THE DPIA OF AN ENTERPRISE CONTACT TRACING SOLUTION: LESSONS LEARNED AT THE CROSSROADS OF CYBERSECURITY AND DATA PROTECTION

Majid Mollaeefar [a], **Roberto Carbone** [a], **and Silvio Ranise** [a, b]
[a] FBK-Center for Cybersecurity, Trento, Italy.
[b] Department of Mathematics, University of Trento, Trento, Italy.

{mmollaeefar, carbone, ranise}@fbk.eu

ABSTRACT

In the wake of the COVID-19 pandemic, Enterprise Contact Tracing (ECT) systems have become integral to public health management within organizations, presenting challenges in cybersecurity and data protection compliance, particularly under the General Data Protection Regulation (GDPR). This paper explores the landscape of ECT systems, focusing on key requirements and their interplay with security, privacy, and compliance challenges. We emphasize the necessity of integrating security and privacy by design principles into ECT systems to ensure compliance with GDPR principles and to effectively safeguard sensitive data. Central to our work is the execution of a Data Protection Impact Assessment (DPIA) for an ECT solution called `Trace4Safe`, encompassing a systematic evaluation of potential security and privacy threats. To address these challenges, we employ a methodology based on the Multi-Stakeholder Risk Minimization Problem (MSRMP), aimed at identifying optimal solutions that consider the varied perspectives and objectives of all stakeholders, including data controllers and subjects. We present experimental results from applying the MSRMP methodology to the Trace4Safe solution and explore different strategies to select optimal solutions for risk management. Additionally, we discuss the implications of various threats on protection goals and the importance of selecting a solution that balances risks for all stakeholders. The paper concludes with an analysis of specific mitigation mappings associated with the chosen solution, highlighting the trade-offs and decision-making processes in achieving an effective balance between organizational safety requirements and individual privacy rights.

Keywords Cybersecurity · Data Protection Impact Assessment · COVID-19 · Enterprise Contact Tracing · GDP · Multi-Stakeholder Risk Assessment

1 Introduction

Contagious diseases are typically transmitted from person to person through direct touch with bodily fluids, inhalation of infected individuals' respiratory droplets, or contact with contaminated surfaces and materials. To identify patients and those at higher risk and physically isolate them from the general population during outbreaks of these diseases, it is critical to trace human-to-human interactions in the past [1]. This process is known as contact tracing [2]. Since at least 2007, digital contact tracing has been known and shown to be effective in the first empirical study employing Bluetooth data in 2014 [3, 4]. However, mass adoption has been a significant barrier. The idea gained notoriety during the COVID-19 pandemic, attracting the interest of several countries, such as China, Italy, Singapore, Germany, and big IT enterprises like Microsoft, Google, and Apple.

In the modern era's public health landscape, contact tracing plays a vital role, functioning as a core strategy to mitigate the spread of infectious diseases. Enterprise-level contact tracing solutions present themselves as a robust complement to national-scale initiatives, offering innovative and tailored approaches to disease prevention and control. Notably, enterprise-focused solutions are not only adept at dovetailing with broader, national contact tracing efforts, but they also

serve as valuable tools in adhering to Environmental, Social, and Governance (ESG) frameworks. By aligning safety protocols with sustainable business practices, they create safer workplaces, promote social responsibility, and ensure better governance in managing health crises. This paper focuses on an enterprise-level, exemplified by the innovative Trace4Safe[1] solution, offering proactive, targeted, and customizable mechanisms that distinctly set it apart from national-scale counterparts. Specifically designed for business entities, enterprise contact tracing (ECT) systems like Trace4Safe offer real-time alerts and preventive strategies that are not typically feasible at a national scale, considering the intricate, massive details and processes involved. Moreover, they may include gamification features to encourage adherence to safety protocols, fostering a culture of compliance and shared responsibility within the organization. In essence, the distinctiveness of ECT lies in its scale, proactive nature, real-time alerting capabilities, and the ability to integrate gamification, offering a more customized approach to ensure the safety of employees within the organizational environment, while continuing to play a crucial role in broader public health management.

Contact tracing solutions can use multiple technologies such as short-range wireless technologies like Wi-Fi and Bluetooth, or GPS to trace contacts. Regardless of what technology is used, organizations face security, privacy, and compliance challenges.

-Security. Security is a prime concern in any system that deals with sensitive data. These systems collect and process personal and location data, making them potential targets for malicious attacks aimed at compromising the confidentiality, integrity, or availability of the data—elements collectively referred to as the CIA triad. Attacks could range from unauthorized access to the data, their alteration or destruction, causing severe implications such as privacy invasion, data manipulation, or misinformation. As a consequence, deploying robust security mechanisms, including encryption, authentication, secure storage, and access control, becomes imperative to protect the data and maintain the system's trustworthiness.

-Privacy. In the context of contact tracing, privacy extends beyond merely securing the data; it involves adhering to stringent data protection norms. It encompasses the respect of individuals' rights and freedoms concerning the processing of their personal data. According to the General Data Protection Regulation (GDPR) [5], principles like purpose limitation, transparency, and data minimization must be observed. Privacy challenges can escalate as contact tracing inherently collects sensitive personal information, often containing personally identifiable information (PII). Therefore, it is essential for such systems to be designed to mitigate the risks of privacy invasion while adhering to the principles of GDPR. Moreover, privacy significantly overlaps with security, specifically in relation to the CIA triad. Confidentiality, which aims to prevent unauthorized disclosure, that when compromised in contact tracing, could inadvertently expose personal health data, posing privacy risks and potentially leading to discrimination. Integrity also holds relevance to privacy because it significantly contributes to data accuracy; In the context of contact tracing, if violated, it could cause panic and misinformed health responses due to incorrect contact or health status alerts. Similarly, availability, ensuring data accessibility when necessary, is intertwined with privacy. In the event of contact tracing data becoming unavailable due to system failures or cyberattacks, it could result in delayed exposure alerts, infringing on individuals' right to timely healthcare. Hence, the overlap of privacy and security challenges in contact tracing necessitates holistic solutions that adequately address both aspects.

-Compliance. Compliance, according to GDPR, means the system must meet the requirements for properly handling personal data in a given system that processes personal information, and infringements of *data processing principles and data subjects' rights* are subject to administrative fines[2].

There exist several reasons to enforce privacy in the context of proximity contact tracing applications where sensitive data are collected from users. Therefore, it is crucial to follow Security by Design (SbD) and Privacy by Design (PbD) principles, which are increasingly identified as necessary for dealing with design faults that may jeopardize security or privacy in the system [6]. Indeed, these approaches (i.e., SbD and PbD) help to address the *security* and *privacy* challenges mentioned above. Furthermore, their application is also envisaged by the GDPR, as it demands Data Protection by Design (DPbD) for any systems that involve personal data in their processing. To meet the *compliance* challenge, the GDPR requires conducting a Data Protection Impact Assessment (DPIA) to reduce risks and preserve individuals' rights and freedom.

In addition to the challenges outlined above, enterprises may have to contend with further complexities when considering the implementation of contact tracing solutions within their organizations. These complications could span from balancing business requirements to satisfying the ESG framework while operating within budget constraints. Addressing these challenges requires a comprehensive approach that factors in the diverse needs of all stakeholders, from the enterprise to its employees and the public at large. Thus, organizations acting as data controllers must adopt a strategy,

[1]The "Trace4Safe" project represents a solution designed to monitor and control social distancing during the COVID-19 pandemic through the use of contact tracing systems.

[2]According to article 83.5.a and b, the fines can be up to 20,000,000 EUR, or in the case of an undertaking, up to 4% of the total worldwide annual turnover of the preceding financial year, whichever is higher.

procedure, or framework that helps them make informed decisions about the deployment of security and privacy mechanisms. The goal is to strike a balance, facilitating a more favorable trade-off between their necessities and the rights of the data subjects. The task of providing a more favorable trade-off between organizations' needs and those of their users is non-trivial, as it may require searching through all possible subsets of a set of available controls and identifying those that minimize the risks of all stakeholders. Since stakeholders may have different perceptions of the risks (especially when considering the impact of threats), conflicting goals may arise that require finding the best possible tradeoffs among the various needs [7]. To do this, we use the methodology in [7] that solves this problem based on the notion of Pareto optimality to formulate the trade-off identified above. However, the designer may be left with the problem of selecting one particular solution among the many Pareto optimal ones. This is a tricky task that requires some ingenuity. We make two main contributions in this paper: first, we explain how the methodology developed in [7] can be integrated with performing a DPIA in a real-world enterprise contact tracing solution, and then, to help designers select the optimal solution, we propose an approach to do this by exploiting some insights made available by the methodology.

Plan of the paper. In Section 2, we discuss enterprise contact tracing solutions, the objectives, and common processes within these systems. Following that, we introduce `Trace4Safe` as an example of an ECT solution. The requirements and their relationship with security, privacy, and compliance within these kinds of systems will be discussed (in Section 2.1). In Section 3, we conduct a data protection impact assessment. To do that, we briefly overview the methodology proposed in [7] by applying it to the Trace4Safe scenario. Section 4 presents the experimental results of solving the MSRMP instance of Trace4Safe, and we propose some strategies (in Section 4.1) in order to select the optimal solution. The related work is discussed in Section 5. We conclude the paper with a summary of the main contributions and some hints for future work (Section 6).

