Probabilistic Analysis and Predictions of Component-Based Real-Time Systems

Anders Möller[‡⋆] Mikael Nolin[‡⋆] Ian Peake[†] Heinz W. Schmidt[†]

[‡]MRTC, Mälardalen University, Västerås, Sweden
[†]Monash University, Melbourne, Australia
[⋆]CC Systems, Uppsala, Sweden

E-mail: Anders.Moller@mdh.se

Abstract

Software components are suitable vehicles to introduce advanced analysis techniques in a software-engineering context for embedded control software; a feat that has yet to be fully accomplished.

We are adopting a component-based approach to control software development. We are extending and combining methods from disparate disciplines, such as probabilistic reliability predictions, stochastic scheduling analysis and software component technologies. We study theories and methods for probabilistic modelling, analysis, and prediction of control software executing in resource-constrained embedded computers.

Combining the behaviour models and the architectural model of a component assembly we are deriving stochastic properties, such as reliability, expected delays, and resource consumption. Using components as the fundamental unit of reuse, we employ run-time monitoring techniques to extract probabilistic models of the component behaviour.

1 Introduction

Developers of embedded real-time control-system face challenges of (*i*) high demands on reliability and performance (*ii*) requirements on lowered product cost, and (*iii*) supporting many configurations, variants and suppliers. Computer systems offer the performance needed for the functions requested, but at the same time product reliability and safety must not suffer. Unfortunately computers add new sources of failures, and therefore can cause lessened product reliability. This yields a strong focus on the ability to *model*, *predict*, and *verify* software functionality, reliability, and safety. However, many companies lack the technical support to verify their system behaviour (both with respect to functional and extra-functional aspects of the software) - and black-box testing together with manual code inspections are usually the only methods used to "confirm" system functionality and reliability.

Also, in order to keep the software development costs within budget, more and more Original Equipment Manufacturers (OEMs) use sub-contractors (and/or Commercial-Off-The-Shelf (COTS) Components) to develop various parts of their systems. This increases complexity of system analysis and further jeopardises software system trust. Due to the potentially high (economic and/or safety) impact of software failures (e.g., passenger safety in a car) – predictable software becomes increasingly important.

The real-time systems of interest to us are distributed, and therefore require dealing with parallel behaviour (or at least communication of a sequential component with parallel components in its deployment environment). Our work focuses initially on the sequential components and the behaviour of multiple components executing on the same embedded controller. Therefore asynchrony and scheduling issues are inside the boundaries of our project while true parallelism is outside.

2 Project Aims

In this project, we study theories and methods for probabilistic modelling, analysis, and prediction of control software executing in resource-constrained embedded computers. The overall goal is to come up with methods that will result in a decrease in system life-cycle costs, and an increase in system quality.

Using components as the fundamental unit of reuse, we employ run-time monitoring techniques to observe the systems and extract probabilistic models of the component behaviour [9]. Combining the behaviour models and the architectural model of a component assembly, we derive stochastic properties, such as reliability, expected delays, and resource consumption. We are extending and combining methods from disparate disciplines, such as probabilistic reliability-predictions [1, 7] stochastic scheduling-analysis [5, 6, 11], run-time monitoring [2, 9, 10], and software-component technologies [3, 4].

The data obtained from monitoring the components are

used to address the following key areas in software engineering of embedded control-systems:

