Review of the thesis
"Network-wide Security Analysis"
submitted by
Hidda Marakkala Gayan Ruchika de Silva

The thesis focuses on the problems of building an effective model for dynamic networks and developing an analysis method to evaluate their reachability and security properties. In the first part of the thesis, the author discusses three network modeling methods two of which are new. These methods are then employed in the second part of the work aiming to develop methods for analysis of reachability, device configurations, routing, filtering, and service quality. The research is based on using modern mathematical tools, in particular those provided by graph theory, category theory, and relational algebra.

With 163 pages of text divided into nine chapters, the thesis is quite comprehensive. The first chapter, Introduction, gives all information an introduction should provide including an overview, motivation, significance of the research, and the methodology used.

The second chapter gives a state-of-the-art description containing sufficiently detailed information about the related research work performed by other authors.

The next, core, chapter describes three methods of network modeling. The first of them builds on using the so-called abstract network graph and provides a convenient tool for an easy computing of available network paths. The second method uses the model given by a special graph, the so-called filtering network graph. This method is completely new and may advantageously be used to find hidden paths. The filtering network graph is obtained from the abstract one by a procedure whose basic step is a transformation of the network topology. This transformation may be given by a collection of graph transformations and, therefore, graph transformations are studied in a separate subchapter. Here, nontrivial results and considerations of graph theory and category theory are used to get an algorithm for transforming the abstract network graph into the filtering one. The last of the three models dealt with, the forwarding device network model, is obtained as a combination of the abstract network graph and the routing information base. This model, too, is new and provides a useful tool for network analysis as shown further in the thesis. At the end of the chapter, the advantages and disadvantages of individual network models are discussed.

In the fourth chapter, Reachability Analysis, formal methods are used for checking the reachability of dynamic networks. The main purpose of the methods is to evaluate the network reachability in an event of a link or device failure. For this reachability analysis, the abstract network model discussed in the previous section is used. As the main result of the section, a formal model, called Modified Topology Table, briefly MTT, is developed, which covers all available paths and network states so that it may be used for predicting the reachability of any given network state. An overview is also given of the state loss graph patterns, i.e., the patterns of abstract graphs of the networks that are not producing the entire state space when constructing the MTT. Such graphs are to be avoided when using MTT. This problem may be solved by employing the filtering network graph to obtain the situation with MTT containing the full state space and all paths.

The fifth chapter deals with a method of modeling and analyzing device configurations and packet transformations based on constraints. Constraint data are expressed in terms of constrained relations and, therefore, relational algebra is used as a tool for constructing the constraint data model. Since any device configuration can be transformed into a logical formula, each input-to-output conversion may be expressed as a logical operation with a set of predicates. It follows that a logical programming language like Prolog may be used to implement the modeling and analysis.

The next chapter is devoted to an analysis of routing based on device configuration. The model with forwarding devices and routing information bases discussed in chapter 3 is used as the network model for the analysis. The method developed gives a network depiction convenient for predicting the communication paths in the network thus properly configuring routing and avoiding problems such as floods.

In the seventh chapter, a complete analysis method is developed for evaluating the network security policy versus the configured firewall filtering rules. The method is an important contribution to the firewall policy analysis.

Analysis with simulation tools is dealt with in the eight chapter. The chapter focuses on semi-automatic analysis of networks configured with a dynamic routing protocol. The approach uses a combination of formal methods and simulation. The simulation model developed is based on using the abstract network model from chapter 3 and the reachability analysis described in chapter 4. An experiment is presented to demonstrate the correctness of the analysis method.

In the last chapter, conclusions and suggestions for the future work are presented.

The problems solved in the thesis represent highly topical issues in the field of network-wide security analysis of computer networks and, therefore, the subject of the thesis fully corresponds to the the Computer Science and Engineering Ph.D. study programme. The results attained are non-trivial, numerous and give new, useful tools for modeling dynamic networks and analyzing their reachability and security properties. In particular, the method developed based on using filtering network graphs provides a new, efficient tool for network modeling with significant applications, e.g., in constraint-based analysis. The method obtained as a combination of the abstract network model and the routing information base is an important and useful contribution to the field, too. It may be used, as shown in the thesis, for creating new, efficient methods of routing analysis. Also the other results presented in the thesis and mentioned above concerning the analysis of reachability, device configurations, routing, filtering and service quality are interesting and valuable because they provide useful and efficient analysis methods. The high quality of the results is also corroborated by the fact that most of them are contained in the papers written by the author himself or in collaboration with some other authors and published in renowned scientific journals or proceedings of specialized international conferences.

The thesis is presented in a clear, decent form and the entire work is well structured. The English used is very good, there are only a few misprints and negligible language errors in the text. But I found several inaccuracies and confusions in the description of the mathematical background used. The list of them is given below.

1) The Definition 3.1 should be reformulated into the following form: "Let V be a finite set and let $E(V) = \{\{u, v\}|u, v \in V, u \neq\}$. A pair ... is called a graph, the

elements of V are called **vertices** of G and the elements of E are called **edges** of G.

2) In Definitions 3.2 and 3.7, "E_A : ..." and "E_F : ..." should be replaced by "$E_A \subseteq$ ..." and "$E_F \subseteq$ ...", respectively.

3) The formulation of the last item of Definition 3.2 is confusing, it should be formulated as follows: "... is a function assigning a configuration to each node representing a device."

4) In Definition 3.3, "Partial function" should be written instead of "Function".

5) Definition 3.9 defines the concept of a graph, not a graph production as written by a mistake. Therefore, "production" should be deleted. The concept of a graph defined here means what is often called a directed multigraph in the graph literature. To avoid a confusion, I propose to delete the definition 3.1 because, otherwise, there would be two different definitions of a graph.

6) Conversely, Definition 3.11 defines the concept of a (typed) graph production, not (typed) graph as wrongly written. Thus, "graph production" should be written instead of "graph" at the beginning of the Definition.

7) Since the definition of a graph is presented (Definition 3.9), for reasons of consistency, it would be convenient to also include a definition of the used concept of a graph morphism.

8) While computer scientists are usually well familiar with the graph theory, this may not be true for category theory. It would, therefore, be desirable to present a brief definitions of a category and a double pushout.

The imprecisions listed above have a marginal character only and do not influence substantially the overall high quality of the thesis.

To draw a conclusion, I would like to express my opinion that the quality of the thesis shows the author's advanced level of scientific knowledge based on his independent research. The methods used in the thesis are clearly shown and the results are presented, interpreted and discussed within the framework of the current knowledge of the topic. The bibliography included is complete, containing all significant papers and books closely related to the problems dealt with.

For the above mentioned reasons, I recommend to accept the thesis as a proof of Ing. Gayan de Silva's scientific capability and competence necessary for being awarded an academic degree of Ph.D.

Brno, December 30, 2011

Prof. RNDr. Josef Šlapal, CSc.
Deapartment of Mathematics
Faculty of Mechanical Engineering
Brno University of Technology
616 69 Brno, Czech Republic