creates. The nodes on the incoming arcs will be the trigger events or data that is provided by some other operating unit.

Those servers that do not have outgoing arcs are considered to produce product-material.

We have been developed a software solution. There is a graphic interface for operators to provide their network topology. The topology is stored in XML format. The represented graph is transformed into a P-graph, and saved into an other XML file, that is compatible with P-graph studio. P-graph studio is executed. P-graph studio runs the SSG algorithm, and provides a maximal solution. The optimal solution is determined, and the corresponding robustness score is calculated using a python script. The solution is transformed back to the original network topology.

Our method to calculate robustness score:

Every node has a robustness score.

We can calculate the robustness score of the current node using the formula:

The output robustness of „Logical or” connection:

$$p_1+p_2-p_{1*}p_2$$

The output robustness of „Logical and” connection:

$$p_1*p_2$$

We have chosen a network topology of an enterprise accounting company to compare the robustness values using two different methods. The first method was a subjective method based on interrogating the administrator and two employers about the network. We used the technique described in the paper of Kresimir Solic and others [4]. After querying, the robustness value was quite similar using our method.

Conclusion

We gave an objective and measurable definition for security and reliability of computer networks, both bare metal and cloud-based systems. We have also been developing a complex tool which calculates the actual robustness, security and reliability level of computer networks from a given topology and software set. A network administrator should be able to have a visual overview of different setups and would be able to see how the reliability level changes. This way, one can optimize their systems, and improve the reliability level, which can result in a cost efficient system setup as well. The process is fully automatized. The necessary mathematical models are developed, the transformation from network topology to P-graph is formalized. We compared our method on a previous method, and the result for robustness was similar. The application P-graphs in optimization of computer networks was very efficient.

References

[1] Botond Bertok, Karoly Kalauz, Zoltan Sule, Ferenc Friedler: Combinatorial Algorithm for Synthesizing Redundant Structures to Increase Reliability of Supply Chains: Application to Biodiesel Supply, Industrial & Engineering Chemistry Research, Vol. 52 Issue 1, pp. 181-186, 2013.

[2] Friedler, F., K. Tarjan, Y. W. Huang, and L. T. Fan, Combinatorial Algorithms for Process Synthesis, Computers Chem. Engng, 16, S313-320, 1992

[3] Ivo Raedts, Marija Petkovic1, Yaroslav S. Usenko, Jan Martijn van der Werf, Jan FrisoGroote, Lou Somers:

Transformation of BPMN models for Behaviour Analysis, Verification and Validation of Enterprise Information Systems, pp. 126-137, Funchal, Madeira, Portugal 2007.

[4] Kresimir Solic, Hrvoje Ocevcic, Marin Golub: The information systems' security level assessment model based on an ontology and evidential reasoning approach, Computers & Security, Vol. 55, pp. 100-112, 2015.