Transformer for Predictive and Prescriptive Process Monitoring in IT Service Management (Extended Abstract)

Marc C. Hennig [1]

[1] *University of Applied Sciences Munich, Lothstr. 34, Munich, 80335, Germany*

Abstract

Increasingly complex IT environments, requirements, and organizations in IT service management make the development of advanced predictive technologies necessary. Therefore, this paper outlines a Ph.D. project to develop a pipeline supporting novel IT service management approaches using state-of-the-art predictive and prescriptive process monitoring based on transformer neural networks.

Keywords

IT Service Management, ITSM, AITSM, AIOps, Transformer

1. Introduction

IT service management (ITSM) [1] is the processual approach to providing information technology support and services assisting the key activities in organizations and therefore serving the overall organizational goal achievement. ITSM processes define a continuous improvement process framework [1] and operating model for IT organizations by determining services based on customer requirements that are provided by IT infrastructure components. Legal and economic dimensions are added to the services as contractual obligations in service level agreements (SLA). Adherence to SLAs and continual improvement are among the core success factors in ITSM [2]. ITSM processes are complex and difficult to analyze since they are usually embedded in a multi-layered technical and organizational landscape with a high level of specialization [3], [4].

Two emerging fields are assisting in the successful provision of IT services: artificial intelligence-driven ITSM (AITSM) [5] and artificial intelligence for IT operations (AIOps) [6]. AIOps is a data-driven approach for analyzing data to provide operators with the information required to operate complex IT systems efficiently. At the same time, AITSM is the automation, support, and improvement of ITSM processes using machine learning (ML). Insights into processes are important requirements in both approaches.

2. Motivation and Problem

As a collection of interconnected processes embedded in a multi-faceted environment, ITSM is challenged by different stakeholders and demands that must be aligned. Therefore, process instances in ITSM are often characterized by a high degree of flexibility, dependencies, and knowledge intensity, which are necessary to react to disruptions and changes in services and the socio-technological IT organization. The organization comprises so-called configuration items (CI), which are mainly human resources with their responsibilities and hard- and software components. The configuration management database (CMDB) is one of the core elements in an ITSM ecosystem and contains information regarding CIs and their interdependencies and hence offers valuable contextual information. In ITSM, especially the assessment and improvement of the processes and their proper orchestration are central pain points since extracting practical insights on the processes are difficult to obtain which hinders process maturity and hence the service resilience and quality [4], [5]. CMDBs

ICPM 2022 Doctoral Consortium and Tool Demonstration Track, October 23–28, 2022, Bolzano, Italy
EMAIL: mhennig@hm.edu

have been used to analyze some parts of ITSM processes [7]–[9] but have not yet been systematically utilized in combination with event log data.

Due to the unique properties of ITSM processes as complex service processes, this Ph.D. project shall enable transformer neural networks [10] in ITSM and establish a connection between predictive and prescriptive process monitoring and the intersecting fields of AIOps and AITSM. This is particularly interesting since the results which are usually delivered by process monitoring solutions [11], including the next event, its point of occurrence, and the overall duration of a process instance are valuable insights for these fields to provide operational support. Furthermore, concrete recommendations and interactive optimizations on these variables can be derived using prescriptive process monitoring [12] to enable improvements in ongoing instances. Despite the thematic overlaps between predictive and prescriptive process monitoring, AIOps, and AITSM, systematic studies on the intersection are missing.

Predictive and prescriptive process monitoring could benefit from transformer models, a relatively novel approach in ML mainly used in natural language processing (NLP) and other sequence-related tasks. They have outperformed traditional recurrent neural networks and their derivatives in several areas. However, in predictive and prescriptive process monitoring, transformers have only been covered sparsely so far [12]–[14].

Hence, this project aims to understand the fields AITSM and AIOps as novel approaches in ITSM and to integrate predictive and prescriptive process monitoring based on transformer models therein.

3. Research Questions

The success factors for ITSM, as outlined in the previous sections, namely the ability to conform to predefined time and quality-bound SLAs, are directly influenced by process intelligence and the ability to derive actionable insights.

To address these challenges, an end-to-end pipeline for process monitoring using transformer models shall be envisioned, which initiates three research questions. First, the data must be collected and preprocessed; second, the event log must be fed into the transformer model to receive predictions; finally, workable insights should be derived. The insights should benefit the service quality by providing operators with immediate AIOps information to handle events and long-term AITSM support to foster change, resilience, and improvement across the service's life cycle.

Initially, there is the collection and preprocessing of the data, including the extraction and preparation of information from concurrent process instances, like incident management and change management, and other ITSM sources, e.g., the graph-like CMDBs. The inclusion of exogenous data [15], [16] in addition to the event log is deemed necessary to fully capture how an IT organization's performance is affected by the workload and events occurring in different processes and process instances and to identify the key influence factors of service resilience and performance. Additionally, this might be required to attain a sufficient model quality [17]. Therefore, it must be figured out how the data can be optimally prepared for predictive tasks in transformer models and how event data and contextual information from other sources like CMDBs can be integrated.