2 Enterprise Contact Tracing Solutions

The sanitary and socio-economic crisis produced by the COVID-19 outbreak has significantly influenced many different aspects of our lives, changing both personal and business interests. Workplaces, in particular (industrial sites, offices, stores, and so on), were required to ensure that a one-meter physical separation and rigorous adherence to hygiene requirements were adhered to in order to protect the health of employees and clients alike. Production lines at manufacturing enterprises were among the first activities permitted to recommence operations. Since the beginning of the lockdown, they have been actively exploring new solutions to put in place in order to be prepared for a safe reopening. The aim of these solutions is to support businesses and best manage the situation determined by the COVID-19 pandemic. In particular, they target the following objectives:

1. **To monitor and enforce physical distancing**: this service should focus on the education towards proper and safe behavior at work, based on the current physical distancing regulations. The system will then provide immediate feedback to the involved parties when such distancing is not respected. The feedback could come directly from the Token (e.g., buzzer, vibration, or visual feedback) or from some dedicated device (e.g., light or screen).
2. **To analyze the risks associated with a COVID-19 outbreak**: analyze the risks related to a possible COVID-19 outbreak under the current working environment and business processes. This will be based on the analysis of the current contact network and will identify critical aspects to be addressed and reviewed.
3. **To best manage a possible COVID-19 outbreak**: preserve business productivity, confining the outbreak to the involved parties only and to the production areas affected. The system will then allow businesses to immediately react to the outbreak, identifying all parties involved and supporting them in managing the case.

An ECT solution is primarily based on a wearable device that workers carry during working hours. The wearable device monitors social distancing and traces contacts. If a worker tests positive for a contagious disease, the wearable device would have a record of other devices that have been in close proximity, enabling quick and efficient contact tracing. Hence, individuals who might have been exposed to the disease can be rapidly identified, allowing for a swift response to minimize further spread. The data exchange in such contact tracing systems undergoes four distinct phases, each contributing vital input to the risk assessment process. The phases are as below:

-**Registration**. During this phase, the worker (employee or individual who will be wearing the wearable device) must provide some basic information (such as his or her name, surname, email address, and role in the organization) to the COVID safety person who is responsible for registering users in the system and assigning them a token.

-**During Working Shifts**. The system may collect data from the worker's wearable device (token) in two ways; the Real Time Locating System (RTLS) and the Point to Point (P2P) approaches. In dense environments, with employees working in close proximity, RTLS provides the most accurate and reliable approach to monitor social distancing and

Table 1: Contact tracing system requirements categorized by relevance to **S**ecurity, **P**rivacy, and **C**ompliance challenges.

Type	Requirements	S	P	C
General	GR_1. Ensuring security by design	5	-	-
	GR_2. Ensuring privacy by design	-	5	-
	GR_3. Implementing data minimization	5	5	5
	GR_4. Ensuring data confidentiality, integrity, and availability	5	-	5
	GR_5. Preventing abuse of collected information	-	5	5
	GR_6. Protecting health information from disclosure	5	5	5
	GR_7. Achieving user trust and transparency	-	5	5
	GR_8. Handling user consent and opt-out	-	5	5
	GR_9. Guaranteeing purpose limitation	-	5	5
	GR_{10}. Avoiding user tracking to maintain anonymity	-	5	5
	GR_{11}. Preventing exposure of user's location	-	5	-
Enterprise	ER_1. Implementing robust anonymization techniques	5	5	-
	ER_2. Mitigating potential for de-anonymization	5	5	-
	ER_3. Managing stigmatization and discrimination risks	-	5	5
	ER_4. Mitigating surveillance perception	5	5	-
	ER_5. Aligning with ESG and business requirements	-	-	5

trace contacts. Conversely, in sparse environments where workers are isolated most of the time, a P2P approach should be preferred due to its limited costs and complexity.

- In the RTLS approach, the token collects data on the worker's location for measuring physical distancing.
- In the P2P approach, tokens measure the physical distance between workers. When two tokens are closer than the defined distance, they record this as contact. This data is stored in the token's memory until it is uploaded to a gateway, which then forwards it to an edge server or the cloud for later processing.

-**Reporting by Users**. If a worker becomes sick, exhibits symptoms, or tests positive for a sickness, they can report this to the team leader through various communication methods. The system uses this information to create a network of contacts and inform the COVID safety responsible person, who then prepares a plan for isolating the sick person, limiting their contact with others, and contacting local health authorities.

-**Sending Report by the System**. If a worker breaches safe work guidelines, the system can provide information about this to the team leader or directly to the worker. This feedback helps to reinforce safety protocols and can be used as part of a gamification approach to encourage compliance.

Trace4Safe is an instance of an ECT solution created in the context of a project funded by the European Institute of Innovation and Technology (EIT)[3] to contribute to COVID-19 pandemic management.

2.1 Security, Privacy and GDPR Compliance in ECT

Contact tracing solutions often necessitate the use of a *Token* to monitor individuals' interactions over time. As such, a comprehensive understanding of privacy implications and system compliance with current regulations is crucial. According to Articles 35 and 36 of the GDPR, a Data Protection Impact Assessment (DPIA) must be undertaken if processing personal data presents a substantial risk to the freedom and rights of natural persons. The European Data Protection Board (EDPB) mandates that a DPIA should precede the deployment of a contact tracing system given the associated high risk, which encompasses health data, expected large-scale adoption, systematic monitoring, and the utilization of a new technological solution [8]. It is imperative that contact tracing systems, embody several GDPR-compliant privacy and security measures from the outset. In the context of ECT solutions, contact data such as Token IDs, timestamps, and, in the case of the RTLS approach, location data, is considered personally identifiable information (PII). Any disclosure of such data could potentially lead to the identification of individuals, thus exacerbating the risks.

Table 1 listed a set of requirements in two categories; which are general in any contact tracing system and specific for ETC solutions. Such systems must consider security (GR_1) and privacy (GR_2) by design, forming a strong foundational layer of protection. This approach to design enables effective data minimization (GR_3), ensures data confidentiality, integrity, and availability (GR_4), helps safeguard against the abuse of collected information (GR_5), and shields health information from unauthorized disclosure (GR_6). Furthermore, contact tracing systems must establish user trust

[3] `https://eit.europa.eu/`

and maintain transparency (GR_7), handle user consent and opt-out procedures diligently (GR_8), and ensure that the purposes for which the data are used are clearly defined and strictly limited (GR_9). Maintaining user anonymity is crucial, thus tracking of users must be avoided (GR_{10}), and steps must be taken to prevent the exposure of users' location (GR_{11}).

ECT solutions operate within a distinctive environment, presenting unique challenges and necessitating the consideration of specific requirements. In particular, these systems must implement robust anonymization techniques (ER_1) to effectively protect users' privacy. This is especially critical in ECT solutions where the pool of individuals is relatively smaller compared to national-level contact tracing applications, and the patterns of interactions are often recurrent and predictable. Even when explicit identifiers are removed, the unique characteristics of these interaction patterns could potentially lead to the indirect identification of individuals (ER_2). A notable example might be the revelation of an employee's identity through their repeated close proximity to a specific workstation during certain hours. As such, it is vital to develop and implement robust anonymization techniques tailored to the unique characteristics of workplace environments. Addressing privacy concerns is not limited to the implementation of robust anonymization. ECT systems must also ensure that they do not contribute to tracking users, revealing users' interactions, exposing visited locations, or identifying COVID-19 positive cases [9], all of which can violate the principles of privacy by design. Of particular concern is the perception of the contact tracing system as an intrusive surveillance tool. Continual monitoring of employees' movements and interactions can lead to discomfort among employees, thereby necessitating the need to mitigate such perceptions of surveillance (ER_4). ECT solutions must also manage the risks of stigmatization and discrimination (ER_3). Identifying individuals as infected or at risk can potentially lead to negative outcomes, both within and outside the workplace. These risks could deter individuals from participating in the system or from being transparent about their health status. Thus, it is crucial that ECT solutions are designed with mechanisms to prevent such occurrences. Moreover, it is of utmost importance that ECT solutions align with ESG and other business requirements (ER_5). Such alignment not only ensures compliance with standards and regulations but also bolsters the acceptability and adoption of the solution in an enterprise setting.

In light of these requirements, ECT solutions must carefully consider the unique characteristics and requirements of the workplace environment in their development and implementation. Failing to consider any of these critical aspects could lead to privacy breaches, reduced system adoption, and decreased effectiveness of contact tracing efforts in controlling the spread of infectious diseases like COVID-19.

