USING FORMAL ANALYSIS APPROACH ON NETWORKS WITH DYNAMIC BEHAVIORS

Gayan de Silva
Doctoral Degree Programme, FIT BUT
E-mail: xdesil00@stud.fit.vutbr.cz

Supervised by: Miroslav Švéda, and Petr Matoušek
E-mail: {sveda, matousp}@fit.vutbr.cz

ABSTRACT

Modern computer networks are complex and it is impossible for a human to predict all potentially dangerous situations. Formal modeling and analysis help to overcome this problem. In this paper, a novel approach of formal modeling and analysis of reachability and security properties over dynamic networks is introduced. Dynamic networks are configured with dynamic routing protocols such as RIP, OSFP, EIGRP, or BGP. In case of a device or a link failure, consequent topology changes appear and response of the network can be different. In this paper, we show how network reachability properties can be analysed using formal approach.

1 INTRODUCTION

Current scanning and testing tools are useful for analysis of the existing stable networks, or for analysis of a network after its topology changed. These tools cannot predict different network topologies in case of link or device failures. Therefore, formal network analysis becomes a demand and necessity for the industry. This requirement was identified and addressed by the ANSA project [5]. The ANSA project has shown that it is possible to develop a model that combines static and dynamic behavior using previously introduced techniques described in [4], [1], or [2]. This paper uses a unified model defined in [4], but employs different approach for analysis. Unlike [5], the analysis does not pre-compute all possible routing configurations in order to verify network reachability properties.

2 FORMAL MODEL OF THE NETWORK

The aim of this section is to built a formal model of the network. The network is described using extended graph theory, as shown in Figure 1.

Definition 1 *(Network).* A network is defined as a tuple $N = \langle R, L \rangle$, where R is a finite set of network devices, and $L \subseteq RxR$ is a finite set of physical connections between a given device to its adjacent devices, such that every physical link between two adjacent devices R_i and R_j is a pair of channels $l_{ij} = \langle R_i, R_j \rangle$ and $l_{ji} = \langle R_j, R_i \rangle$.

In real networks, there are many different network devices. For our analysis, we consider every end-point device like a PC or a Web server as a router with one interface. For our network in

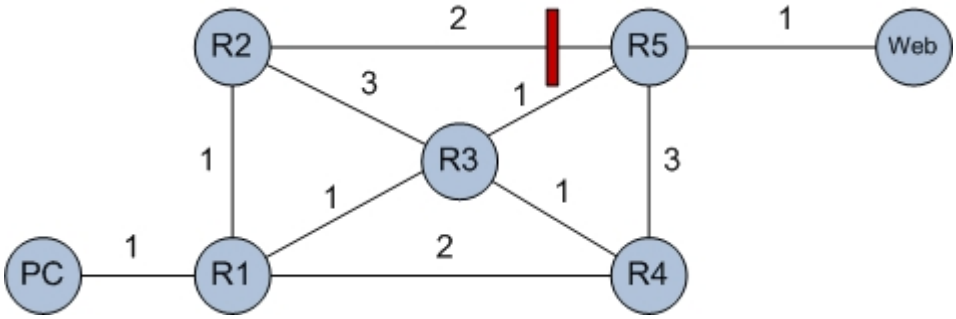

Figure 1: Example of Network Topology - routers, links, an ACL

Fig.1, $R = \{ R_1, R_2, R_3, R_4, R_5, PC, Web\}$, and $L = \{ l_{pc1}, l_{12}, l_{13}, l_{14}, l_{25}, l_{23}, l_{34}, l_{35}, l_{45}, l_{5web}\}$.

Definition 2 *(Next-Hop $NH(R,l)$).* Next-hop is an adjacent router over a given link. Formally, $NH(R_i, l_{ij}) = \{ R_j \in R \mid R_i \xrightarrow{l_{ij}} R_j\}$. For R_1 from our example, $NH(R_1, l_{1p}) = \{PC\}, NH(R_1, l_{12}) = \{R2\}, NH(R_1, l_{13}) = \{R3\}$, and $NH(R_1, l_{14}) = \{R4\}$. Considering point-to-point connection in our model, $NH(R,l)$ is always one router.

