

Period Optimization and Priority Assignment for Cause Effect Chains

Jian-Jia Chen

Department of Computer Science
TU Dortmund University, Germany
jian-jia.chen@tu-dortmund.de

Mario Günzel

Department of Computer Science
TU Dortmund University, Germany
mario.guenzel@tu-dortmund.de

I. INTRODUCTION

In many applications of distributed cyber-physical systems, a sequence of tasks has to be executed for a certain functionality. For example, from a sensor to an actuator, the first task reads the sensor value (cause), the second task processes the data, and the third task produces an output for the actuator (an effect is triggered). The *end-to-end* timing properties of the sequence of tasks (the so-called cause-effect chain) are very important.

End-to-end latency analysis determines the maximum time for the data flow between the **cause** (external activity) and the **effect** (actuation based on the external activity). Davare et al. [5] explored such end-to-end latency and optimization of *cause-effect chains* in 2007, extended by Feiertag et al. [7] in 2009. Industrial and practical applications of cause-effect chains can be found in AUTOSAR Timing Extensions [1], communication-centric designs by Hamann et al. [14] from Robert Bosch, and Robot Operating System 2 (ROS2) [19].

In the last two decades, extensive results to analyze the end-to-end latency of cause-effect chains have been provided, e.g., [2]–[6], [8]–[10], [12], [13], [15]–[18]. A tutorial by Günzel et al. [11] has been recently published by the authors of this work. Despite extensive results have been provided in the literature for end-to-end latency analyses, they typically assume the periods and the priorities of the tasks are given. Essentially, the periods of the tasks should be assigned just short enough to meet the end-to-end latency constraints of the cause-effect chains. It is surprising that this problem is much less studied. In this work, we discuss an open problem by exploring a simple scenario which is already complicated enough to be handled optimally.

II. OPEN PROBLEM FOR PERIOD AND PRIORITY ASSIGNMENTS

A *cause-effect chain* $E = (\tau_1 \rightarrow \tau_2 \rightarrow \dots \rightarrow \tau_n)$ is a data path through a finite sequence of periodic tasks $\tau_i \in \mathbb{T}$. A periodic task τ is represented by $\tau = (C_\tau, T_\tau, \phi_\tau) \in \mathbb{R}^3$, where the first job of τ is released at its phase ϕ_τ , the subsequent jobs are released periodically with a period T_τ , and every job of τ executes for at most C_τ , denoted as worst-case execution time (WCET), time units. We note that the WCETs of the tasks are specified as part of the input, but the periods and phases of the tasks are not specified as part of the input.

In this work, we limit our discussions with two cause-effect chains with $E_1 = (\tau_{1,1} \rightarrow \tau_{1,2} \rightarrow \dots \rightarrow \tau_{1,n_1})$ and $E_2 = (\tau_{2,1} \rightarrow \tau_{2,2} \rightarrow \dots \rightarrow \tau_{1,n_2})$ to demonstrate why this is an interesting open problem to study. We are given end-to-end latency constraints of the cause-effect chains, denoted as L_1 for E_1 and L_2 for E_2 . The problem is to study how to assign the periods, phases, and priorities of the tasks so that the end-to-end latency constraints are fulfilled. We assume a fixed-priority preemptive scheduler, in which a task is assigned a fixed priority for the rest of this work.

We focus on implicit communication, i.e., each job's read-event is when the job starts and the job's write-event when the job finishes. In this regard, there are two analyses to derive the end-to-end latency $\text{Lat}(E)$ of a cause-effect chain E .

Theorem 1. *If tasks of the cause-effect chain $E = (\tau_1 \rightarrow \dots \rightarrow \tau_n)$ communicate via implicit communication. Then, the end-to-end latency $\text{Lat}(E)$ is upper bounded by $\text{Lat}(E) \leq \sum_{i=1}^n T_{\tau_i} + R_{\tau_i}$, where R_{τ_i} is the worst-case response time of task τ_i , due to Davare et al. [5, Section 2.1]. A tighter analysis by Dürr et al. [6, Theorem 5.4] shows that $\text{Lat}(E)$ is upper bounded by*

$$\text{Lat}(E) \leq \sum_{i=1}^n T_{\tau_i} + R_{\tau_i} - \sum_{i=1}^{n-1} [P_i] \cdot \min(R_{\tau_i}, T_{\tau_{i+1}}), \quad (1)$$

where $[P_i] = 1$ if τ_i and τ_{i+1} run on the same processor and τ_{i+1} has lower priority than τ_i , and $[P_i] = 0$ otherwise.