3 DPIA within Trace4Safe

The previous section provided an investigation into the security, privacy, and compliance challenges inherent in contact tracing systems, focusing specifically on the ETC solution, this section delves into the process and methodology of a DPIA within the Trace4Safe solution. In light of these challenges, the necessity for a DPIA becomes paramount. Conducting a thorough investigation into the potential security and privacy threats in the Trace4Safe scenario forms the crux of this assessment. Our approach to the DPIA begins with the identification of possible vulnerabilities within the system, leading to the construction of threat scenarios. Alongside each of these, we identify potential mitigating strategies designed to reduce the risk of these threats. Subsequent to this, we perform an evaluation of the security and privacy risks associated with the identified threats within the Trace4Safe contact tracing scenario.

In preparation for conducting this assessment, we sought to gain a comprehensive understanding of the Trace4Safe solution's architecture. Figure 1 provides an overview of the scenario in which the solution operates, illustrating the primary actors and their interactions with the system components. These actors include the COVID safety responsible person, representing the service provider (i.e., the organization utilizing the Trace4Safe product), the end-users (i.e., employees), and the team leader, who has authorized access to the contact tracing data. Each communication channel, where data collection and transfer occur, has been numbered for ease of reference.

Simultaneously, we developed a questionnaire and circulated it among project partners, where it served to elucidate privacy issues at each phase (the four phases of data exchange identified in Section 2) of the scenario. The questions were mapped to a privacy principle (e.g., Data quality) and a privacy target (e.g., ensuring data minimization), derived from both GDPR legal principles and existing privacy-related literature [10, 11]. These principles and targets align with data protection goals introduced in [12] and can aid in identifying mitigation controls.

In this section, we go one step further by analyzing security and privacy risks associated with the Trace4Safe solution by adopting the methodology developed in [7]. To this end, in Section 3.1, we briefly overview and describe the main components of the methodology.

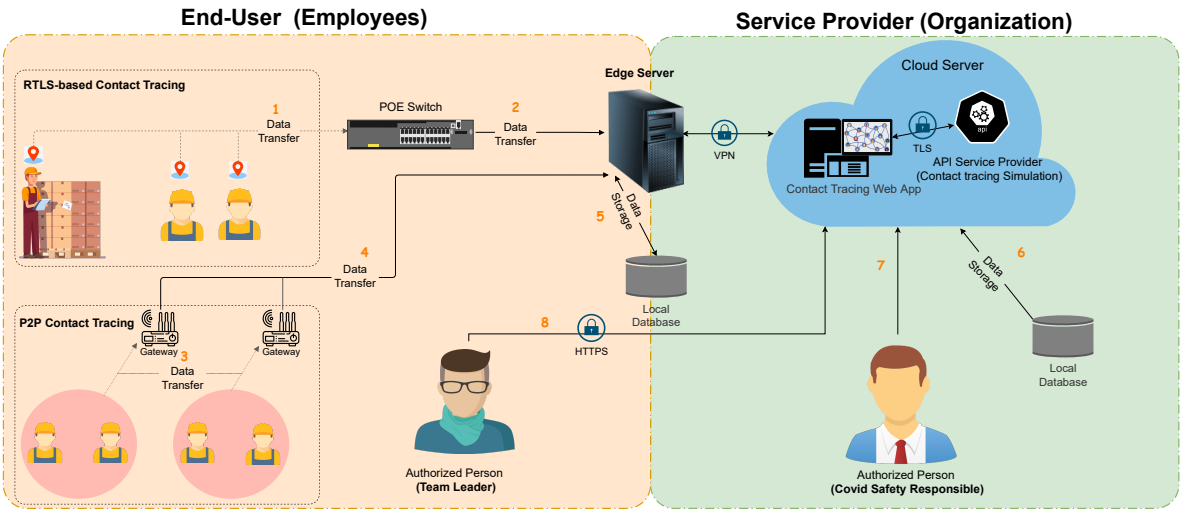

Figure 1: System's components and the main actors in Trace4Safe.

3.1 The Methodology

The Multi-Stakeholder Risk Minimization Problem (MSRMP) introduced in [7] to assist in the definition of the best (with respect to all the stakeholders involved in the system) Risk Management Policies (RMPs)—as an appropriate set of security controls to mitigate the identified set of threats—in the fundamental step of selecting mitigations for risk management. As observed in [7], it is useful in a DPIA to be able to evaluate different configurations of mitigations that allow for to reduction of the risks related to data protection as perceived by the set $\mathcal{S}$, of stakeholders involved in the system including the data controller and data subject. For this, it is crucial to identify the set $\mathcal{G}$ of protection goals, the set $\mathcal{T}$ of threats, and the set $\mathcal{C}$ of controls mitigating the threats in $\mathcal{T}$. As explained in [7], this allows us to specify the following multi-objective optimization problem (MOOP):

$$min_{h\mu_T i_{T2T}} \ \langle \ _{T2T} \ i_s(T) \cdot (1 - \frac{_{c2C_T}\mu_T(c)}{jC_Tj}) \rangle_{s2S} \tag{1}$$

where, μ_T represents a mitigation mapping, $\mathcal{C}_T$ is a family of controls associated to $T \in \mathcal{T}$, and

$$i_s(T) = \frac{1}{|il_{max}|} \sum_{p2P} al_p^s(T) \times PW_p^s \tag{2}$$

The solutions of (1) correspond to Pareto optimal configurations of mitigations, striking a trade-off between various stakeholders. However, we are left with the problem of defining $i_s(T)$, which represents the impact level threat T for each stakeholder $s \in \mathcal{S}$. In this formula, al_p^s represents the aversion level for each protection criterion p and stakeholder s, PW_p^s is the *weight* assigned to each criterion, and il_{max} donates to the maximum impact level. In the following, we explain how this defines in our ECT scenario, but before that, we briefly explain how the input requirements are going to obtain:

-A list of Threats ($\mathcal{T}$). We identified security and privacy threats to the Trace4Safe solution and their consequences (See Table 8 in Appendix A). We also mapped each of these threats to the corresponding affected components/channels reported in Figure 1. For security threats, we adopted Microsoft STRIDE threats [13], and we used LINDDUN [14] for privacy threats. Although Table 8 listed fifteen security and privacy threats, for this assessment, we assumed that the local databases on the edge server and cloud keep data securely and hence neglected the threats associated with database breaches caused by unauthorized access. Also, some privacy threats, such as purpose limitation and transparency, do not affect personal data. In Table 2, five threats (out of the fifteen threats considered in Table 8) are chosen for the risk assessment because they have a more adverse impact on either the contact tracing network (e.g., manipulating the contact network) or have some consequences on users, like tracking and identification.

-A list of Controls ($\mathcal{C}$). We identified mitigation controls for each threat and reported them in the second column of Table 2. For instance, *Regularly checking Token's battery* is considered as a security mitigation control for the *Battery*

drain attacks (i.e., T3), whereas *Facilitating the report by workers to the system* is considered as a privacy mitigation control for the *intervenability threat* (i.e., T5) by exploiting the "privacy target" concept defined in [10].

-A list of Stakeholders ($\mathcal{S}$). As can be seen in Figure 1, we have two main stakeholders: (i) the organization as the service provider who plans to deploy the Trace4Safe solution in its organization, and (ii) the end-users, who are the organization's employees, and their contact activity will be collected by carrying a wearable device. From the GDPR perspective, the organization is the data controller and end-users are the data subjects.

-A set of Protection criteria ($\mathcal{P}$). Stakeholders have different protection criteria when they want to evaluate the risk of potential threats. For instance, organizations as the data controller typically adopt business impact criteria, such as financial or reputation impact, whereas data subjects evaluate risk on the basis of their impact on their personal sphere. We consider that the *financial situation* and *reputational* are the impact criteria for the organization, which are linked to indirect and direct pecuniary loss or damage deriving from privacy violations, and by taking inspiration from [10, 15], we have specified the following protection criteria for the employees in the system who are concerned about:

- **Health condition**: Since the contact tracing scenario has a direct relation with the health condition of employees, they are concerned about their health, any attack to disrupt contact data may result in wrongly tracing the contacts. Consequently, it impacts on the health condition of employees.
- **Individual freedom**: Employees are worried about being tracked, as we enumerated some privacy issues in Section 2.1. This could lead to profiling of user behaviors, such as relationships and interactions, and these may result in losing their individual freedom.
- **Social situation**: This criterion may be compromised when an information disclosure happens as a consequence of a data breach. This can result in more severe health concerns and make disease outbreaks more difficult to control. As a consequence, employees may face discrimination at the moment of the COVID-19 pandemic or in the future.

Additionally, each stakeholder has different preferences, which result in different importance given to different protection criteria; for this reason, as can be seen in (2) a family $\{PW_p^s\}_{p2P,s2S}$ of weights are associated to a protection criterion p for each stakeholder s. Therefore, in the Trace4Safe scenario, after meeting and discussion with the project partners, we consider the values $\{0.7, 0.3\}$ and $\{0.5, 0.2, 0.3\}$ as weights for the organization's and employees' protection criteria, respectively.

