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. [7] 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. [9] 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. [6] 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. [10] present a methodology for merging logs at the trace and activity level using rules and methods to discover the process. Claes et al. [4] provide techniques for performing merge operations in inter-organizational environments. This paper indicates rules for merging data in order to perform process mining algorithms.

EDI (electronic data interchange) standards pemit business communication of documents. Among these standards, the notion of process is not explicitly specified, impeding organizations from applying Business Process Management (BPM) methods in business collaboration environments. Engel et al.[8] extended process mining techniques by discovering interaction sequences between business partners based on exchanged documents. Lo et al.[13] have provided and developed a framework for data exchange designed even in intra-organizational situations. This framework is based on blockchain and decentralized public key infrastructure technologies. These technologies ensure scalability, reliability data security, and data privacy.

Xie et al.[14] propose an architecture for the internet of things based on Trusted Execution Environment and blockchain. The combination of these technologies enables data communication and use in an off-chain environment. The proposed architecture aims to solve data and identity security problems in the process of data sharing.

Hussain et al.[11] present a tool for privacy protection and data management in organizations. the tool is designed for the management, protection, sharing and re-sharing of data among multiple collaborating companies. This tool allows data encryption to be configured according to the privacy obligations dictated by the context of a system's use.

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. [3] in their study created a framework called Regov that allows for the exchange of sensitive

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

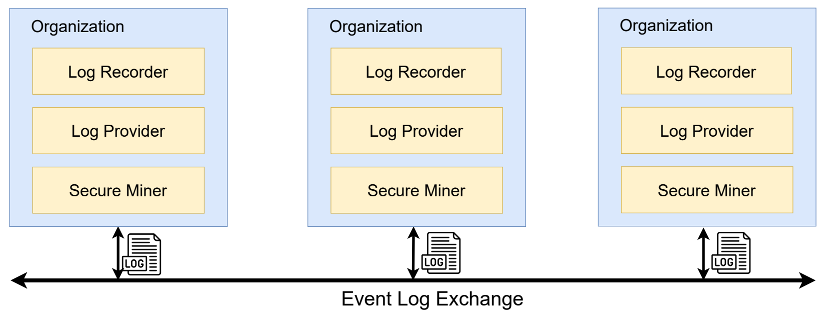

Fig. 1: High-level architectural overview of the framework.

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.

4 Design

In this section, we present the high-level architecture underlying our solution. We take into account the main functionalities of each component avoiding details on the employed technologies discussed in the next sections. Once introduced the architecture, we focus on the `Secure Miner` component that represent the core of our contribution.

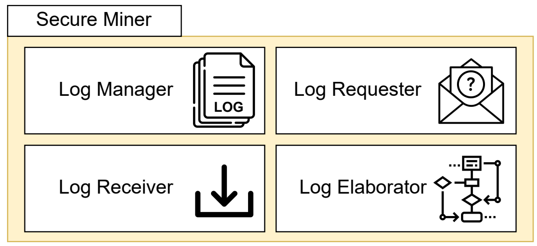

Fig. 2: Modules of the Secure Miner component

4.1 Architecture at large

Our architecture involves networks of nodes controlled by different `Organization`s exchanging their event logs. `Organization`s in the same network collaborate to achieve a common objective and compose business processes whose event logs are scattered across multiple places. Therefore, each `Organization` produces event logs recording the operations executed to complete a business process. The hospital, the specialized clinic, and the pharmaceutical company mentioned in the running example provide an example of partner `Organization`s. An `Organization` may assume one of the following two different roles or both: *provider*, if it delivers local 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.

In Fig. 1, we propose a high level schematization of our solution. Each `Organization` embeds four main components, which we describe next: the `Log Recorder`, the `Log Provider` and the `Secure Miner`.

The maintenance of event logs is the core task performed by `Log Recorder`. This component registers the events taking place in the `Organization`. The `Log Recorder` is queried by the local `Log Provider` for event logs to be fed into `Secure Miner`s.

The `Log Provider` component delivers on-demand data to `Secure Miner`s. It controls access to owned event logs by authenticating data requests generated by miners. `Log Provider`s reject demands from unauthorizated parties and only permit `Secure Miners` of partner `Organizations` to use the data.

