Privacy Modeling and Enhancement Methodology for JCAS-Enabled Railway Applications

Yevhen Zolotavkin, Prajnamaya Dass, Stefan Köpsell

Barkhausen Institut, Dresden, Germany
{prajnamaya.dass|yevhen.zolotavkin |stefan.koepsell}@barkhauseninstitut.org

***Abstract*—In this paper, we present a new risk-based methodology for assessing privacy impacts in joint communication and sensing (JCAS)-based systems, with a focus on providing privacy-enhancing measures for 6G applications. To illustrate the methodology's applicability, we explore a railway use case where a JCAS system monitors level crossings for obstacles. Additionally, we offer a short demonstration of how the methodology applies to this use case scenario.**

***Index Terms*—Privacy, JCAS, railway, level crossing.**

I. Introduction

Due to its highly distributed topology, railway infrastructure is an important test ground for Joint Communication and Sensing (JCAS), where, besides functional requirements, privacy should also be tested and improved [1]. To better understand the privacy-limiting conditions of JCAS application in railway systems, we focus on JCAS-assisted level crossing (LX) monitoring, whose primary functionality is to produce early warnings about hazardous situations like the one depicted on fig. 1 [2].

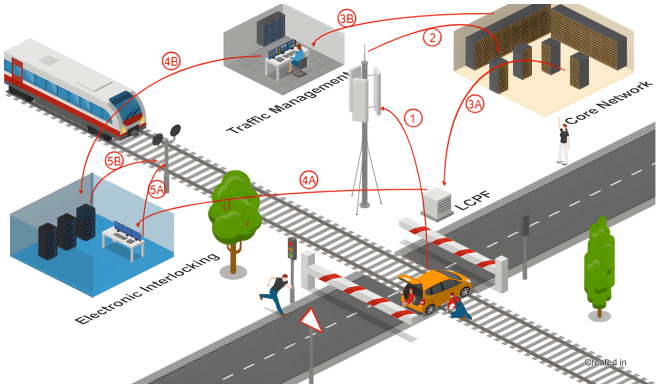

Fig. 1. OV-1 diagram of the level crossing use case.

We describe five main steps and distinguish between two modes (A and B) in implementing Sensing as a Service (SaaS) for LX monitoring. For both modes, the obstacle detection area between the barriers is sensed, (1), and the obtained data is sent for further pre-processing (e.g., clutter removal), (2). The following steps differ for A and B: a short message (e.g., a few Bytes) indicating the presence of the obstacle is sent to the Level Crossing Protection Facility (LCPF) at (3A), while richer sensing data (e.g., point clouds) is sent to the Traffic Management System (TMS) at (3B). Electronic Interlocking receives different (but short) warning commands from LCPF and TMS at (4A) and (4B), respectively. Based on those commands, corresponding encodings are sent to change the light signal at (5A) and (5B), respectively.

The described steps are focused on safety, while Personally Identifiable Information (PII) of natural persons in the sensed area also requires protection. Efficient PII protection involves reasoning about the components of the use case system architecture and the principles of the system's operation. The key questions that need to be answered by the privacy engineers are as follows: *i*) What are the privacy-related characteristics common to all system's building blocks (e.g., components) contributing to the system's privacy as a whole? *ii*) How can the holistic system privacy be expressed? *iii*) What are the steps leading to systematic privacy improvements?

To answer these questions, we propose a new system-wide and generalizable risk-based methodology for privacy modeling and enhancement. The rest of the paper is structured as follows. In section II, we outline the major blocks of our methodology. This is followed by the first such block in section III, where we elaborate on the steps of privacy modeling based on architectural descriptions and producing privacy views. The second block, dealing with the privacy enhancements for the modeled privacy views, is detailed in section IV and is followed by a short demonstration of the methodology's applicability in section V. Discussion about the specifics of the methodology and further steps concludes the paper in section VI.

II. Preliminaries

The methodology discussed further is based on 1) privacy-related modeling and 2) a corresponding constrained optimization approach. The diagram for our methodology is presented in fig. 2. It uses the principle IDEF0 notation, which can be interpreted as follows: Activity boxes represent functions (operations, processes) that accept inputs on the left of each box and produce outputs on the right. This transformation may be additionally controlled (constrained, instructed) by the inputs on the top. It can also consume resources (or rely on mechanisms and infrastructure) shown entering the box from the bottom. For example, the fig. 2 describes how the initial system (e.g., JCAS-LX) coupled with resources (arrows on the left) can be transformed into a privacy-enhanced system (arrow on the right). The privacy modeling is controlled by the principles of architectural description

(e.g., frameworks such as UAF and DoDAF) and privacy-related principles. The activity's output consists of a privacy view (privacy model) of the system and resource-dependent constraints applicable to that view. Privacy enhancement is a terminal activity that accepts the privacy view as input and constraints as a control to efficiently consume resources, allowing the production of an enhanced system.

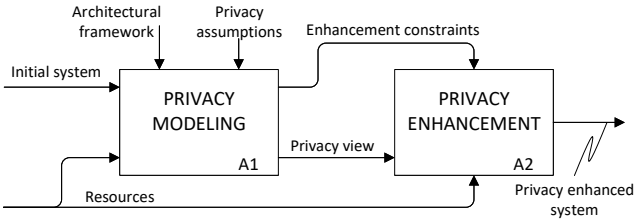

Fig. 2. Methodology-explaining IDEF0 diagram.

III. Privacy Modeling

The diagram for privacy modeling is presented in fig. 3. The distinctive feature of the diagram is the dependence of the privacy model on expert judgment about privacy characteristics. In turn, this judgment depends on the architectural view(s) of JCAS-LX produced for the expert. The benefit of the model is its ability to complement judgements about privacy characteristics and ensure their consistency: all this allows to provide risk-based interpretations for JCAS-LX. Modeling of constraints (based on resources) is not covered in this paper.