-A list of protection Goals ($\mathcal{G}$). The Standard Data protection Model (SDM) [12] uses the term "data protection goals" to describe certain categories of requirements derived from data protection law. These requirements are aimed at properties of lawful processing operations, which have to be ensured by technical and organizational measures. The SDM specifies CIA triad, *Unlinkability* & *Data minimization*, *Transparency*, and *Intervenability* as six protection goals of data protection. These goals align with our outlined requirements (in Table 1) for contact tracing systems. The SDM's CIA triad directly correlates with GR_1, GR_2, and GR_4. The *Unlinkability* & *Data minimization* goal of the SDM resonates with GR_3 and, in the ECT context, further emphasizes ER_1 and ER_2. The SDM's *Transparency* goal aligns with GR_7, GR_8, and GR_9, and for ECT, it also emphasizes ER_4. Lastly, the SDM's principle of *Intervenability* harmonizes with GR_{10}, GR_{11}, and, in the ECT environment, is crucial for addressing ER_3.

In summary, the artifacts enumerated above are considered input requirements for the methodology. A simplified representation of the methodology is illustrated in Figure 2, which consists of two main phases (the yellow and green boxes), namely: *Association* and *Evaluation*. In the following, we explain the processes and their outputs within these two phases.

3.1.1 Association Phase

This phase is semi-automated and defines the association between controls and threats $\{\mathcal{C}_T\}_{T2T}$, the aversion level mapping AI_p^s for each protection criteria $p \in \mathcal{P}$ and stakeholder $s \in \mathcal{S}$, and the observation weight (OW_T) for each threat $T \in \mathcal{T}$. The first two outputs of this phase are obtained with human intervention as the risk analyst needs to identify which security controls are effective for each threat and which is the level of aversion of each stakeholder for a given protection criteria to be violated, whereas the last one is automatically derived after the risk analyst has specified which protection goals are affected by each threat. The output of each process in this phase—i.e., emphasized with dashed lines border in Figure 2—consider as inputs for the next phase. Below, we briefly describe these three processes:

-Threat-Controls association. In this process, we must identify mitigation controls for each threat. As we reported earlier in Table 2, fifteen security and privacy mitigation controls have been identified for the selected threats. For instance, the controls *Using pseudo-random identifiers and changing over time* and *using Encryption methods* can mitigate the risk of *Tracking and Eavesdropping* threat (T1).

Table 2: Selected threats and their mitigation controls.

Threats ($\mathcal{T}$)	Mitigation Controls $\{\mathcal{C}_T\}_{T\in\{T1,T2,T3,T4,T5\}}$
T1- Tracking and Eavesdropping	c_1) Using time dependent pseudo-random identifiers c_2) Using Encryption methods
T2- Data Tampering	c_3) Encrypting data-at-rest and data-in-transit by using encrypted connections such as SSL, TSL, HTTPS, etc. c_4) Implementing access control mechanisms c_5) Implement multifactor authentication
T3- Unlimited Data Storage	c_6) Implementing deletion mechanism c_7) Ensuring data minimization c_8) Using anonymization techniques for data-at-rest
T4- Battery drain Attack (DoS Attack)	c_9) Network monitoring c_{10}) Regularly checking Token's battery c_{11}) Validating contacts on Token's side c_{12}) Define a rate-limitation (e.g., control the rate of requests sent or received by a Token).
T5- Intervenability Threat	c_{13}) Facilitating the report by workers to the system c_{14}) Handling the workers' change requests c_{15}) Informing the workers about data processing (e.g., providing information about their daily activities by mentioning the contacts, locations, etc.)

-Threat-Protection Criteria association. This association defines impact levels according to stakeholders. For each identified threat, we must estimate the adverse impact level on each protection criterion of the stakeholders. We considered the aversion level has a value between 0 and 4 (call that 0 means no, 1 low, 2 moderate, 3 critical, and 4 catastrophic impact) for each protection criterion p and stakeholder s. However, for the sake of brevity, we reported the result of this association containing the aversion levels for each threat on the stakeholders' criteria in Table 3. The

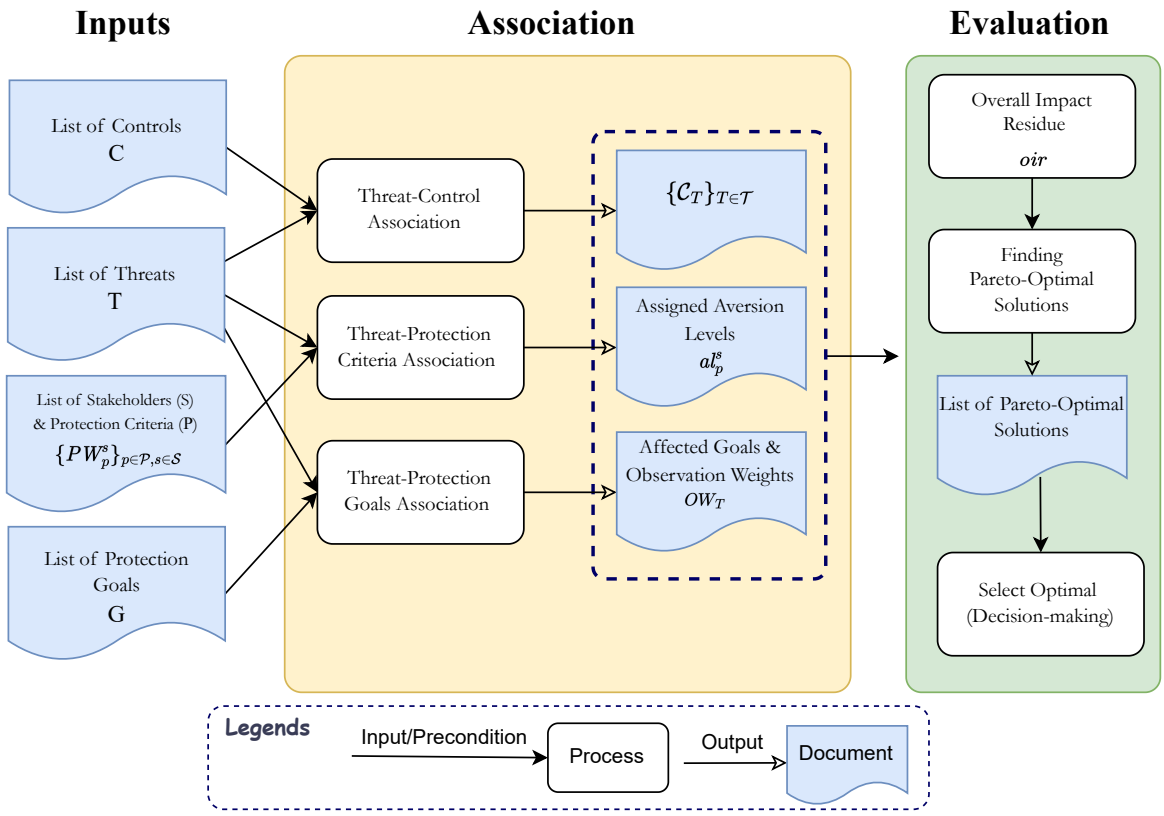

Figure 2: A simplified representation of the proposed methodology in [7] with an additional process for selection of the optimal solution.

Table 3: The estimated aversion levels for each selected threat on the stakeholders' protection criteria, and the assigned weight to each stakeholder's preference.

Stakeholder	Protection Criteria	Weights	Aversion Level				
			T1	T2	T3	T4	T5
Organization	Financial situation	0.7	3	3	3	3	1
	Reputational situation	0.3	3	2	1	2	1
Employee	Health condition	0.5	0	4	0	3	3
	Individual freedom	0.2	4	0	0	0	2
	Social situation	0.3	3	2	1	0	2

whole analysis is reported in Appendix B. For instance, according to Table 3 and also as reported in Appendix B, the aversion level for threat T1 and from the point of view of *Employee* for the protection criterion *Individual freedom* is 4, and the *weight* assigned to it is 0.2.

-Threat-Protection Goals association. This process specifies the relation between the protection goals and threats. The purpose of this process is to identify how many goals each threat is impacting. In Table 4, we specified the association between the identified threats and the protection goals. For instance, T1 affects goals **C**onfidentiality and **U**nlinkability & data minimization, while leaving unaffected the rest of goals. The observation weight value reported in the last column is computed by using the following formula:

$$OW_T = \frac{AG_T}{\sum_{T2T} AG_T} \tag{3}$$

in other words, it measures how much a threat T is pervasive for the goals in $\mathcal{G}$, where AG_T is the number of goals in $\mathcal{G}$ affected by a threat $T \in \mathcal{T}$. For example, in Table 4, the observation weight OW_{T2} is $1/8$, where **I**ntegrity is the only affected goal by T2, and the total number of affected goals is 8.

3.1.2 Evaluation Phase

The second phase is fully automated and aims to define and solve an instance of (1). This requires using the outputs of the first phase in order to define the impact level mapping i_s for each stakeholder $s \in \mathcal{S}$. However, solving the instance of (1) is considered a subjective assessment [7], and in order to have a less subjective assessment, they defined an approach where they measure the amplitude of the impact on each goal of a given threat. For that, (i) it requires computing the normalized threat criticality level (NTC) to quantify the severity of a threat using:

$$NTC_T = \frac{OW_T \times x_T}{\sum_{T2T}(OW_T \times x_T)} \tag{4}$$

where x_T represents the residual risk and is the replacement for $1 - \frac{c2c_T \mu_T(c)}{jc_T j}$ for each threat T. This is determined by applying a certain set of controls or, equivalently, by selecting a specific mitigation mapping μ_T from the set $\{0, 0.5, 1\}$.