Let us first focus on one cause-effect chain. With the timing analysis by Davare et al. [5, Section 2.1], it seems quite natural to simply minimize $\sum_{i=1}^n T_{\tau_i} + R_{\tau_i}$. In fact, Davare et al. [5] put their focus also on geometric programming to assign the periods of the tasks in the cause effect chain.

However, there is a natural pitfall by adopting the analysis in Theorem 1. If all the tasks in the cause-effect chain have the same phase 0 and the same period T , we can simply organize the execution of the tasks to pass the data sequentially in the chain by also assigning the priorities so that τ_i has higher priority than τ_{i+1} . Therefore, the end-to-end latency $\text{Lat}(E)$ is upper bounded by $T + \sum_{i=1}^n C_{\tau_i}$ if $T \geq \sum_{i=1}^n C_{\tau_i}$. It is not difficult to see that this is a tight latency bound. Therefore, as long as $T + \sum_{i=1}^n C_{\tau_i}$ is no more than the latency constraint and $T \geq \sum_{i=1}^n C_{\tau_i}$, the period assignment and priority assignment is feasible. This seems to be an optimal solution when there is only one cause-effect chain with implicit communication.

Let us now extend this to two cause-effect chains. Without loss of generality, let $L_1 \leq L_2$. In case **none of the tasks in E_1 and E_2 is the same**, these two cause-effect chains are independent from each other. Since $L_1 \leq L_2$, it is quite natural to assign the tasks in E_1 higher priorities than the tasks in E_2 . Furthermore, to reduce the impact on E_2 , we should simply assign a common period T to the tasks in E_1 such that $T + \sum_{i=1}^{n_1} C_{1,\tau_i} = L_1$ with $T \geq \sum_{i=1}^{n_1} C_{1,\tau_i}$. Similarly, the tasks in E_2 should also be executed by following the priority order so that the data written from $\tau_{2,i}$ can be read as soon as possible by $\tau_{2,i+1}$, if all the tasks in E_2 has the same period and phase. Therefore, by setting their periods to $\hat{T}$ and phases to 0, we know that the $\text{Lat}(E_2)$ is upper bounded by $\hat{T} + R_{2,n_2}$, where R_{2,n_2} is the worst-case response time of τ_{2,n_2} under the above period and priority assignment. If $\hat{T} + R_{2,n_2} \leq L_2$, we derive a feasible solution. This seems to work but we miss some details:

- Should T be simply set to $L - \sum_{i=1}^{n_1} C_{1,\tau_i}$ when $\sum_{i=1}^{n_1} C_{1,\tau_i} \leq L/2$?

Under this assignment, T is maximized and therefore R_{2,n_2} is minimized. Although we do not have a proof yet, this seems to be an optimal solution.

The open problem is to investigate the same problem setting under a more realistic scenario that **some tasks appear in both E_1 and E_2** . It is already unclear how to optimally handle the scenario with only one task appeared in both E_1 and E_2 . One simple heuristic is to first derive a common period T of the tasks in E_1 like the above procedure, and, then, find a common period $\hat{T}$ of the remaining tasks in E_2 such that the end-to-end latency based on Theorem 1 holds. We will discuss more details about this scenario during the open problems seminar.