- **Model-extraction**. By monitoring component-based software, information about the component properties can be extracted. The properties are used to annotate the component models. In this way, extra-functional component properties are obtained and refined without engineering effort. These model properties provide a basis for trust in components and can be used for system-level predictions.
- **System-level predictions**. By using models of the component behaviour, key properties such as reliability, timing behaviour and resource consumption can be predicted. Early prediction of such properties can be used to guide system designers and aid in dimensioning hardware resources, hence reduce the development effort and increase quality.
- **System-level testing and debugging**. By monitoring individual components and component interactions, errors can be found and traced. Monitoring can also be used to support replay debugging [10], where erroneous system-executions are recreated in a lab environment to allow tracing of bugs. Testing and debugging of embedded systems are notoriously difficult and time-consuming. Hence, increased support for test and debug has great potential for cost savings.
- **Run-time contract checking**. The run-time monitoring allows surveillance of third party components. Both functional (e.g. range of output values) and extra-functional (e.g. memory usage) properties can be monitored. During acceptance testing, the contract checking is used to validate that a component does not violate its specification. In systems that fail after deployment, logs from the contract checking can be used in post-mortem analysis. The possibility for post-mortem assessment of contract breaches is likely to increase the willingness to deploy third-party components in critical systems, since such assessments can be coupled to financial liabilities.
- **Observability**. Computer systems in general and embedded systems in particular, are infamous for the difficulty of observing their internal behaviour. This has drawbacks throughout the debugging, testing and maintenance phases. Systems that are unobservable become very difficult to analyse and validate. Also after deployment, observability is an important feature, allowing inspection and performance tuning of running systems.

To realize the above goals, this project contains two main strands of research:

✓ Development of monitoring techniques suitable for resource-constrained embedded systems. Monitoring these systems require lightweight techniques to be developed. Also, the number of metrics (and their quality/granularity) should be limited. This presents a trade-off between achieved results of the monitoring and the resources consumed.

✓ Development of techniques to predict system properties from component properties. We focus our research on stochastic methods such as reliability prediction and probabilistic scheduling-analysis. Whether to deploy stochastic or deterministic (worst-case) predictions is a strategic decision for each project. However, the majority of systems are not designed for worst-case scenarios; it is not possible to motivate the cost incurred by dimensioning and validating the system for the worst-case scenario. Hence, our focus is on stochastic methods that allow designers to make well founded trade-offs between, e.g., reliability and production cost.

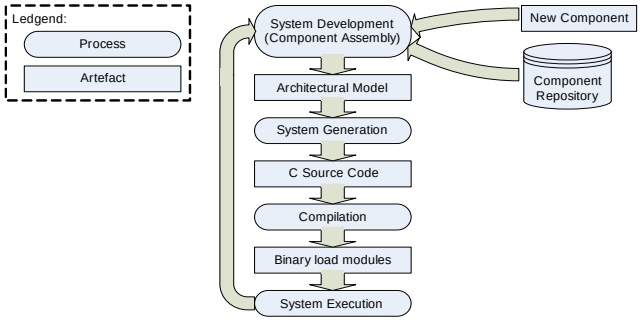

Figure 1. Traditional CBD Processes for Embedded Systems

3 Component-Based Development

Component-Based Development (CBD) for resource-constrained embedded real-time systems differs from general CBD for Internet/office applications where components are individually compiled and deployed. These components are loaded at run-time, and bindings between components are dynamically created by a middleware. Figure 1 illustrates the typical process when using CBD for *embedded systems*. Systems are developed by assembling components from a repository, or by using newly developed components. In the next step a system generator generates source code (typically C code) where component bindings are explicitly represented and connected. Finally the code is compiled and the system executed. Upon the detection of bugs or undesired system properties, the system designer goes back and modifies the system.

Using the results of this project, a much more attractive process can be obtained. Figure 2 illustrates the new possibilities with highlighted activities. Firstly, using analysis, early assessments about system properties can be made, also – if the system is redesigned undesired properties can be detected. Secondly, the system generator can be augmented to automatically insert the instrumentation code needed to per-

form run-time monitoring. Finally, dynamic properties for the components are extracted during run-time. These properties are then stored together with the components in the repository and can be used in subsequent development cycles.

4 Research Topics

This project draws upon existing results from several research directions. The main challenges will include the identification of suitable techniques and a modification of these techniques in order to better suit component-based development of embedded systems.

4.1 Software Component Monitoring

Current approaches rely on software instrumentation to produce logs. Part of this instrumentation can be placed from operating system functions, which minimises the manual effort for instrumentation. However, not all the data required can be extracted in operating system calls, and manual instrumentation of application code is required. Furthermore, to make sure that all the needed data is collected – expert knowledge of the application is required.

Component-based development gives us the opportunity to automatically add code instrumentation. This is possible if the component model makes explicit all data flows and state variables, and this is the case for the component models considered in this project [4].

