Inter-organizational Process Mining through Trusted Execution Environments

Davide Basile[1 [0000−1111−2222−3333]], Luca Barbaro[1 [0000−0002−2975−5330]],
Valerio Goretti[1 [0000−0001−9714−4278]], and
Claudio Di Ciccio[1 [2222−−3333−4444−5555]]

Sapienza University of Rome

Abstract. ...

1 Introduction

In today's interconnected business landscape, organizations are constantly seeking ways to enhance their operational efficiency, increase their performance, and gain valuable insights to improve their processes. In this context, the availability of worthwhile information plays a key role. One of the primary obstacles lies in securely and reliably accessing and utilizing data from various companies or business units, ensuring that all involved parties can derive substantial benefits from it. Traditional approaches to sharing sensitive data across organizational boundaries often involve concerns related to privacy, data integrity, and the need for mutual trust. This paper introduces a novel approach that leverages Trusted Execution Environment (TEE) technology to facilitate the exchange of event logs among different companies or business units. TEE provides a secure and isolated environment within a computer system, ensuring the confidentiality, integrity, and privacy of data and code execution. By utilizing TEE, organizations can exchange event logs in a trusted and privacy-preserving manner, enabling them to harness valuable information from external processes to enhance, monitor, or modify their own processes effectively.

The proposed methodology combines the power of TEE with process mining techniques to unlock the potential of inter-organizational event log exchange. Process mining is a data-driven approach that extracts knowledge and insights from event logs to gain a comprehensive understanding of business processes. By incorporating external event logs from trusted sources, organizations can obtain a broader and more accurate view of the end-to-end processes they are involved in, leading to better process optimization, performance monitoring, and decision-making. By bridging the gap between different organizations, the utilization of TEE-enabled event log exchange for process mining offers promising opportunities for collaborative learning, benchmarking, and continuous process improvement. The findings and insights derived from this research have the potential to revolutionize the way organizations approach process optimization and enable them to make more informed decisions based on a holistic view of their processes.

2 Related Work

The literature proposes several studies that consider process mining techniques in inter-organizational environments. Van Der Aalst [1] shows that inter-organizational processes can be divided according to different dimensions making identifiable challenges of inter-organizational process extractions. Elkoumy et al. [8] propose a tool that allows independent parts of an organization to perform process mining operations by revealing only the result. This tool is called Shareprom and exploits the features of secure multi-party computation (MPC). Engel et al. [11] present EDImine Framework, which allows to apply process mining operations for inter-organizational processes supported by the EDI standard[1] and evaluate their performance using business information. Elkoumy et al. [7] propose an architecture based on MPC. This architecture aims to perform process mining operations without sharing their data or trusting third parties.

Applying process mining techniques in intra-organizational contexts requires merging the event logs of the organizations participating in the process. Hernandez-Resendiz et al. [12] present a methodology for merging logs at the trace and activity level using rules and methods to discover the process. [5] [9] [10] [3] [13]

Once the data exchange has taken place, it is critical that the data be stored in a trusted part of the consumer's device. Basile et al. [4] in their study created a framework called Regov that allows for the exchange of sensitive information in a decentralized web context, ensuring usage control-based data access and usage. To ensure control over the consumer's device, Davide et al. use trusted execution environment that allows storage and utilization management of retrieved resources.

The application of process mining in an inter-organizational scenario is infrequent due to concerns related to privacy and confidentiality, integrity and data heterogeneity. To overcome these problems, a large number of techniques have been proposed. Federated process mining [2] aims to effectively manage the problem of privacy-preserving. Using federated data sources, event data can be transparently mapped between multi-source autonomous provider to monitor, analyze, and improve processes across organizations.

3 Motivating Scenario

In the fast-evolving landscape of healthcare, seamless collaboration between multiple organizations is essential to ensure the highest standard of patient care. We delve into the application of Trusted Execution Environment (TEE) to facilitate the secure exchange of event logs between three pivotal actors: an esteemed hospital, a specialized clinic, and a leading pharmaceutical company. This innovative approach fosters a robust and trustworthy ecosystem where sensitive patient data can be shared securely, promoting seamless collaboration for the betterment of patient outcomes.