Research Question 1: What is the proper way to collect and preprocess the event log to account for the complexities of ITSM processes and leverage additional data sources like CMDBs to allow for further processing in transformer models?

Secondly, an architecture using transformer models is to be developed based on the previous research [13], [14]. This architecture should accommodate the event logs' unique sequential properties and underlying processes. Specifically, the non-continuous time and the non-equidistant time intervals between the discrete events make processes different from usual timelines and pose an interesting challenge for transformers. Other approaches than the often-used positional embedding and a special trace encoding [18] might be necessary to make the dependencies between activities and timestamps workable for transformer models [19].

Research Question 2: How can a transformer architecture be designed to be suitable for ITSM event logs and additional data sources to provide predictions and optimizations with continuously generated events from real-world applications?

Finally, the core factors influencing the performance of process instances in ITSM must be detected based on the data, hence adding explanations to the mostly black box [20] results of transformer models. Explainability is essential to enable proactive improvement of the process, the organization, and the analysis of problem sources [21], which has not yet been achieved in prescriptive process monitoring [12].

Research Question 3: How can the root causes impacting the performance within process instances be derived from complementing the predictions and be used for continuous improvement and operational support?

The results of these steps shall then be combined into a pipeline for process monitoring on ITSM event logs that can be used to support IT operations as an AITSM and AIOps solution.

4. Research Methodology

This Ph.D. project will follow the design science research process [22]. The different research questions will be worked on iteratively to individually create tangible artifacts and assess their impact on the predictive performance of transformers. The development of artifacts starts with the preprocessing, progresses to the model that provides descriptive predictions, and finalizes with the extraction of explanatory insights.

First, the exact challenges of each research question will be identified in detail to infer tangible objectives, which serve as the base for the requirement definition and tracking of goal attainment. These objectives are then synthesized into a system design employing literature research from online databases to extract the latest field findings. During the literature analysis, recent and relevant conference papers will be preferred, followed by other publications and preprints. This literature research will draw special attention to other ML domains, like NLP, to see whether knowledge from these areas can be leveraged in this use case.

The theoretical views on the problem and system design are subsequently used to develop the artifacts and establish a practical understanding of the narrowed problems defined by the research questions and their solutions. The artifacts will then be demonstrated on real-world data to prove their applicability and usefulness to the problem domain.

To conclude the projected approach, the artifacts are tested using evaluation strategies appropriate for the artifacts [23]. The evaluation includes the use of quality measures for ML tasks to ensure that the training and results are valid. In benchmarks against other published and available ML models in predictive and prescriptive process monitoring, it will be verified whether improvements could be reached. The evaluation shall be done on different event logs to ensure the applicability and transferability of the models. Finally, the artifact is evaluated from a functional perspective by comparing the defined aims and requirements with the created artifact to ensure the goals are reached.

5. References

[1] M. Marrone and L. M. Kolbe, “Uncovering ITIL claims: IT executives’ perception on benefits and Business-IT alignment,” *Inf Syst E-Bus Manage*, vol. 9, no. 3, pp. 363–380, Sep. 2011, doi: 10.1007/s10257-010-0131-7.

[2] Yajun Zhang, Jinlong Zhang, and Jiangtao Chen, “Critical Success Factors in IT Service Management Implementation: People, Process, and Technology Perspectives,” in *2013 International Conference on Service Sciences (ICSS)*, Shenzhen, Apr. 2013, pp. 64–68. doi: 10.1109/ICSS.2013.38.

[3] S. Morana, T. Gerards, and A. Mädche, "ITSM ProcessGuide – A Process Guidance System for IT Service Management," in *New Horizons in Design Science: Broadening the Research Agenda*, Cham, 2015, vol. 9073, pp. 406–410. doi: 10.1007/978-3-319-18714-3_32.

[4] J. Serrano, J. Faustino, D. Adriano, R. Pereira, and M. da Silva, "An IT Service Management Literature Review: Challenges, Benefits, Opportunities and Implementation Practices," *Information*, vol. 12, no. 3, p. 111, Mar. 2021, doi: 10.3390/info12030111.

[5] H. Mao, T. Zhang, and Q. Tang, "Research Framework for Determining How Artificial Intelligence Enables Information Technology Service Management for Business Model Resilience," *Sustainability*, vol. 13, no. 20, p. 11496, Oct. 2021, doi: 10.3390/su132011496.

[6] Y. Dang, Q. Lin, and P. Huang, "AIOps: Real-World Challenges and Research Innovations," in *2019 IEEE/ACM 41st International Conference on Software Engineering: Companion Proceedings (ICSE-Companion)*, Montreal, QC, Canada, May 2019, pp. 4–5. doi: 10.1109/ICSE-Companion.2019.00023.