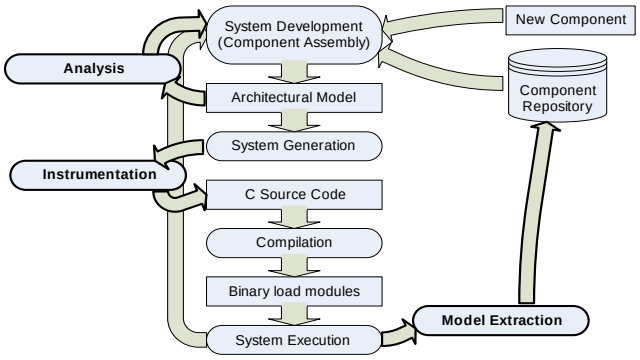

Figure 2. Our proposal to CBD for Embedded Systems

Some specific questions studied within this project are:

- Monitoring embedded systems require resource conservative monitoring techniques. Are current techniques good enough? Can they be modified to better suit embedded real-time systems?
- Since resources are constrained, the number of metrics (and their quality/granularity) should be limited. This presents a trade-off between achieved results of the monitoring and the resources consumed. A key problem in this project is to identify the metrics that have the highest beneficial impact on the software engineering process.
- Ideally, metrics should be collected and refined throughout the life-cycle of a component. However, resource limitations may hinder this ideal solution since different metrics could be more important during different phases. This presents a trade-off of which metrics to obtain during what stage. A goal of this project is to identify which metrics are most important during what stages.
- The problem that arises with monitoring is the large number of executions necessary to establish a reasonable statistical confidence. Due to the complex behaviour of software components, traditional statistical methods to estimate the statistical validity, such as confidence intervals, may have to be used together with more domain specific validity metrics. We will investigate to what extent traditional statistical methods are suitable and, if needed, try to find methods to complement them.

4.2 Reliability Predictions

We focus on reliability predictions using Markov models, as described in, e.g., [1, 7]. However, such methods require that a system has a well-defined start and end state. Unfortunately, this is not the case for reactive control-systems. In addition, the Markov behaviour of software components has to be mapped to the execution histories and their stochastic properties. This requires blending automata models with Markov chains.

Our prior research, e.g. [7, 8], has shown that reasonably accurate prediction of these system properties can be achieved by parameterisation. Component-level, extra-functional, models are variable with formal parameters ranging over different behaviours or different extra-functional variations.

Interesting research questions include:

- How can the modelling techniques be adapted/extended in order to fit reactive systems, without start and end states? One possible solution could be to introduce additional super-states, defining the start and end states. This, however, needs to be combined with a way of finding recurrent states in the reactive control-system, i.e. system states that will re-occur during execution.
- To facilitate architecture-based reliability predictions, software components need to be equipped with reliability figures. These figures depend on the context in which the component is used. The usage-profile of a specific component in a certain application has influence on the reliability. What type of reliability measure best suits the control-system domain, and what does this imply in terms of changes in the analysis techniques?
- One basic idea is to transfer the architecture description of the component-based system into a Finite State Automata

(FSA), and use reliability figures of the components to represent system reliability. Is it possible to adapt this method, in order to be more fine-grained, e.g., by representing also the actual components with FSAs? In this case, reliability methods can be used to find the weakest part of each of the used components.

- By using component impact analysis, software developers can be guided to put focus on the part of the application that is most crucial in terms of reliability and usage. How can the architecture-based reliability predictions, using a Markov chain representation, be extended in order to guide control-system developers during the design-phase?

4.3 Stochastic Schedulability Analysis

Traditional methods for stochastic schedulability-analysis assume that execution-times can be described by known probability functions. However, such a description is often misleading since execution times cannot take on arbitrary values. Typically, execution times are clustered around a small set of "probability peaks" which is not easy to represent with general probability functions.

On the other hand, if the representation of the execution-times and their probabilities is too detailed, statistical analysis becomes infeasible due to the combinatorial explosion in the number of possible combinations of execution times.