Table 4: Association between threats and protection goals (the transparency goal is excluded in this table since it is not affected by any of these threats). A mark "×" indicates if a goal is affected by a specific threat.

Threat	Protection Goals					Observation
	C	I	A	U	In	Weight
T1	×			×		2/8
T2		×				1/8
T3	×			×		2/8
T4		×	×			2/8
T5					×	1/8

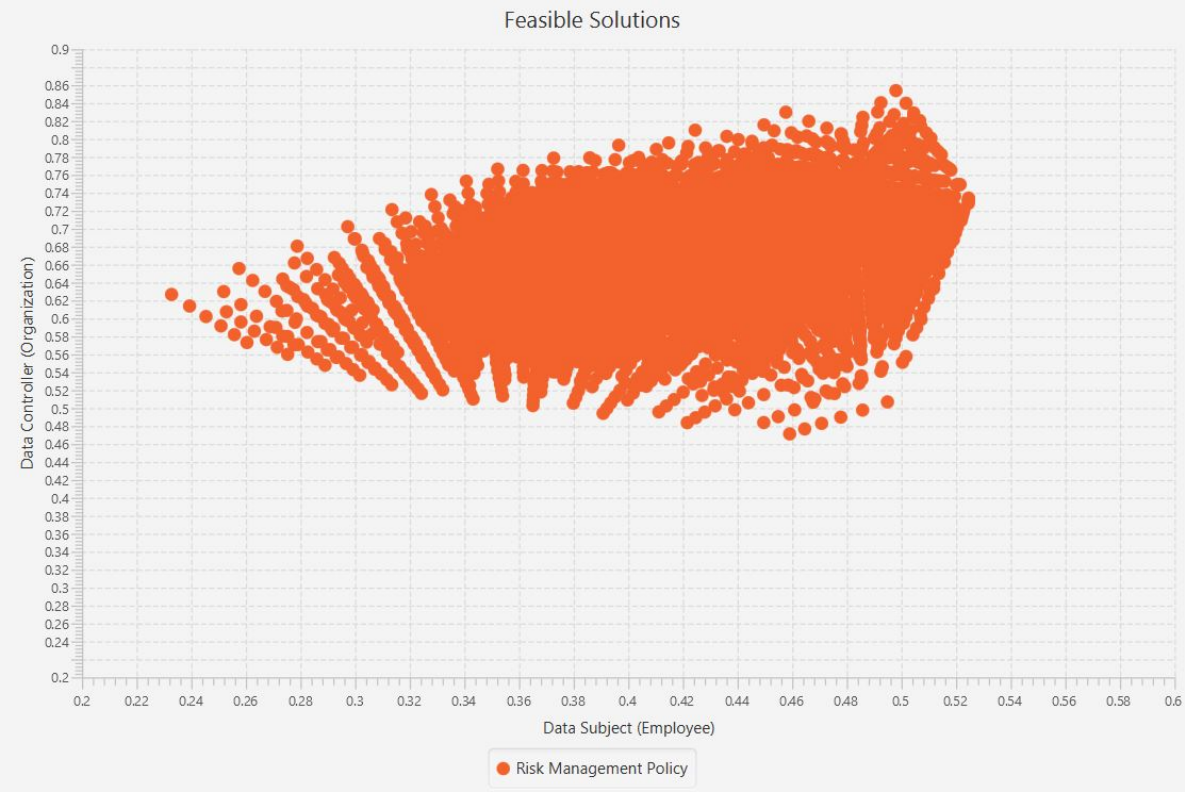

Figure 3: All feasible solutions in Trace4Safe scenario.

In this set, 0 indicates the control is not implemented, while 0.5 and 1 signify that the control is partially and fully effective in mitigating the threat T, respectively.

And (ii), use the NTC to weight the function i_s for each protection goal. The overall impact residue (oir) for a stakeholder s is defined as:

$$oir(s) = \sum_{G2G} \left(\frac{\sum_{T2T} \xi_{T,G} \times NTC_T \times i_s(T)}{\#(\mathcal{T}, G)} \right) \tag{5}$$

where $\xi_{T,G}$ is 1 if threat T compromises goal G and 0 otherwise. $\#(\mathcal{T}, G)$ is the number of threats that impact goal G. For example, if only one threat affects a goal, $\#(\mathcal{T}, G)$ is 1.

Now we are in the position to solve the following multi-objective risk minimization problem:

$$min_{h_{\mu_T i_{T2T}}} \quad \langle oir(s) \rangle_{s2S} \tag{6}$$

Briefly speaking, a combination of these three values ($i_s(T)$, x_T, and OW_T) will produce the overall impact residue (oir). By applying the Pareto optimality notion to the produced feasible set, a list of Pareto optimal solutions will be obtained. At this point, we can choose the best solution, often referred to as the decision-making process, from all the identified Pareto solutions.

4 Solving MSRMP instance of Trace4Safe

This section presents some experimental results from the developed tool in [7] for the Trace4Safe scenario. The input data for the tool are the list of threats and their mitigation controls (see Table 2), the stakeholders and their protection criteria introduced in Section 3.1.1 along with the aversion levels (see Table 3), and the association between threats and protection goals reported in Table 4. By applying the tool on Trace4Safe Scenario, the obtained feasible set contains 6,912 solutions. In Figure 3, we plotted the all feasible solutions and each point represents a solution under an RMP. The x-axis depicts the risk exposure level of the data subject (employee), while the y-axis depicts the risk exposure level of the data controller (organization).

After generating the feasible set of solutions, we need to apply the Pareto optimality algorithm to find the optimal solutions. Table 5 lists twenty-two found Pareto solutions, where the overall risk residue for the employee and the organization are reported in the second and third columns. In this table, the retrieved risk residuals (i.e., x_T) for these twenty-two solutions for each threat are reported in column fifth, which shows the residual risks under the identified Pareto solutions. We now face the task of choosing an optimal solution, a step inherent to optimization problems. In the following, we will explore various strategies for this selection process.

4.1 Optimal Selection Strategies

Selecting an optimal solution can be approached through various strategies. We propose three distinct strategies for this purpose:

Strategy 1. (i) The primary focus is on solutions with fewer residual risks (x_T) equal to 1, as this indicates a significant contribution towards mitigating threats. When several solutions satisfy this criterion, (ii) the observation weight value for the threats becomes the deciding factor (see Table 4). A threat that impacts multiple goals is inherently more pervasive. Therefore, the ideal solution should target the threat with the highest observation weight, ensuring it has the lowest residual risk.

Upon examining Table 5 and applying the first criterion, solutions ranging from S_{13} to S_{22} emerge as potential choices, each having only one residual risk set to 1. Among these, only S_{18} has a residual risk of 1 for threat T5, while the rest are associated with T3. Delving into the second criterion, T3, characterized as the *unlimited data storage* threat, affects two protection goals, namely *confidentiality* and *Unlinkability & data minimization*. On the other hand, T5, or the *intervenability* threat, solely impacts *intervenability*, in other words, T3 has a higher OW value than T5. Consequently, S_{18}, emphasized in gray, stands out as the optimal selection under this strategy, presenting overall risks of 0.3424 and 0.5154 for the data subject and data controller, respectively.

Strategy 2. This strategy employs a linear combination approach using a weighted sum function. In this method, the weighted sum function is applied directly to the identified Pareto solutions as:

$$Sum\ (S_j) = \sum_{i=1}^{jTj} (OW_{T_i} \times x_{T_i}) \tag{7}$$

Table 5: Pareto's solutions for Trace4Safe scenario along with their retrieved risk residual values. Legend; PS: Pareto Solution, DS: Data subject, DC: Data Controller.

