Report on PhD thesis "Integration Paradigms for Ensemble-based Smart Cyber-Physical Systems"

Candidate: *Vladimír Matěna*
Reviewer: *Jan Carlson, Professor, Mälardalen University, Sweden*
Date: *August 20, 2018*

Content and Contributions

The PhD thesis addresses the challenges of developing software for cyber-physical systems where a large number of loosely coupled embedded systems, interacting with each other and their environment, collaborate to provide the desired functionality. Specifically, it focuses on the class of *ensemble-based systems*, where system functionality is defined through dynamically constructed groups consisting of components that satisfy some given constraint and as a result share information in order to contribute to the orchestration if the overall system behavior. Such systems imply significant demands on the communication between the nodes of the system, in order to correctly form ensembles and to share information within them, which is problematic in case of many cyber-physical systems where communication has to be done over networks that are unreliable, heterogenous and continuously changing. The thesis investigates methods to limit and adjust communication in order to achieve as high system performance as possible given the limitations of the underlying networks.

The first chapters introduce the domain and gives a high-level motivation to the thesis topic, and this is then elaborated further, giving a more in-depth description of three industrial domains where the work is applicable and existing solutions and technologies in the area. The discussion leads up to a number of challenges that must be addressed in order for the ensemble-based paradigm to meet the demands of the domains. From these challenges, three specific research questions are identified, focusing on network communication supporting the ensemble-based approach in cyber-physical systems.

Next, a long list of representative use-cases from multiple application domains is presented, focusing on the domains of local and global vehicle coordination. This is followed by concrete descriptions how ensembles can be used to define system behaviors needed for a selection of these use-cases.

The main contribution of the thesis is presented in Chapters 8, 9 and 10. First, a grouping mechanism is introduced by which communication can be limited to statically identified system subsets in a way that ensembles are still correctly formed. Although this is a fairly straightforward approach, it has the potential to significantly reduce the communication requirements in many system instances. Next, this approach is extended with a concept of dynamically changing communication boundaries, which may omit certain ensemble possibilities but ideally only those with low fitness or those that are difficult to synchronize due to network conditions. The third contribution is a mathematical formulation of properties relevant to an analysis the temporal properties, in particular constraints for ensuring that all deadlines are met.

Chapters 11 to 13 describe three concrete implementations that include the ideas presented in the thesis, based on three different programming languages (Java, C++ and Python) two of which are new implementations and one being an extension of a previously developed framework. This is followed by a description of a developed testbed supporting the experimentation and evaluation of the adaptation mechanisms proposed in the thesis.

Based on the testbed three experiments are presented, addressing the methods introduced in Chapters 8, 9 and 10, respectively.

Presentation and writing style

There are quite many minor mistakes, such as grammar errors and typos. Although they do not affect the understandability of the main argumentation much, they sometimes cause unnecessary confusion.

The overall structure of the thesis is mostly clear, but the separation of previously existing solutions and the specific contributions of the thesis could be emphasized more. For example, Chapter 6 gives an overview of the solutions used to address the thesis goals, then Chapter 7 presents two concepts (bipartite and intelligent ensembles) that are not part of these new solutions but published previously by other researchers, before Chapters 8-10 return again to the contributions of the thesis.

Regarding the detailed presentation, it would have been helpful if the concrete proposed new solutions, as well as the claims about their benefits and positive properties were more clearly separated from the general texts discussing the problem, possible approaches and specific examples.

Concerns and Questions

The research questions and goals are formulated at a very high-level, making them descriptions of general topics of interest rather than questions to which the contributions of the thesis can provide well-validated answers.

The descriptions of the technical contributions are often based on quite abstract explanations of the proposed method, complemented and illustrated by a concrete example. This is good for conveying the overall idea, but in some places, there is a lack of technical details defining exactly what the proposed method is. This is especially true in Chapters 8 and 9. For example, the general idea of communication groups is well described, but *"related to each other in terms of a specific knowledge value"* leaves some questions. Are the group definitions limited to checking if a single knowledge value is equal (as in the example) or can other relations be used? The example also defines the group based on the knowledge of the coordinator. It that always the case, and would it make any difference it we for the group from *member.destination.CityDistrict*?

For some parts of the related work, in particular 14.2, it is not clear why certain frameworks or component models have been included. For most of them, nothing is said about their network architecture (which is the title of the section), and since most of them don't support dynamic ensemble creation, they seem outside the scope of the research questions of this thesis.

Finally, I really miss a clear, explicit summary of the contributions of the thesis in terms of (partial) answers to the research questions. The conclusion chapter summarizes the work done but does not return to the research questions.

Detailed questions:

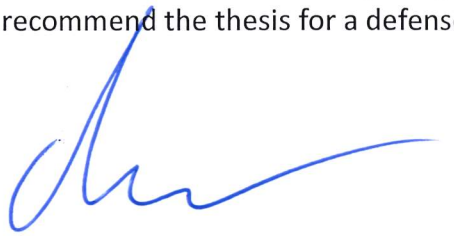
1. To what extent has the two part of research question Q2 been answered by the contributions in the thesis?

2. If the test bed is a scientific result on its own (not just a means to validate the other results of the thesis), how has it been validated? Can you claim that this particular solution is a satisfying answer to Q3?
3. The research questions are stated in a very general way, but most of the work has been done in the context of a specific language and a specific use case scenario. To what extent are the different results in the thesis generalizable to other ensemble frameworks and to other use cases?
4. In 9.4.2, it is stated that the communication boundary is an over-approximation of the membership **by design**. What mechanisms are enforcing this?
5. What are the assumptions behind the equations in Chapter 10? For example, the equations in 10.3 and 10.4 assume loss-less communication, but this assumption is not explicitly mentioned until 10.5. Are there any other implicit assumptions?
6. Equation 10.4 seems incorrect. As mentioned in the text below, the sum of higher priority tasks in the interval **plus the execution time of the task itself**, must be lower than the period, but the equation only includes the higher priority tasks. Would the equation be correct if T-hat is redefined to be the tasks of higher **or equal** priority?
7. What is the overall conclusion from the result from the experiment in 13.3? Section 13.3.2 states that the results indicate that optimization of the system utility by setting different communication parameters is possible, but that seems to be a very weak overall conclusion (it is hard to imagine any result that would show that such an optimization is impossible).

Overall Judgment

Ensemble-based development of cyber-physical systems is a promising approach, especially for loosely coupled systems. The thesis' investigations of the relaxation of maintaining a fully synchronized system state, at the cost of very high network demands, and instead form ensembles and coordinate their constituents based on a limited communication is an important contribution towards making this paradigm applicable in more complex, realistic settings. Although implemented in the DEECo framework, the developed methods seem generally applicable also in other ensemble frameworks. The thesis is based on several peer-reviewed publications, including one Journal article, which indicates a relevance of the work to the scientific community.

The thesis could be more precise in isolating the specific contributions, and would benefit from a more in-depth validation and evaluation of the proposed approach in relation to the formulated research questions, but in summary, the thesis shows the candidate's capability of creative scientific work, of general relevance to the scientific community. Thus, I recommend the thesis for a defense and judge the candidate worthy of a PhD degree.



Jan Carlson, Professor, Mälardalen University, Sweden
August 20, 2018