- We investigate histogram-based schedulability-analysis. A histogram can be adapted to provide the desired level of detail, thus it gives the possibility to reduce complexity of the analysis.
- When performing stochastic schedulability-analysis there are two major sources of errors: (1) errors in the models used (e.g. due to not monitoring the component long enough), and (2) errors due to simplifications/abstractions made during the analysis. The effect of both these errors needs to be quantified and bounded. This means that it is not enough to calculate the result of the schedulability analysis, but also a level of confidence that account for both error source is needed.

5 Conclusion

This project will provide novel models, theories and techniques to monitor and predict the behaviour of resource-constrained embedded control-systems. This class of systems constitutes an increasing fraction of value in many products, such as in vehicles and automation robotics. We fuse scientific methods from disparate disciplines, such as probabilistic reliability-predictions, stochastic scheduling-analysis, run-time monitoring, and software-component technologies. This will result both in novel methods and extensions to existing, well established, methods.

This project takes a unique approach in that we, (*i*) compared to other component-technology projects not focus on developing our own component technology. Rather we focus on the ability to use the architectural- and component-models to facilitate system-level analysis. Our goal is to automate the whole process, thus we do not introduce additional burdens for the software-engineers. And, (*ii*), compared to other analysis-techniques project, we do not focus on generally applicable methods. Rather we constrain our research to methods suitable for component-based systems.

References

[1] R. C. Cheung. A User-Oriented Software Reliability Model. *IEEE Transactions on Software Engineering*, 6(2):118–125, 1980. Special collection from COMPSAC 1978.

[2] S. Chodrow. Run-time monitoring of real-time systems. In *Proc. of IEEE 12^{th} Real-Time Systems Symposium*, pages 74–83, December 1991. San Antonio, USA.

[3] I. Crnkovic. Component-Based Approach for Embedded Systems. In *the 9^{th} International Workshop on Component-Oriented Programming*, June 2004. Oslo, Norway.

[4] M. Åkerholm, A. Möller, H. Hansson, and M. Nolin. Towards a Dependable Component Technology for Embedded System Applications. In *Proceedings of the 10^{th} IEEE International Workshop on Object-oriented Real-Time Dependable Systems*, February 2005. Sedona, Arizona, USA.

[5] J. P. Lehoczky. Real-Time Queuing Network Theory. In *Proceedings of the 18^{th} IEEE Real-Time Systems Symposium (RTSS'97)*, pages 58–67, San Francisco, CA, USA, December 1997. IEEE Computer Society.

[6] T. Nolte, A. Möller, and M. Nolin. Using Components to Facilitate Stochastic Schedulability. In *Proceedings of the 24^{th} Real-Time System Symposium – Work-in-Progress Session*. IEEE Computer Society, December 2003. Cancun, Mexico.

[7] R. H. Reussner, I. H. Poernomo, and H. W. Schmidt. Reasoning about Software Architectures with Contractually Specified Components. *Component-Based Software Quality. LNCS 2693, Springer-Verlag*, pages 287 – 325, 2003.

[8] H. W. Schmidt. Trustworthy components: compositionality and prediction. *Journal of Systems and Software, Elsevier Science Inc*, 65(3):215–225, 2003.

[9] D. Sundmark, A. Möller, and M. Nolin. Monitored Software Components – A Novel Software Engineering Approach –. In *Proceedings of the 11^{th} Asia-Pasific Software Engineering Conference, Workshop on Software Architectures and Component Technologies*, November 2004. Pusan, Korea.

[10] H. Thane, D. Sundmark, J. Huselius, and A. Pettersson. Replay Debugging of Real-Time Systems Using Time Machines. In *Proceedings of Parallel and Distributed Systems: Testing and Debugging*, pages 288 – 295. ACM, April 2003.

[11] T. S. Tia, Z. Deng, M. Shankar, M. Storch, J. Sun, L. C. Wu, and J. S. Liu. Probabilistic Performance Guarantee for Real-Time Tasks with Varying Computation Times. In *Proceedings of the 1^{st} IEEE Real-Time Technology and Applications Symposium (RTAS'95)*, pages 164–173, Chicago, IL, USA, May 1995. IEEE Computer Society.