Definition 3 *(Cost Function).* For a given link l , we define a function $C(l) : L \rightarrow N$. In network terminology, it is called a metric. C is computed based on the policy of dynamic routing protocols to select the best path. It uses link distance, bandwidth, delay, packet loss rate, or their combination. For routing protocol RIP [3], the cost function is mapped to a distance measured by number of hops on the entire path, e.g., $C(l_{25}) = 2$, $C(l_{23}) = 3$, etc.

Definition 4 *(Filtering Function).* $F_l(p) : ACL \rightarrow boolean$ is a filtering function that for a given packet evaluates filtering rules ACLs over link l. ACL stands for Access Control List, a list of filtering rules (firewall rules). Formal definition and evaluation of $F_l(p)$ can be found in [5]. Here, we demonstrate F on an example. Suppose an ACL composed of two rules *1) access-list 101 permit tcp any host Web dst-port eq https*, and *2) access-list 101 deny ip any any* applied in router R_5 over link l_{25}. Suppose network reachability verified for a packet p = {source-ip = PC, destination-ip = Web, source-port=any, destination-port=www}. Thus, $F_{l_{25}}(p) = 0$ (e.g., False, deny), because http connection is not allowed by the ACL. For further operations, see [5].

Formal specification of our network model according to Fig.1 is composed of links, costs and filtering functions, as follows:

R_i	R_j	$C(l_{ij})$	$F_{l_{ij}}(p)$	R_i	R_j	$C(l_{ij})$	$F_{l_{ij}}(p)$
PC	R1	1	1	R2	R5	2	0
R1	R2	1	1	R3	R4	1	1
R1	R3	1	1	R3	R5	1	1
R1	R4	2	1	R4	R5	3	1
R2	R3	3	1	R5	Web	1	1

3 ANALYSIS APPROACH

This section explains in detail the way how results of network reachability can be obtained for the formal model defined above. Verification of network reachability properties of the model

can be seen as the analysis of two sets of points, namely critical and universal (see further), and relations between them. The main objective of a dynamic routing protocol is to find the best available path for the network data transfer. Since routing protocols cannot establish a virtual path without a physical connection, building routing tables and then analyzing the best path for every combination of links states (up and down) as proposed in [5] is too complex in time and space that it seems to be unfeasible for larger networks. In our approach, all available physical paths are examined and the best path is selected using the cost function. This straightforward analysis of links and routers on a path employs sets of critical points (CP) and universal points (UP). Knowing these sets, we can determine a set of routers that are critical for network reachability. We can also select links and routers which does not play any role in topology changes. In our approach, routing analysis is hidden under the next-hop and cost function.

3.1 COMPUTING REACHABILITY

This section defines sets of critical and universal points, and relations between them. Using this formal model and considering the network status (links up/down), network reachability and security properties can be easily concluded.

Definition 5 *(Path π).* A path π is defined as a sequence of links and routers along an available physical connection between a source and a destination. There can be multiple paths. Paths are computed by concatenation of physical links between adjacent devices starting from source to destination. Let R_0 be a source, and R_n be a destination of path π, then, the k-th existing path between R_0 and R_n is defined as follows: $\pi^k_{<R_0,R_n>} = R_0 l_1 R_1 .....R_{i-1} l_i R_i .....R_{n-1} l_n R_n$ such that $\forall i, 0 \leq i \leq n$: $NH(R_{i-1}, l_i) = R_i$. An example of two paths between PC and Web is given: $\pi^2_{<pc,web>} = PCl_{pc1}R_1 l_{12} R_2 l_{25} R_5 l_{5web} web$, and $\pi^4_{<pc,web>} = PCl_{pc1}R_1 l_{12} R_2 l_{23} R_3 l_{35} R_5 l_{5web} web$.

While identifying paths in the graph model, loops has to be eliminated. If $NH(R_{i-1}, l_i)$ is matching device R_j and R_j has been previously passed along the path π, then there is a loop on π. In another words, only unique R's and L's should be included in π, i.e., path $\pi = R_0 l_1 R_1 l_2 \ldots R_{n-1} l_n R_n$ has no loops iff $\forall i, j : 0 \leq i, j \leq n : R_i = R_j$, and $l_i = l_j$ if and only if $i = j$.