			Retrieved risk residual				
PS	DS	DC	x_{T_1}	x_{T_2}	x_{T_3}	x_{T_4}	x_{T_5}
S_1	0.3243	0.5164	0.25	1	1	0.125	1
S_2	0.3134	0.5261	0.25	1	1	0.125	0.83
S_3	0.3017	0.5366	0.25	1	1	0.125	0.66
S_4	0.289	0.5478	0.25	1	1	0.125	0.5
S_5	0.2753	0.5599	0.25	1	1	0.125	0.33
S_6	0.2604	0.5731	0.25	1	1	0.125	0.16
S_7	0.3432	0.5101	0.25	1	0.83	0.125	1
S_8	0.3651	0.5028	0.25	1	0.66	0.125	1
S_9	0.3908	0.4942	0.25	1	0.5	0.125	1
S_{10}	0.4216	0.4838	0.25	1	0.33	0.125	1
S_{11}	0.4591	0.4713	0.25	1	0.16	0.125	1
S_{12}	0.323	0.5216	0.25	0.83	1	0.125	1
S_{13}	0.3116	0.5318	0.25	0.83	1	0.125	0.83
S_{14}	0.2993	0.5429	0.25	0.83	1	0.125	0.66
S_{15}	0.286	0.5548	0.25	0.83	1	0.125	0.5
S_{16}	0.2715	0.5678	0.25	0.83	1	0.125	0.33
S_{17}	0.2558	0.5819	0.25	0.83	1	0.125	0.16
S_{18}	0.3424	0.5154	0.25	0.83	0.83	0.125	1
S_{19}	0.2509	0.5916	0.25	0.66	1	0.125	0.16
S_{20}	0.2454	0.6021	0.25	0.5	1	0.125	0.16
S_{21}	0.2394	0.6138	0.25	0.33	1	0.125	0.16
S_{22}	0.2328	0.6267	0.25	0.16	1	0.125	0.16

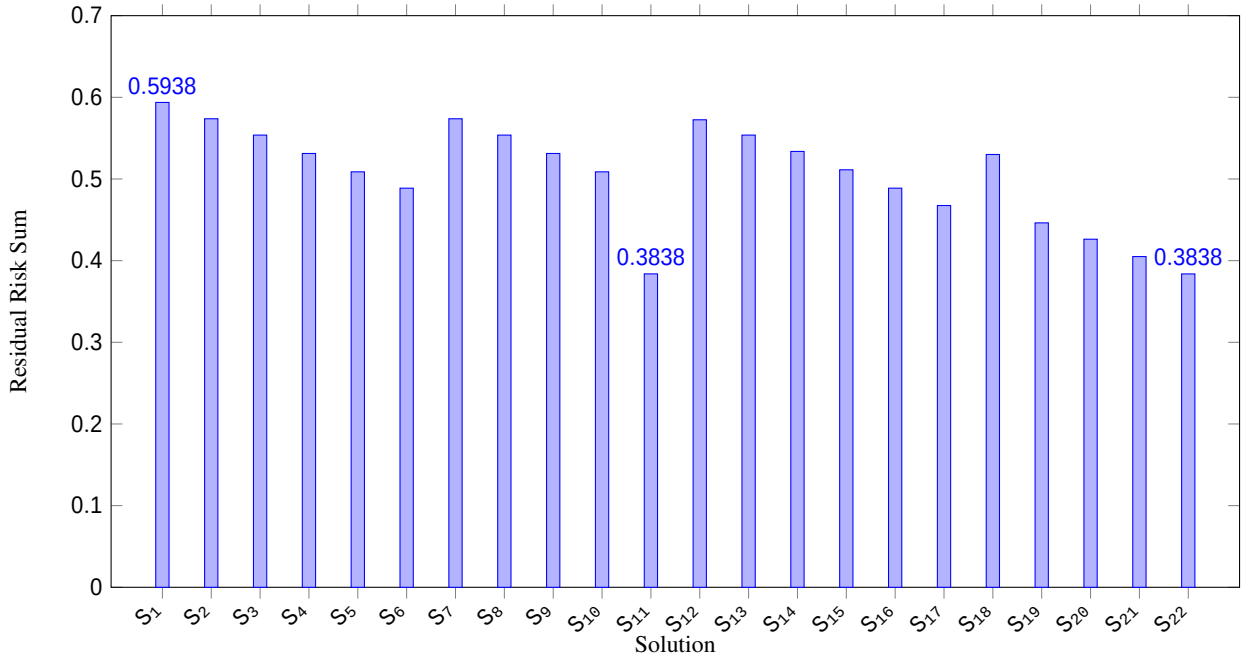

Figure 4: Bar chart representing the risk levels of identified solutions based on Strategy 2.

Here, j represents the index number of the solution, and OW_{T_i} represents the observation weight obtained for each threat, as detailed in Table 4. Using this function, we can systematically evaluate and rank the solutions based on their computed sum values. The resulting sum values for all identified Pareto solutions are depicted in Figure 4. This bar chart visualizes the risk residual sum levels for each solution. Notably, both S_{11} and S_{22} exhibit the lowest risk residual sum at 0.3838, positioning them as the optimal choices under this strategy. In contrast, S_1 has a sum of 0.5938, marking it as the least favorable. Consequently, the overall risks for the data subject and data controller when considering S_{11} are 0.4591 and 0.4713, respectively, while for S_{22}, they are 0.2328 and 0.6267.

Strategy 3. This strategy involves lexicographical ordering based on the OW_T values. In lexicographical ordering, solutions are ranked based on a sequence of criteria, starting with the most significant criterion. If two solutions are identical concerning the first criterion, the second criterion is used, and so on.

In our context, the ordering begins with the threats having the highest observation weights where the threats T1, T3, and T4 have the highest observation weights (i.e., 2/8). However, since the residual risk values for T1 and T4 are the same across all solutions, the primary differentiator becomes T3. In cases where solutions have identical residual risks for T3,

Table 6: The identified Pareto solutions ordered lexicographically.

	$x^*_{T_1}$	$x^*_{T_2}$	$x^*_{T_3}$	$x^*_{T_4}$	$x^*_{T_5}$
S_{11}	0.25	1	0.16	0.125	1
S_{10}	0.25	1	0.33	0.125	1
S_{9}	0.25	1	0.5	0.125	1
S_{8}	0.25	1	0.66	0.125	1
S_{18}	0.25	0.83	0.83	0.125	1
S_{7}	0.25	1	0.83	0.125	1
⋮	⋮	⋮	⋮	⋮	⋮
S_{22}	0.25	0.16	1	0.125	0.16
⋮	⋮	⋮	⋮	⋮	⋮
S_{1}	0.25	1	1	0.125	1

the next threats (based on their observation weights) are considered to differentiate between the solutions. For instance, both S_{18} and S_7 have a residual risk of 0.83 for threat T3. However, when we consider the next threat, T2, S_{18} has a more favorable residual risk than S_7, making S_{18} a better choice in this specific comparison. As shown in Table 6, the solutions are arranged lexicographically. Among the solutions, S_{11} has the lowest residual risk for the most critical threat, T3, making it the most favorable solution based on this strategy. On the other hand, S_1 is considered the least favorable due to its higher residual risks associated with the most critical threats.

4.2 Discussion

In the ECT scenario, understanding the threats and their implications on the protection goals is crucial. The identified threats, ranging from "*Tracking and Eavesdropping*" to "*Intervenability*", have varied impacts on the protection goals, as detailed in Table 4. Notably, threat T2 solely impacts the integrity goal, but it carries significant consequences where tampering data could lead to misinformation regarding COVID-19 exposure, posing a direct health risk to users, which elevates the importance of addressing T2 effectively. However, upon examining the outcomes of the strategies employed, solution S_{11} emerges as a balanced choice, presenting risks of 0.4591 for the data subject and 0.4713 for the data controller. In contrast, solution S_{18}, although advantageous for the data subject with a risk of 0.3424, poses a marginally increased risk of 0.5154 for the data controller. Meanwhile, solution S_{22}, while being the best for the data subject at 0.2328, raises concerns for the data controller with its risk value of 0.6267. Across the strategies, solution S_{11} consistently stands out. Yet, given the severe health implications associated with threat T2, it is imperative to select a solution that robustly addresses this threat. Solution 18 strikes a balance, adeptly reducing the risks linked to T2 and maintaining an equitable risk distribution for both involved parties. Factoring in the health consequences and the harmonized risk profile, solution S_{18} is posited as the most judicious choice for our ECT scenario.

To conclude the discussion, the third column of Table 7 illustrates five possible mitigation mappings associated with the solution S_{18}. Notice that all mitigation mappings associated with this solution suggest avoiding implementing any mitigation control for threat T5.

5 Related Work

This section reviews the literature on contact tracing systems with a particular focus on the integration of cybersecurity and the importance of Data Protection Impact Assessments.

-Cybersecurity in Contact Tracing Applications. The integration of cybersecurity in contact tracing applications is crucial to ensure user trust and effective implementation. Several research in this field have explored various

Table 7: Threats with associated security controls (first two columns) together with five possible mitigation mappings associated to the optimal solution (third column), and residual risk of optimal solution (fourth column). Legend: each control is associated to a mitigation level among three possible values ○ = 0 (the control has not been selected for implementation), ◐ = 0.5 (the control has been selected for implementation, but it is only partially effective to mitigate T), or ● = 1 (the control has been selected for implementation, and it is fully effective to mitigate T).