[1] https://edicomgroup.com/learning-center/edi/standards

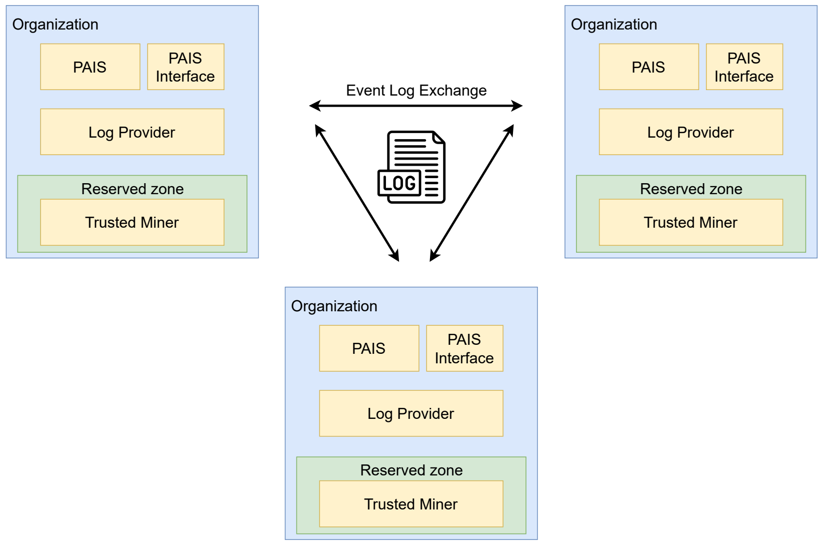

Fig. 1: High-level architectural overview of the framework.

4 Design

In this section, we present the high-level architecture underlying our solution. We take into account each component individually. After that, we focus on the trusted miner component that represent the core of our contribution.

4.1 Architecture at large

Our architecture involves networks of nodes controlled by different `Organizations` exchanging their event logs. `Organizations` in the same network collaborate to reach a common objective composing business processes whose event logs are scattered across multiple places. The hospital, the specialized clinic, and the pharmaceutic company mentioned in the running example, provide an example of partner `Organizations`. An `Organization` may assume one of the following two different roles or both: *provider*, if it delivers event logs to be collaboratively mined; a *miner* whenever it applies process mining algorithms using local event logs in combination with ones generated by providers. Each `Organization` is associated with an asymmetric public/private key pair through which it authenticates messages sent to other `Organizations`.

In Fig. 1, we propose a high level schematization of our solution. Each `Organization` embeds four main components, which we describe next: the Process-Aware Information System (`PAIS`) and the `PAIS Interface`, the `Log Provider` and the `Trusted Miner`.

The `PAIS Interface` collects the logic to interact with the Process-aware Information Sytem (`PAIS`) of the `Organization`. `PAISs` systems help `Organizations`

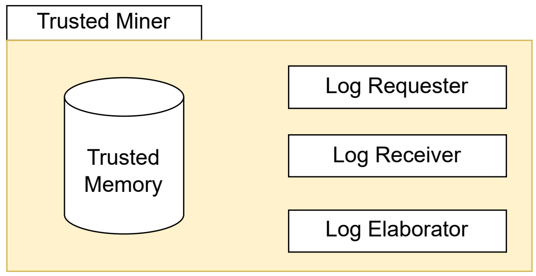

Fig. 2: Modules of the Trusted Miner component

to handle business processes including accounting and resource management. The maintenance of event logs is one of the core tasks performed by these systems [6]. In our architecture, we generalize the interaction with these systems through the `PAIS Interface`. The `PAIS Interface` is queried by the local `Log Provider` for event logs to be fed into `Trusted Miner`s.

The `Log Provider` component delivers on-demand data to `Trusted Miner`s. It controls access to resources by verifying the identity of the miner using asymetric encryption.

The `Trusted Miner` shelters external event logs inside an `Organization`'s system by preserving data confidentiality and integrity. We provide an in deepth focus on this component as follow.

4.2 Trusted Miners in Reserved Zones

The main goal of the `Trusted Miner` is to allow `Organization`s to execute process mining algorithms using event logs retrieved from partner `Organization`s, offering fair data utilization to log providers. `Trusted Miners` are located in `Reserved Zone`s of the `Organization` system. The `Reserved Zone` isolates the`Trusted Miner` and guarantees tamperproofing and data confidentiality. In Fig. 2, we show an high level schematization of `Trusted Miner`s in which we distinguish four different modules: the `Trusted Memory`, the `Log Requester`, the `Log Receiver`, and the `Log Elaborator`.

Data contained in event logs belonging to provider `Organizaition`s are stored by miners in the `Trusted Memory`. This module relies on an hardware tecniques which makes event log manipuluations impossible from outside the `Reserved Zone`. Modules of the `Trusted Miner` are the only entities enabled to access data stored in the `Trusted Memory`.

The `Log Requester` and the `Log Receiver` are the fundamental modules that we employee during the event log exchange. `Log Requester`s initalize the exchange procedure and sends authenticable data requests to the `Data Provision` module of log providers. The `Log Receiver` collects event logs sent by `Log Providers` and store them in the `Trusted Memory`. This module offers methodologies that allow `Log Providers` to perform remote attestation through wich it verifies that the `Log Requester` is trusted and is actualy running in the `Reserved Zone`.

The `Log Elaborator` is the core module of the `Trusted Miner`. It collects the logic to execute process mining algorithms in the `Reserved Zone`. It support the integration of *process discovery* [?], *conformance checking* [?] and *performance analysis* [?] process mining tecniques. When activated, the `Log Elaborator` accesses external event logs stored inside the `Reserved Zone` and integrates them with the local event log via the merging procedure. During the merging, the `Log Elaborator` enriches local traces with events belonging to logs retrieved from partner `Organization`s.