The `Secure 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 Secure Miners

The primary objective of the `Secure Miner` is to allow `Organization`s to execute process mining algorithms using event logs retrieved from partner `Organization`s, ensuring fair data utilization to log providers. `Secure Miners` leverage isolated execution contexts that guarantee tamperproofing and data confidentiality. In Fig. 2, we show an high level schematization of `Secure Miner`s in which we distinguish four different modules: the `Log Manager`, the `Log Requester`, the `Log Receiver`, and the `Log Elaborator`

Event logs belonging to partner `Organizaition`s are stored in the isolated execution context of the `Secure Miner`. We handle these data via the `Log Manager` that makes event log access not practicable from outside the `Secure Miner`'s execution context. Thus, the `Log Manager` prevents external parties from having direct access to event logs. These unauthorized entities include the owner of the miner `Organization` system.

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 entrust them to the `Log Manager`. When collecting data, `Log Receiver`s prove their trustworthiness to `Log Provider`s delivering evidences that certifies the `Secure Miner`'s execution context.

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

5 Realization

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

5.1 Deployment

As follow, we bridge the gap between high-level system architecture and its practical realization. Fig. 3 depicts a *UML deployment diagram* [12] that aims to help with understanding the instantiated infrastrucuture.

The `Organization Machine` represent the physical computation *device* embracing the software and hardware entities of the company. The `Log Recorder`, the `Log Provider` and `Secure Miner` are included in the `Organization Machine` as abstract *components* . These logical elements incorporate the core functionalities already discussed in Section 4. The `Organization Machine` is characterized by two *execution environment*s namely the `Operative System` and the `Trusted Execution Environment`.

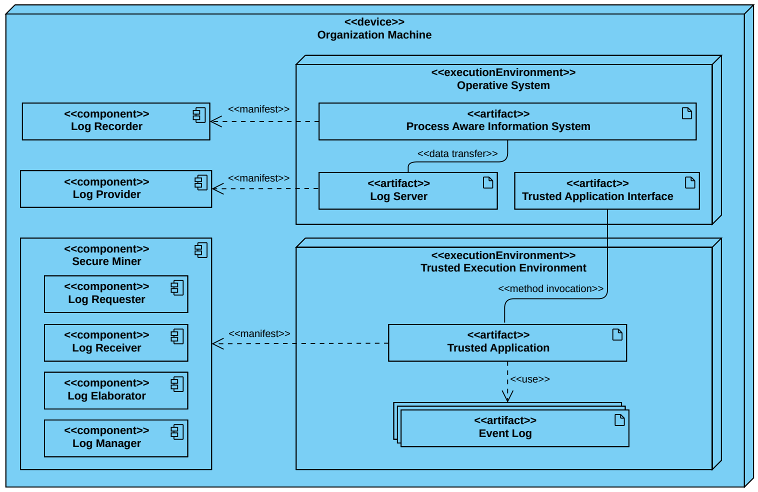

Fig. 3: UML deployment diagram.

Software entities that we expose to the users of the `Organization Machine` run inside the `Operative System`. We mainfest the functionalities offered by the `Log Recorder` in the `Process Aware Information System` [5]. These systems help users to handle business processes including accounting and resource management. In our solution, the `Process-Aware Information System` provides the `Log Server` access to event logs. `Log Servers` are web services which processes remote data request and provides event log to miners. We build this entities upon existing web standards such as HTTP, FTP and Goopher.

`Trusted Execution Environment`s are the core technologies of our solution. Application and data located in the Trusted Execution Environment are protected through hardware-encryption tecniques in a reserved zone of the `Organization Machine`'s CPU.

5.2 Workflow

5.3 Implementation

6 Evaluation

6.1 Convergence study

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. 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
4. 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)
5. 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
6. 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)
7. 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)

8. 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)
9. 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)
10. 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)
11. Hussain, A., Lasrado, L.A., Mukkamala, R.R., Tanveer, U.: Sharing is caring–design and demonstration of a data privacy tool for interorganizational transfer of data. Procedia Computer Science **181**, 394–402 (2021)
12. Koch, N., Kraus, A.: The expressive power of uml-based web engineering. In: Second International Workshop on Web-oriented Software Technology (IWWOST02). vol. 16, pp. 40–41. Citeseer (2002)
13. Lo, N.W., Chen, S.C., Chang, S.C.: A flexible electronic data exchange framework based on consortium blockchain. Journal of Internet Technology **21**(5), 1313–1324 (2020)
14. 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