Threats ($\mathcal{T}$)	Controls $\{\mathcal{C}_T\}_{T2\,fT1,T2,T3,T4,T5g}$	Possible Mitigation Combinations					x_T
T1	c_1	●	◐	◐	●	●	0.25
	c_2	◐	●	●	◐	◐	
T2	c_3	◐	○	○	◐	○	0.83
	c_4	○	◐	○	○	◐	
	c_5	○	○	◐	○	○	
T3	c_6	○	○	◐	○	○	0.83
	c_7	○	◐	○	◐	○	
	c_8	◐	○	○	○	◐	
T4	c_9	●	◐	●	●	●	0.16
	c_{10}	●	●	◐	◐	●	
	c_{11}	●	●	●	●	◐	
	c_{12}	◐	●	●	●	●	
T5	c_{13}	○	○	○	○	○	1
	c_{14}	○	○	○	○	○	
	c_{15}	○	○	○	○	○	

aspects ranging from technical challenges to privacy concerns. Concerning technical challenges and solutions, the authors in [16] developed a privacy-preserving smartphone-based contact tracing method, highlighting the balance between privacy and epidemic control. Gupta et al. [17] presented a structured framework for analyzing contact tracing applications, focusing on security features, data privacy, and security vetting. Decentralized solutions in contact tracing applications have been found to be more privacy-preserving, as they distribute information across wireless networks [18]. Sahraoui et al. [19] proposed a new method based on online social networks for real-time contact detection and forecasting. Research has also focused on the ethical implications and the need for transparent citizen engagement in the deployment of contact tracing applications, where the authors in [20] presented an overview of the development of contact tracing and, suggested a reflection and possible solutions for the ethical and sustainable deployment through a more active and transparent citizen engagement. Additionally, the public's acceptance of contact tracing applications varies across cultures and socio-demographic strata, with varying levels of trust in data privacy [21].

-DPIA in Contact Tracing Applications. DPIAs for contact tracing applications, particularly in enterprise contexts, provide critical insights into the interplay of privacy, security, and legal compliance. For instance, Kouliaridis et al. [22] performed a detailed study of official Android contact tracing apps deployed by European countries, where their analysis results demonstrated that while overall these apps are well-engineered, they are not free of weaknesses, vulnerabilities, and misconfigurations that may ultimately put the user security and privacy at risk. Another exhaustive work done in Android applications by Hatamian et al. [23] who conducted a comprehensive analysis of 28 Android-based contact tracing apps, focusing on code quality, security, and privacy vulnerabilities, and adherence to legal requirements where their findings revealed the need for developers to take more cautionary steps to ensure code quality and address security and privacy vulnerabilities while consciously following legal requirements. Sun et al. [24] developed an automated security and privacy assessment tool called `COVIDGUARDIAN`, which combines identification and analysis of Personal Identification Information (PII), static program analysis and data flow analysis, to determine security and privacy weaknesses. Authors in [25] analyzed the NHSX contact tracing app's DPIA, pointing out significant data protection issues. The study highlighted the need for improvements in the DPIA process, addressing user rights, and monitoring concerns. A scientific DPIA on three published contact tracing app designs conducted by authors in [26], revealed weaknesses and risks, including insufficient anonymization, informed consent, and purpose-binding.

6 Conclusion

In this paper, we investigated the complex balance between cybersecurity, data protection, and the operational efficiency of Enterprise Contact Tracing (ECT) systems, with a particular focus on the Trace4Safe system amidst the COVID-19 pandemic. We identified key requirements and the interplay of security, privacy, and compliance challenges, emphasizing the necessity of integrating security and privacy by design principles to ensure GDPR compliance and safeguard sensitive data effectively. We executed a DPIA for the Trace4Safe system, by employing a methodology based on the Multi-Stakeholder Risk Minimization Problem (MSRMP), which has been pivotal in identifying optimal solutions that consider the varied perspectives and objectives of all stakeholders involved. The experimental results obtained from applying the MSRMP methodology to Trace4Safe, alongside exploring different strategies to select optimal solutions for risk management, furthered our understanding of the complexities inherent in ECT systems. Moreover, we discussed the implications of various threats to protection goals and the importance of selecting a solution that balances risks for all stakeholders which shows the dynamic and evolving nature of ECT. We concluded our work by analyzing specific mitigation mappings associated with the chosen solution, underscoring the trade-offs and decision-making processes essential for achieving an effective balance between organizational safety requirements and individual privacy rights.

References

[1] Thamer Altuwaiyan, Mohammad Hadian, and Xiaohui Liang. Epic: efficient privacy-preserving contact tracing for infection detection, 2018.

[2] World Health Organization et al. Contact tracing during an outbreak of ebola virus disease, 2014.

[3] Shamshul Bahri. Enhancing quality of data through automated sars contact tracing method using rfid technology, 2007.

[4] Katayoun Farrahi, Remi Emonet, and Manuel Cebrian. Epidemic contact tracing via communication traces. *PloS one*, 9(5):e95133, 2014.

[5] Regulation (eu) 2016/679 of the EUROPEAN parliament and of the council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. `https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN`, 2016.

[6] Majed Alshammari and Andrew Simpson. Towards a principled approach for engineering privacy by design, 2017.

[7] Majid Mollaeefar and Silvio Ranise. Identifying and quantifying trade-offs in multi-stakeholder risk evaluation with applications to the data protection impact assessment of the gdpr. *Computers & Security*, page 103206, 2023.

[8] European Data Protection Board. Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the covid-19 outbreak, 2020.

[9] EPFL and ETH Zurich advance digital contact tracing project. DP3T–decentralized privacy-preserving proximity tracing. `https://actu.epfl.ch/news/epfl-and-eth-zurich-advance-digital-contacttracin/(lastvisited9/9/2021)`, 2020.

[10] Marie Caroline Oetzel and Sarah Spiekermann. A systematic methodology for privacy impact assessments: a design science approach, 2014.

[11] Marie Caroline Oetzel, Sarah Spiekermann, Ingrid Grüning, Harald Kelter, and Sabine Mull. Privacy impact assessment guideline for rfid applications, 2011.

[12] Unabhängiges Landeszentrum für Datenschutz. The standard data protection model: A concept for inspection and consultation on the basis of unified protection goals. `https://www.datenschutzzentrum.de/uploads/sdm/SDM-Methodology_V2.0b.pdf`, 2020.

[13] Adam Shostack. *Threat modeling: Designing for security*. John Wiley & Sons, 2014.

[14] Kim Wuyts and Wouter Joosen. Linddun privacy threat modeling: a tutorial. `https://www.linddun.org/linddun`, 2015.

[15] Majid Mollaeefar, Alberto Siena, and Silvio Ranise. Multi-stakeholder cybersecurity risk assessment for data protection. In *Proceedings of the 17th International Joint Conference on e-Business and Telecommunications - Volume 3: SECRYPT*, pages 349–356. INSTICC, SciTePress, 2020.

[16] Tyler M. Yasaka, B. Lehrich, and Ronald Sahyouni. Peer-to-peer contact tracing: Development of a privacy-preserving smartphone app. *JMIR mHealth and uHealth*, 8, 2020.

[17] Rajan Gupta, G. Pandey, Poonam Chaudhary, and S. Pal. Technological and analytical review of contact tracing apps for covid-19 management. *Journal of Location Based Services*, 15:198 – 237, 2021.

[18] Viktoriia Shubina, Sylvia Holcer, Michael Gould, and E. Lohan. Survey of decentralized solutions with mobile devices for user location tracking, proximity detection, and contact tracing in the covid-19 era. *Data*, 5:87, 2020.

[19] Yesin Sahraoui, Ludovica De Lucia, A. Vegni, Kerrache Chaker Abdelaziz, M. Amadeo, and A. Korichi. Traceme: Real-time contact tracing and early prevention of covid-19 based on online social networks. *2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*, pages 893–896, 2022.

[20] Stéphane Roche. Smile, you're being traced! some thoughts about the ethical issues of digital contact tracing applications. *Journal of Location Based Services*, 14:71 – 91, 2020.

[21] My Zetterholm, Yan-Jin Lin, and P. Jokela. Digital contact tracing applications during covid-19: A scoping review about public acceptance. *Informatics*, 8:48, 2021.

[22] Vasileios Kouliaridis, G. Kambourakis, and Dimitrios Geneiatakis. Dissecting contact tracing apps in the android platform. *PLoS ONE*, 16, 2020.

[23] Majid Hatamian, Samuel Wairimu, Nurul Momen, and Lothar Fritsch. A privacy and security analysis of early-deployed covid-19 contact tracing android apps. *Empirical Software Engineering*, 26, 2021.

[24] Ruoxi Sun, Wei Wang, Minhui Xue, Gareth Tyson, S. Çamtepe, and D. Ranasinghe. An empirical assessment of global covid-19 contact tracing applications. *2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE)*, pages 1085–1097, 2021.

[25] Michael Veale. Analysis of the nhsx contact tracing app 'isle of wight' data protection impact assessment. 2020.

[26] Kirsten Bock, Christian Kühne, Rainer Mühlhoff, Meto R. Ost, Jörg Pohle, and Rainer Rehak. Data protection impact assessment for the corona app. *ArXiv*, abs/2101.07292, 2020.

Appendix

A Security and Privacy Threats

The identified threat scenarios along with their type (security or privacy) and the consequences are reported in Table 8. Each of these threats is mapped to the corresponding affected components/channels reported in Figure 1.

Table 8: Trac4Safe Scenario: The identified threats along with their consequences.

