Diploma Thésis Review

Thesis Author: Bc. Peter Hrošš̆o

Thesis Title: **Black-Box Attack on Network Intrusion Detection Systems**

Reviewer: RNDr. Petr Somol, Ph.D. (ÚTIA AV ČR, Cisco Systems)

I read Mr. Hrošš̆o's Master Thesis with great interest. The topic of the thesis touches subjects that are no doubt in focus of cyber security teams all over the world. Despite the importance of this family of topics, the amount of general knowledge can still be considered as very limited in terms of deeper formal understanding of the key questions. Mr. Hrošš̆o's thesis contains material of both experimental and formal depth and as such is a valuable addition to the study of the subject.

More specifically, the thesis aims at defining and optimization of a network attack so as to maximize attack strength while preventing it from being discovered by a behavioral analysis based detector. The scope of the thesis is thus limited to a particular problem setting. The setting is realistic though, and for a master thesis more than sufficient.

Solving the problem as stated in the introduction of the thesis proved to require non-trivial investigations in various fields of computer science. The flow of analysis as captured in the thesis shows also some slightly surprising intermediate results and the respective reactions of the author - in all cases to the good of the ultimate goal.

The particular details that caught my attention include the slight twist between pages 15 and 16 – the chapter on Anomaly detection starts by putting emphasis on the advantages of convex detectors and ensembles of detectors, while right in the next page I read that the employed detector showed to be non-convex. The following investigations then include analysis of the problems caused by non-convexity of the particular detector in use; the analysis is correct, I would just prefer to at least comment on where would the original assumptions of convexity lead, if fulfilled by the employed tools. Can some such tools be at least suggested for future investigations ?

In page 18 I read that 5 time windows were used in all experiments. Is there any particular reason for the constant 5 ? Is there space for optimization of this constant ?

In page 21 where you assume that the detection function should not decrease after injection of attack flows – this looks expectable and the experiments indeed do confirm it, but I seem to be somewhat lost regarding what assumptions need to be made about the attack flows. I do not say the claims are incorrect, I would just prefer a slightly more elaborate description of the context in which this is true. The attack would have to be quite clearly recognizable for this to work – is it so under all circumstances ?

When Scenarios 1 and 2 were defined in pages 24 and 25, I immediately wonder whether it would be possible to define next scenarios 3, 4... with each next differing just by one more window shift to the right of the interval on which

model2 and thresholds2 are computed. Can you comment on whether this would further improve robustness at the cost of less accurate model fit, until there would be no overlap of window intervals ?

Having finished reading chapter 3 it came to my mind that given the assumption of behavioral-analysis based detection, the detector learned state depends on the complete traffic in the considered network. Is not this a challenge for the adversary and the proposed method as a whole ? Do I understand it right that the adversary would need snapshots of the complete network traffic, to be able to train the detector for use in attack strength optimization ?

The first paragraph in page 40 is confusing. It is not clear whether a real attack has been captured or whether it was a simulated one.

The comparison of Exhaustive search and Gradient search as described in page 43 makes the reader ask why was it necessary to define both procedures in slightly incompatible way when the purpose would be their ultimate comparison. What difference would it make if the Exhaustive search was allowed to generate combinations with permitted zero number of flows ? I understand this is not a serious problem as it would not affect the results in any crucial way, but the clarity could be improved by such unification, if it indeed is possible.

In Conclusions, it would perhaps make sense to extend the discussion and touch the subject of extensibility of the work to other setups. For instance, does the author see any chance to define similar algorithmic framework for detectors based on other than behavioral analysis ? Or is there any restrictions in case of other behavioral analysis based detectors than the ones employed here ? I understand this would touch a fairly distant subject, but a very short author's take would be beneficial.

The thesis is written in above-average English given the fact that the author is not a native English speaker. I noticed just isolated mistakes like "less IPs", meaning "fewer IPs", and one or two somewhat confusing paragraphs. The structure of the thesis is well defined and generally makes sense, while being occasionally demanding to follow due to the number of involved sub-results and their interactions and dependencies. The claims are well illustrated by graphs, tables and illustrations, correctly references from text. The list of references is reasonable though not extensive.

To conclude: I fully recommend this thesis to be accepted as Diploma Thesis. As for the final marks I hesitated between A and B, while I lean more towards B (very good) as the material is presented in somewhat raw but otherwise intelligently assembled form. The ideas are both relevant and generally correct, while opening many related research questions. All this illustrates the ability of the author to dig deep into the problem and scout for all its relevant aspects and consequences. I believe the author should consider continuing his work and aim for doctoral degree – this thesis illustrates well his promising capabilities in this respect.

Petr Somol
Prague, 26th May 2014