Now, we extend our definition of paths to an evaluated path. Let $P = \{\pi, C(\pi), F_\pi(p)\}$ be an evaluated path, where $C(\pi)$ is a cost function over path π, and $F_\pi(p)$ is a filtering function over path π. Total cost along the path π can be expressed as follows: $C(\pi) = C(l_0) + C(l_1) + \ldots + C(l_n)$, where $l_0, l_1, \ldots, l_n \in \pi$. Two examples of Fig. 1 are given: $C(\pi^2) = C(l_{pc1}) + C(l_{12}) + C(l_{25}) + C(l_{5web}) = 1 + 1 + 2 + 1 = 5$, $C(\pi^4) = C(l_{pc1}) + C(l_{12}) + C(l_{23}) + C(l_{35}) + C(l_{5web}) = 1 + 1 + 3 + 1 + 1 = 7$.

Filtering function $F_\pi(p))$ over path π is defined similarly: $F_\pi(p) = F_{l_0}(p) \wedge F_{l_1}(p) \wedge \ldots \wedge F_{l_n}(p)$, where $l_0, l_1, \ldots, l_n \in \pi$. The result of the total filtering function is conjuction of filtering functions over links along the path π. For our example and packet p defined above, $F_{\pi^2}(p) = F_{l_{pc1}}(p) \wedge F_{l_{12}}(p) \wedge F_{l_{25}}(p) \wedge F_{l_{5web}}(p) = 1 \wedge 1 \wedge 0 \wedge 1 = 0$, or $F_{\pi^4}(p) = F_{l_{pc1}}(p) \wedge F_{l_{12}}(p) \wedge F_{l_{23}}(p) \wedge F_{l_{35}}(p) \wedge F_{l_{5web}}(p) = 1 \wedge 1 \wedge 1 \wedge 1 \wedge 1 = 1$.

It can be shown that the results can be computed easily using simple algorithm. For our example in Fig. 1 and packet p, P is computed in the following table.

k	$C(\pi^k)$	$F_{\pi^k}(p)$	π^k
1	4	1	$PC, l_{pc1}, R_1, l_{13}, R_3, l_{35}, R_5, l_{5web}, web$
2	5	0	$PC, l_{pc1}, R_1, l_{12}, R_2, l_{25}, R_5, l_{5web}, web$
3	6	1	$PC, l_{pc1}, R_1, l_{14}, R_4, l_{43}, R_3, l_{35}, R_5, l_{5web}, web$
4	7	1	$PC, l_{pc1}, R_1, l_{12}, R_2, l_{23}, R_3, l_{35}, R_5, l_{5web}, web$
5	7	1	$PC, l_{pc1}, R_1, l_{13}, R_3, l_{34}, R_4, l_{45}, R_5, l_{5web}, web$
6	7	1	$PC, l_{pc1}, R_1, l_{14}, R_4, l_{45}, R_5, l_{5web}, web$
7	8	1	$PC, l_{pc1}, R_1, l_{13}, R_3, l_{32}, R_2, l_{25}, R_5, l_{5web}, web$
8	10	1	$PC, l_{pc1}, R_1, l_{12}, R_2, l_{23}, R_3, l_{34}, R_4, l_{45}, R_5, l_{5web}, web$
9	10	1	$PC, l_{pc1}, R_1, l_{14}, R_4, l_{43}, R_3, l_{32}, R_2, l_{25}, R_5, l_{5web}, web$

3.2 REACHABILITY ANALYSIS

This section describes analysis of the define network model. It mainly concludes of reachability, security and the topology changes using the constructed formal model. For the analysis, we define two sets.

Definition 7 *(Critical Points CP).* CP is a subset of routers and links that are present on every possible path. The set is defined as follow. $CP = \{R^c, L^c\}$, where R^c represents critical devices and L^c represents the critical links such, that $R^c = \{r \mid \forall \pi \in P : r \in \pi\}$, and $L^c = \{l \mid \forall \pi \in P : l \in \pi\}$.

Definition 8 *(Universal Points UP).* UP=$\{R^u, L^u\}$ is a subset of routers and links that are not present on available paths connecting devices under network reachability analysis. The links and routers of UP have no effect on topology changes and behaviour of routing protocols. This is useful to indicate whether there will or will not be topology change due to link or device failures. The set is defined as follows. First, lets define sets R', and L' such, that $R' = \{r \mid \exists \pi \in P : r \in \pi\}$, and $L' = \{l \mid \exists \pi \in P : l \in \pi\}$. The universal devices and links are complements to R', resp., L', i.e., $R^u = R - R'$, resp. and $L^u = L - L'$.