REFERENCES

- [1] AUTOSAR. Specification of timing extensions (AUTOSAR CP R22-11). https://www.autosar.org/fileadmin/standards/R22-11/CP/AUTOSAR_TPS_TimingExtensions.pdf, 2022. Accessed: 2024-01-18.
- [2] M. Becker, D. Dasari, S. Mubeen, M. Behnam, and T. Nolte. Mechaniser-a timing analysis and synthesis tool for multi-rate effect chains with job-level dependencies. In *Workshop on Analysis Tools and Methodologies for Embedded and Real-time Systems (WATERS)*, 2016.
- [3] M. Becker, D. Dasari, S. Mubeen, M. Behnam, and T. Nolte. Synthesizing job-level dependencies for automotive multi-rate effect chains. In *International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA)*, pages 159–169, 2016.
- [4] H. Choi, M. Karimi, and H. Kim. Chain-based fixed-priority scheduling of loosely-dependent tasks. In *International Conference on Computer Design (ICCD)*. IEEE, 2020.
- [5] A. Davare, Q. Zhu, M. D. Natale, C. Pinello, S. Kanajan, and A. L. Sangiovanni-Vincentelli. Period optimization for hard real-time distributed automotive systems. In *Design Automation Conference, DAC*, pages 278–283, 2007.
- [6] M. Dürr, G. von der Brüggen, K.-H. Chen, and J.-J. Chen. End-to-end timing analysis of sporadic cause-effect chains in distributed systems. *ACM Trans. Embedded Comput. Syst. (Special Issue for CASES)*, 18(5s):58:1–58:24, 2019.
- [7] N. Feiertag, K. Richter, J. Nordlander, and J. Jonsson. A compositional framework for end-to-end path delay calculation of automotive systems under different path semantics. In *Workshop on Compositional Theory and Technology for Real-Time Embedded Systems*, 2009.
- [8] J. Forget, F. Boniol, and C. Pagetti. Verifying end-to-end real-time constraints on multi-periodic models. In *ETFA*, pages 1–8, 2017.
- [9] A. Girault, C. Prevot, S. Quinton, R. Henia, and N. Sordon. Improving and estimating the precision of bounds on the worst-case latency of task chains. *IEEE Trans. on CAD of Integrated Circuits and Systems, (Special Issue for EMSOFT)*, 37(11):2578–2589, 2018.
- [10] M. Günzel, H. Teper, K. Chen, G. von der Brüggen, and J. Chen. On the equivalence of maximum reaction time and maximum data age for cause-effect chains. In *ECRTS*, volume 262 of *LIPICs*, pages 10:1–10:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023.
- [11] M. Günzel, H. Teper, G. von der Brüggen, and J.-J. Chen. End-to-end latency of cause-effect chains: A tutorial. *ACM Trans. Embed. Comput. Syst.*, 24(1):22:1–22:18, 2025.
- [12] M. Günzel, N. Ueter, K. Chen, and J. Chen. Timing analysis of cause-effect chains with heterogeneous communication mechanisms. In *RTNS*, pages 224–234. ACM, 2023.
- [13] M. Günzel, N. Ueter, K. Chen, G. von der Brüggen, and J. Chen. Probabilistic reaction time analysis. *ACM Trans. Embed. Comput. Syst.*, 22(5s):143:1–143:22, 2023.
- [14] A. Hamann, D. Dasari, S. Kramer, M. Pressler, and F. Wurst. Communication centric design in complex automotive embedded systems. In *Euromicro Conference on Real-Time Systems, ECRTS*, pages 10:1–10:20, 2017.
- [15] T. Kloda, A. Bertout, and Y. Sorel. Latency analysis for data chains of real-time periodic tasks. In *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA*, pages 360–367, 2018.
- [16] A. Rajeev, S. Mohalik, M. G. Dixit, D. B. Chokshi, and S. Ramesh. Schedulability and end-to-end latency in distributed ecu networks: formal modeling and precise estimation. In *International Conference on Embedded Software*, pages 129–138, 2010.
- [17] J. Schlatow and R. Ernst. Response-time analysis for task chains in communicating threads. In *IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS)*, pages 245–254, 2016.
- [18] J. Schlatow, M. Möstl, S. Tobuschat, T. Ishigooka, and R. Ernst. Data-age analysis and optimisation for cause-effect chains in automotive control systems. In *IEEE International Symposium on Industrial Embedded Systems (SIES)*, pages 1–9, 2018.
- [19] H. Teper, M. Günzel, N. Ueter, G. von der Brüggen, and J. Chen. End-to-end timing analysis in ROS2. In *RTSS*, pages 53–65. IEEE, 2022.