[7] M. Sarnovsky and J. Surma, "Predictive Models for Support of Incident Management Process in IT Service Management," *AEI*, vol. 18, no. 1, pp. 57–62, Mar. 2018, doi: 10.15546/aeei-2018-0009.

[8] P. Anchuri *et al.*, "Graph mining for discovering infrastructure patterns in configuration management databases," *Knowl Inf Syst*, vol. 33, no. 3, pp. 491–522, Dec. 2012, doi: 10.1007/s10115-012-0528-3.

[9] Haochen Li and Zhiqiang Zhan, "Bussiness-driven automatic IT change management based on machine learning," in *2012 IEEE Network Operations and Management Symposium*, Maui, HI, Apr. 2012, pp. 1374–1377. doi: 10.1109/NOMS.2012.6212078.

[10] A. Vaswani *et al.*, "Attention is All you Need," in *Advances in Neural Information Processing Systems 30*, Long Beach, CA, USA, Dec. 2017, vol. 30, pp. 5998–6008.

[11] C. Di Francescomarino, C. Ghidini, F. M. Maggi, and F. Milani, "Predictive Process Monitoring Methods: Which One Suits Me Best?," in *Business Process Management*, Cham, Switzerland, 2018, vol. 11080, pp. 462–479. doi: 10.1007/978-3-319-98648-7_27.

[12] K. Kubrak, F. Milani, A. Nolte, and M. Dumas, "Prescriptive process monitoring: Quo vadis?," *PeerJ Computer Science*, vol. 8, no. 1097, Sep. 2022, doi: 10.7717/peerj-cs.1097.

[13] Z. A. Bukhsh, A. Saeed, and R. M. Dijkman, "ProcessTransformer: Predictive Business Process Monitoring with Transformer Network." arXiv, Apr. 01, 2021. Accessed: Feb. 11, 2022. [Online]. Available: http://arxiv.org/abs/2104.00721

[14] P. Philipp, R. Jacob, S. Robert, and J. Beyerer, "Predictive Analysis of Business Processes Using Neural Networks with Attention Mechanism," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, Fukuoka, Japan, Feb. 2020, pp. 225–230. doi: 10.1109/ICAIIC48513.2020.9065057.

[15] A. Banham, S. J. J. Leemans, M. T. Wynn, and R. Andrews, "xPM: A Framework for Process Mining with Exogenous Data," in *Process Mining Workshops*, Eindhoven, Netherlands, 2022, vol. 433, pp. 85–97. doi: 10.1007/978-3-030-98581-3_7.

[16] A. Banham, S. J. J. Leemans, M. T. Wynn, R. Andrews, K. B. Laupland, and L. Shinners, "xPM: Enhancing exogenous data visibility," *Artificial Intelligence in Medicine*, vol. 133, p. 102409, Nov. 2022, doi: 10.1016/j.artmed.2022.102409.

[17] F. Folino, M. Guarascio, and L. Pontieri, "Discovering High-Level Performance Models for Ticket Resolution Processes," in *On the Move to Meaningful Internet Systems: OTM 2013 Conferences*, Graz, Austria, 2013, vol. 8185, pp. 275–282. doi: 10.1007/978-3-642-41030-7_18.

[18] S. Barbon Junior, P. Ceravolo, E. Damiani, and G. Marques Tavares, "Evaluating Trace Encoding Methods in Process Mining," in *From Data to Models and Back*, Cham, 2021, vol. 12611, pp. 174–189. doi: 10.1007/978-3-030-70650-0_11.

[19] H. Mei, C. Yang, and J. Eisner, "Transformer Embeddings of Irregularly Spaced Events and Their Participants," in *The Tenth International Conference on Learning Representations*, Sep. 2021. Accessed: May 04, 2022. [Online]. Available: https://openreview.net/forum?id=Rty5g9imm7H

[20] R. Galanti, B. Coma-Puig, M. de Leoni, J. Carmona, and N. Navarin, "Explainable Predictive Process Monitoring," in *2020 2nd International Conference on Process Mining (ICPM)*, Padua, Italy, Oct. 2020, pp. 1–8. doi: 10.1109/ICPM49681.2020.00012.

[21] A. Terra, R. Inam, S. Baskaran, P. Batista, I. Burdick, and E. Fersman, "Explainability Methods for Identifying Root-Cause of SLA Violation Prediction in 5G Network," in *2020 IEEE Global Communications Conference*, Taipei, Taiwan, Dec. 2020, pp. 1–7. doi: 10.1109/GLOBECOM42002.2020.9322496.
[22] P. Johannesson and E. Perjons, *An Introduction to Design Science*, 2nd ed. Cham: Springer International Publishing, 2021. doi: 10.1007/978-3-030-78132-3.
[23] J. Venable, J. Pries-Heje, and R. Baskerville, "FEDS: a Framework for Evaluation in Design Science Research," *European Journal of Information Systems*, vol. 25, no. 1, pp. 77–89, Jan. 2016, doi: 10.1057/ejis.2014.36.