Lets assume a set of failed devices, resp. links, R^f, resp. L^f. Using sets CP and UP, network reachability can be verified under conditions R^f and L^f.

1. If $\exists r \in R^f$ such that $r \in R^c$ or $\exists l \in L^f$ such that $l \in L^c$, then destination is unreachable.
2. If $R^f \subseteq R^u$ and $L^f \subseteq L^u$, then topology is not changed and network uses the best path which has the least cost $C(\pi^k)$.
3. Otherwise, select π^k where $\forall r \in R^f$ and $\forall l \in L^f$ such that $(r,l) \notin \pi^k$.
   Best Path $\pi_{<pc,web>} = \{\pi^k \in P \mid \forall r \in R^f, \forall l \in L^f, \forall i \in < 0, n >: r \notin \pi^k \wedge l \notin \pi^k \wedge C(\pi^k) \leq C(\pi^i)\}$. If $F(\pi^k) = 1$, then the service is reachable.

First two border cases are easy to find. The most interested is the case no. 3, which require more advance algorithms and optimization procedures to improve the efficiency. Its efficient implementation will be a part of our future work.

Our approach is demonstrated on an example. Assume links l_{13} and l_{35} be down. So, available paths are as follows:

k	$C(\pi^k)$	$F(\pi^k)$	π^k
2	5	0	$PC, l_{pc1}, R_1, l_{12}, R_2, l_{25}, R_5, l_{5web}, web$
6	7	1	$PC, l_{pc1}, R_1, l_{14}, R_4, l_{45}, R_5, l_{5web}, web$
8	10	1	$PC, l_{pc1}, R_1, l_{12}, R_2, l_{23}, R_3, l_{34}, R_4, l_{45}, R_5, l_{5web}, web$
9	10	1	$PC, l_{pc1}, R_1, l_{14}, R_4, l_{43}, R_3, l_{32}, R_2, l_{25}, R_5, l_{5web}, web$

Following steps above (omit the blocking path, i.e., $F(\pi^k) = 0$), the best path for the given network state between the PC and the Web is path 6, which has the least cost and match the first row from the top.

4 CONCLUSION

The paper introduces a new approach to analyze dynamic behavior of computer networks without a need to compute routing table for all states of links. The paper has shown steps how to build a formal model by extracting main properties from a network. According to other approaches, mainly [5], routing tables don't need to be computed since we distinguish logical paths and work with available physical paths. In our approach, the cost function is used to select the best path among the available physical paths. This behavior is similar to dynamic routing protocols as they identify the best path after routing tables update. In the analysis, we showed that the network reachability for any status of links (and topology) can be easily computed.

5 FUTURE WORK

Our future work is oriented on research of analysis techniques for complex networks. Also, the feasibility of model checking approaches for the introduced technique will be examined. Once it is identified, the technique will be added into simulation tool OMNeT++ as an alternative approach to simulation since it gives more precise predictions which can be used in real time practical environments of network behaviors for different routing protocols. Finally, we want to incorporate this formal model to a tool which can read on-line device configurations for a given network and verify reachability and security properties as desired by the user, and as defined in network policy and recommendations.

REFERENCES

[1] D. Antoš. *Hardware-constrained Packet Classification*. PhD thesis, Masaryk University, 2006.

[2] M. Christiansen and E. Fleury. An Interval Decision Diagram Based Firewall. In *3rd International Conference on Networking (ICN'04)*. IEEE, February 2004.

[3] C.L.Hedrick. *Routing Information Protocol*. RFC 1058, June 1988.

[4] G. G. Xie et. On static reachability analysis of ip networks. In *INFOCOM*, pages 2170–2183, 2005.

[5] P. Matoušek, J. Ráb, O. Ryšavý, and M. Švéda. A formal model for network-wide security analysis. In *In the 15th IEEE Symposium and Workshop on ECBS*. On-line on http://www.fit.vutbr.cz/research/view_pub.php?id=8554, 2008.