Review of Doctoral Thesis: "OPTIMALIZACE ALGORITMU A DATOVÝCH STRUKTUR PRO VYHLEDÁVÁNÍ REGULÁRNÍCH VÝRAZU S VYUŽITÍM TECHNOLOGIE FPGA"

Thesis author: Ing. JAN KAŠTIL

The submitted doctoral thesis is focused on hardware acceleration of pattern matching in intrusion detection systems designed to guarantee correct and safe function of the computer network. There exists many different detection methods and this thesis discusses real time methods of network traffic analysis. The presented work is topical from the viewpoint of science in this field.

Formally the thesis is divided to ten chapters; first chapter (Introduction) brings detailed overview of state-of-the-art, of the most used methods which are able to detect some intrusion activities in systems. Second chapter presents detailed overview of regular automata, describes methods for construction of deterministic and non-deterministic finite automaton. Third chapter describes requirements for matching the patterns in the area of network security and fourth part of thesis presents some of the common complex architectures based on finite automaton both deterministic and non-deterministic, which are used for regular expression processing. Fifth chapter gives an overview of some hash functions and requirements for implementation of these functions. Sixth part of thesis presents profits of alphabet transformation and describes how it is used for throughput improvement and its demand for hardware resources. Chapter seven deals with experimental evaluation of regular expressions which are used in current network security systems. Main result of the thesis is presented in eights chapter "Perfect hashing based deterministic automata", which describes a novel architecture for implementation of the finite automata with sparse transitions table, based on perfect hashing function. Final two chapters describe and evaluate main experimental results and summarise the conclusion.

It is clear, that there is tremendous amount of work behind this thesis. I have no doubts about knowledge and ability of Ing. Jan Kaštil to solve independently such complex problem, but I also have some comments to the presented work and described algorithms. From the formal point of view I have these remarks:

- Bibliography chapter is not included in the Contents section and the references are not typed in uniform style, author have used maybe all the possible combinations of author names; references are not in accordance with the habitual ISO standard. Some of the references are not traceable.
- List of figures is usually located just after Contents.
- List of shorts, abbreviations, symbols and acronyms is missing.
- There are some sentences relicts in the text, for example last sentence in Czech abstract, or first sentence of Chapter 9; otherwise it is necessary to state that the thesis is written carefully, with minimum mistypes.
- Some "key words" are not written uniformly ("finite automata" vs. "Finite Automata" etc.)
- Some parts of text (Figures, Chapter names, Equation) exceed right text frame border.

Although this thesis is written carefully, elaborately, there are also some objections to the content. After first reading of this work I was not sure what the thesis goal is, so whether the described definitions, algorithms, theorems and experiments fulfilled the thesis targets. It is also not clear what part of the described topics is the author's contribution. Another food for thought is the fact, that the solved topics were not published on the "network" conferences (computer networks, network security). Maybe this is the reason why conference papers published by thesis author have so small reaction of the relevant professional community. My next objection is provoked by the thesis title "Optimization of algorithms and data structures for regular expression matching using FPGA technology". It is a pity that the hardware implementation is described only briefly in chapter 8.1, key word "VHDL" is mentioned only once in the whole thesis text, source code is fully missing so it is not obvious whether the described experiments were realised only on FPGA circuits and who was responsible for the hardware implementation.

Questions for the Defence:

- What part of the work was done by Mr. Kaštil?
- All the experiments described in Chapter 9 were realised on FPGA circuits and which concrete FPGA was used? What tool was used for implementation?
- Is there some opportunity to improve the results by parallel processing in FPGA?
- In Chapter 9.1 there is written that algorithm finding perfect hash functions fails after 10 iterations – what happens after that fail?

Though there are some comments to the assessed doctoral thesis, I am fully convinced that Ing. Jan Kaštil has proven his ability to carry out scientific work, to work independently and that he is capable to come to original results and evaluate these results.

I support Mr. Kaštil's candidacy to receive the PhD degree. Pursuant to Section 47/(4) of Law 111/98 I recommend this thesis for Defence.

Liberec, November 29, 2015 prof. Ing. Zdeněk Plíva, Ph.D.