Threat Scenario	**Channels & Components**	**Type**	**Consequences**
T1- An adversary equipped with a Bluetooth Beacon Tracker can observe tokens, and in the case of token IDs do not change over the time, the attacker can re-identify token holders.	1 2 3 4	PR-Identifiability, Detectability	Tracking users, identifying users, profiling users' behavior, learning about places.
T2- An attacker can eavesdrop on the network traffic by setting up her device close to the gateways when data is uploaded on gateways.	3	SEC-Confidentiality, PR-Identifiability, Detectability	Identifying users.
T3- Tampering data may happen in different phases of data exchanging in the system.	2, 3, 4 Communication channels between edge server and cloud server	SEC-Tampering data	Loss of data integrity.
T4- An unauthorized access to the local stores.	5, 6	SEC-Unauthorized access	Disclosure information.
T5- An unauthorized access to the data/application	7, 8	SEC-Unauthorized access	Disclosure information, Impact on contact tracing network.
T6- A malicious user tries to submit the same data more than once to maliciously impact the protocol execution.	2, 3	SEC-Replay attack	Impact on contact tracing network.
T7- An adversary can set up his/her proximity tracking device, which is equipped with a sensitive antenna and powerful transmitter in a crowded space to increase the range of his/her Bluetooth contacts artificially. Consequently, other tokens consider it as a real contact due to feeling the proximity is in the defined range.	1, 3	SEC- False-positive contacts	Impact on the contact tracing network, receiving wrong notifications (RTLS).
T8- An attacker can make a denial of service to the gateways by sending massive contact messages or sending fake contact messages to impact on construction of the network of contacts which results in the wrong tracing of contacts.	Gateways	SEC- Denial of service	Impact on contact tracing network.
T9- The data is stored for longer and it increases the chance of data abuse and decreases its security.	5, 6	PR- Data longevity (unlimited data storage)	User identification.

Continued on next page

Table 8 – continued from previous page

Threat Scenario	Channels & Components	Type	Consequences
T10- Identifying an entity from a set of collected data, e.g., in our case, identifying positive cases.	5, 6	PR- Identifiability, Linkability	Re-identify users.
T11- An adversary can drain users' device batteries by sending fake contact messages which the victim device assumes as real contacts. The P2P approach may be more vulnerable to this type of attack.	-	SEC-Denial of service	Impact on contact tracing network.
T12- Users' information is shared with a third party or submitted to the health authority without their explicit consent.	-	PR- Policy and consent noncompliance	Non-compliance with the law.
T13- Lack of sufficient and complete description of the service and the operation details (such as data flows, data storage location, transmission methods, etc.) and their impacts on users' data.	-	PR- Lack of transparency	Non-compliance with the law.
T14- Users cannot submit correction requests (in the case of wrongly recorded contacts) that need to be evaluated by the system administrator. There is no implemented procedure in the system to allow the users to notify the system administrator to rectify, erase, or block the wrong registered contacts. For instance, in undefined events, if wrong contacts are uploaded (registered) in the system, it causes the contact tracing network to be created wrongly and results in incorrect notifications.	-	PR- lack of control and inability to rectify or erase the wrong registered contacts	Loss of trust in the system, impact on contact tracing network.
T15-An attacker tries to steal or make a copy of a worker's token. In this scenario, the attacker spoofs the token's identity to impact the contact's network.	-	SEC- Spoofing identity	Impact on contact tracing network.

B Risk Analysis (Aversion Level Estimation)

This section evaluates each of the selected threats (see Table 2) in order to determine their aversion levels. Thus, each threat will be challenged by its potential damage to stakeholders' protection criteria, and for that, we ask the question "what would the impact if threat...?". We considered two protection criteria for the organization (i.e., the data controller), namely: financial and reputational situations, and three protection criteria for the employees (i.e., the data subjects), namely: Health condition, Individual freedom, and Social situation. Table 9 outlines the approach used to evaluate the aversion level of each threat on the protection criteria.

Table 9: Stakeholders' protection criteria and impact levels.

What could be impacted on the protection criteria (for each perspective) if the threat happens?				
Organization		Employee		
Financial situation	Reputational situation	Health condition	Individual freedom	Social situation

0= No impact.
1= Low. The impact of any loss or damage is limited and calculable.
2= Medium. The impact of any loss or damage is considerable.
3= High. The impact of any loss or damage is significant.
4= Catastrophic. The impact of any loss or damage is devastating.

We give an in-depth analysis of the potential impact associated with the identified threats in Table 2 in the following. To accomplish so, we expand our explanation on each threat regarding:

- "What if...?" – The potential damages to the organization and employees can be anticipated.
- Impact Level – The estimated aversion level (e.g., no, low, medium, high, and catastrophic) for the different perspectives.

T1: Tracking and Eavesdropping. If encryption and anonymization measures are not implemented, or if random IDs are not generated and do not change over the time, the likelihood of tracking and eavesdropping threats increases, and the associated parties may face the following consequences:

- The organization's financial situation could be severely harmed, and the Trace4safe deployment effort could fail because employees may lose trust in the system.
- The organization's reputation will severely damage and may lead to losing its employees' trust, and as a consequence, the employee may not use/wear the Token device.
- The health condition of employees is not affected by the lack of mitigation controls for this threat.
- The individual freedom of the employees can be catastrophically affected in the case of tracking their contacts which may lead to losing their freedom.
- The social situation of the employees can be significantly affected due to tracking and identifying users (in particular positive cases) that may lead to discrimination or social pressure.

Aversion-Level of T1: High, High, No, Catastrophic, High

T2: Data Tampering. Security measures such as encrypting data-at-rest and data-in-transit by using encrypted connections such as SSL, TSL, HTTPS, etc., will protect data integrity. Accidental or deliberate attacks on the system can cause to impact data integrity, and the associated parties may face the following consequences:

- The organization's financial situation can be severely affected depending on the consequences, in particular, the rate of infection (the number of positive cases in the system) due to data tampering which the organization may even temporarily lose some employees due to quarantine regulations or in worse case the organization may decide to close some businesses or closing down some premises.

- The organization's reputation can be considerably affected depending on the consequences and how serious the data tampering is.
- The health condition of employees can be seriously affected since data tampering directly impacts the employees' health condition.
- The individual freedom of the employees will not be impacted due to data tampering attacks.
- The social situation of the employees can be considerably affected if their data is distorted (accidentally or deliberately), for instance, by being discriminated against by an unfavorable health condition.

Aversion-Level of T2: High, Medium, Catastrophic, No, Medium

T3: Unlimited Data Storage. The system must store data no longer than necessary need it, and it must explicitly define the expiration time for the collected data (any data related to COVID-19 positive cases and their contacts) because users—especially those who got infected—are concerned about how long their information will be kept in the database. Therefore, if data is stored longer than necessary and no clear rules are implemented to limit data storage, the associated parties may face the following consequences:

- The organization's financial situation can be severely affected by the increasing cost of storage and management of the IT infrastructure.
- It is very unlikely that the organization's reputation would be damaged by excessive storage of employees' contacts.
- The health condition of employees is not affected by excessive collection of their contacts.
- The individual freedom of the employees is not affected by excessive storage of employees' contacts.
- The social situation of the employees is not affected by the excessive storage of employees' contacts. However, as the volume of data stored grows, so does the possibility of data breaches and leaks.

Aversion-Level of T3: High, Low, No, No, Low

T4: Battery drain Attack (DoS Attack). An adversary can drain employees' Token batteries by sending fake contact messages that the victim's device assumes as real contacts. The P2P approach may be more prone to this type of attack. Therefore, this kind of attack can have a significant impact on the contact tracing network, and the associated parties may face the following consequences:

- The organization's financial situation can be severely affected. Because these kinds of attacks may impose some cost on the organization, such as the cost of recovery procedures or the cost of energy might need to recharge the Tokens.
- The organization's reputation can be considerably affected due to this Dos attack where the service might be temporarily unavailable and as a consequence, employees may lose trust in the system.
- The health condition of employees can be severely affected due to unavailable contact tracing services or fake messages that can create a wrong network of contacts.
- The individual freedom of the employees is not affected by this type of attack.
- The social situation of the employees is not affected by this type of attack.

Aversion-Level of T4: High, Medium, High, No, No

T5: Intervenability Threat. Suppose there is no implemented procedure in the system to allow the employees to notify the system administrator to rectify, erase, or block the wrong registered contacts. In that case, it causes the contact tracing network to be created incorrectly and results in incorrect notifications as well as a loss of trust in the system. Therefore under this threat, the associated parties may face the following consequences:

- The organization's financial condition may suffer if employees are unable to rectify their wrong contacts, for example, inaccurate data may result in employees being forced to take unneeded leave or being in quarantine.

- The organization's reputation may be affected if employees cannot rectify their wrong contacts. This results in low-quality data and thus, imperfect contact tracing service.
- The health condition of employees can be severely affected due to not being able to rectify wrong recorded contacts or not being able to report to the system for the case of being somewhere without carrying out the Token.
- The individual freedom of the employees may be slightly affected due to wrong recorded contacts. For instance, they may be forced to stay away from others for a while.
- The social situation of the employees may be slightly affected in case of wrong recorded contacts that may lead to discrimination or social pressure.

Aversion-Level of T5: Low, Low, High, Medium, Medium