5 Realization

In this section we outline the technical aspects concerning the realization of our framework. Therefore we first present technologies and protocols through which we deploy the design principles presented in Section 4. After that, we discuss the interaction workflow between the instantiated components. Finally we present implementation details.

5.1 Deployment

5.2 Workflow

5.3 Implementation

6 Evaluation

6.1 Validation: Volvo

6.2 Convergence Study: Hospital

Settings

Results

7 Conclusion and Future Work

Limitations:

- Both producer and consumer act fairly (so we do not expect to have injected data)
- We do not manage TEE crashes
- We assume a perfect communication channel (no loss, no snap, no corrupted bits)
- Universal clock for event timestamps (cite Event log cleaning for business process analytics by Andreas Solti)

Future Work:

– Declarative models adaptation
– Output inside the TEE, interactions through trusted applications
– Real-world event log data
– Usage policies integration

References

1. van der Aalst, W.M.: Intra-and inter-organizational process mining: Discovering processes within and between organizations. In: The Practice of Enterprise Modeling: 4th IFIP WG 8.1 Working Conference, PoEM 2011 Oslo, Norway, November 2-3, 2011 Proceedings 4. pp. 1–11. Springer (2011)
2. van der Aalst, W.M.: Federated process mining: exploiting event data across organizational boundaries. In: 2021 IEEE International Conference on Smart Data Services (SMDS). pp. 1–7. IEEE (2021)
3. Arfaoui, G., Gharout, S., Traoré, J.: Trusted execution environments: A look under the hood. In: 2014 2nd IEEE international conference on mobile cloud computing, services, and Engineering. pp. 259–266. IEEE (2014)
4. Basile, D., Ciccio, C.D., Goretti, V., Kirrane, S.: Blockchain based resource governance for decentralized web environments. Frontiers in Blockchain **6** (may 2023). https://doi.org/10.3389/fbloc.2023.1141909, https://doi.org/10.3389%2Ffbloc.2023.1141909
5. Claes, J., Poels, G.: Merging event logs for process mining: A rule based merging method and rule suggestion algorithm. Expert Systems with Applications **41**(16), 7291–7306 (2014)
6. Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A.: Fundamentals of Business Process Management, Second Edition. Springer (2018). https://doi.org/10.1007/978-3-662-56509-4
7. Elkoumy, G., Fahrenkrog-Petersen, S.A., Dumas, M., Laud, P., Pankova, A., Weidlich, M.: Secure multi-party computation for inter-organizational process mining. In: Enterprise, Business-Process and Information Systems Modeling: 21st International Conference, BPMDS 2020, 25th International Conference, EMMSAD 2020, Held at CAiSE 2020, Grenoble, France, June 8–9, 2020, Proceedings 21. pp. 166–181. Springer (2020)
8. Elkoumy, G., Fahrenkrog-Petersen, S.A., Dumas, M., Laud, P., Pankova, A., Weidlich, M.: Shareprom: A tool for privacy-preserving inter-organizational process mining. BPM (PhD/Demos) **2673**, 72–76 (2020)
9. Engel, R., Bose, R.J.C.: A case study on analyzing inter-organizational business processes from edi messages using physical activity mining. In: 2014 47th Hawaii International Conference on System Sciences. pp. 3858–3867. IEEE (2014)
10. Engel, R., Krathu, W., Zapletal, M., Pichler, C., van der Aalst, W.M., Werthner, H.: Process mining for electronic data interchange. In: E-Commerce and Web Technologies: 12th International Conference, EC-Web 2011, Toulouse, France, August 30-September 1, 2011. Proceedings 12. pp. 77–88. Springer (2011)
11. Engel, R., Krathu, W., Zapletal, M., Pichler, C., Bose, R.J.C., van der Aalst, W., Werthner, H., Huemer, C.: Analyzing inter-organizational business processes: process mining and business performance analysis using electronic data interchange messages. Information Systems and e-Business Management **14**, 577–612 (2016)

12. Hernandez-Resendiz, J.D., Tello-Leal, E., Marin-Castro, H.M., Ramirez-Alcocer, U.M., Mata-Torres, J.A.: Merging event logs for inter-organizational process mining. In: New Perspectives on Enterprise Decision-Making Applying Artificial Intelligence Techniques, pp. 3–26. Springer (2021)
13. Xie, H., Zheng, J., He, T., Wei, S., Hu, C.: Tebds: A trusted execution environment-and-blockchain-supported iot data sharing system. Future Generation Computer Systems **140**, 321–330 (2023). https://doi.org/https://doi.org/10.1016/j.future.2022.10.016, https://www.sciencedirect.com/science/article/pii/S0167739X22003326