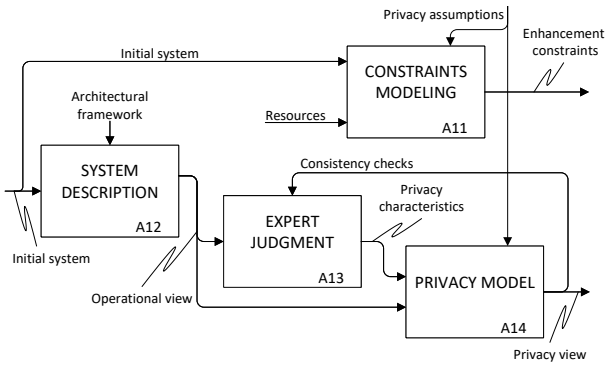

Fig. 3. IDEF0 diagram of 'PRIVACY MODELING' (see A1 in fig. 2).

A. System description

The diagram in fig. 4 is a further detalisation of the OV-1 diagram in fig. 1. It reflects interrelations between components of JCAS and railway systems, demonstrating why PII of natural persons in LX's proximity can be under threat well beyond the typically considered boundaries of JCAS. The diagram consists of Operational Roles (ORs) and Items of Exchange (IEs). ORs are assigned to the Performers on the Operational View (OV) diagram. IEs enable exchange between ORs: IEs may include various elements (e.g., data, signals, energy, etc.) [3].

For further analysis, we specify the boundaries of the considered subsystem: it consists of the JCAS and the Railway components only (excludes physical entities). Depending on

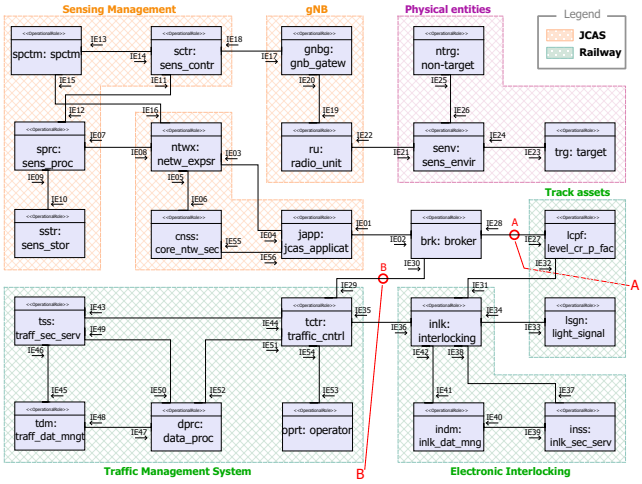

Fig. 4. OV-2 diagram of the JCAS-LX use case.

the mode of operation (`A` or `B`), the broker (`brk`) requests SaaS through the JCAS application (`japp`) interface and uses corresponding network exposure functions (`ntwx`). The request is authorized (based on privacy policies) by SPCTM (`spctm`), after which the sensing control (`sctr`) instructs the gNB gateway (`gnbg`) to use a radio unit (`ru`) for sensing. The gateway pre-processes the received raw sensing data: further processing is decided by sensing control, which typically delegates heavy pattern recognition tasks to sensing processing (`sprc`) and may store (`sstr`) the (intermediate) result. The result is sent through the broker (depending on the mode) to either LCPF (`lcpf`) or TMS (`tctr`). TMS may process rich sensing data in data processing (`dprc`) and/or data management instances (`tdm`), after which the operator (`oprt`) may make a decision to switch a warning (through interlocking, `inlk`) light signal on (`lsgn`). Instead, LCPF can send a short light-switching command and/or message through the interlocking[1].

B. Privacy assumptions

Our assumptions are focused on the amount of PII and access to PII in the context of the architectural view in fig. 4. The amount of PII passing through different architectural components per unit of time (e.g., 1 sec) may differ. This amount depends on the functional needs of the system and non-functional data minimization controls (e.g., based on obfuscation). In addition, PII flows transiting through different components may be dependent (e.g., correlate or be subflows of each other). The authenticity and confidentiality controls guarantee the strength of access control enforcement and information non-disclosure. The authentication and encryption protocols, respectively, supported by these controls, characterize them. The control strength can be increased through corresponding assurance techniques, including formal verification and attestation. However, in contrast to PII,

[1]The description is condensed due to the lack of space.

access enforcements are likely to be entirely independent for different ORs, for example.

1) Notations: We use the following notations to present privacy assumptions. For operational roles (OR), we use set $\mathbb{O}$ of indices i, $\mathbb{O} = \{\forall i\}$. For items of exchange (IE), we use set $\mathbb{I}$ of indices j, $\mathbb{I} = \{\forall j\}$. For every OR_i and IE_j, we introduce PII flow characteristics $\hat{p}_i$ and $\check{p}_j$, respectively, $p \in \mathcal{P}$, $\mathcal{P} \subset \mathbb{N}$. Similarly, for every OR_i and IE_j, we introduce characteristics of access strength $\hat{a}_i$ and $\check{a}_j$, respectively, $a \in \mathcal{A}$, $\mathcal{A} \subset \mathbb{N}$. The characteristics have the following meaning. Smaller p is preferable since it reduces the risk of linking or identifying a person. In contrast, larger a is preferable since it reflects the strength of technological means for access control and information flow control enforcement.

2) PII flows and access control: The following assumptions interpret relations between *a*) PII flows in IEs and PII content in ORs, *b*) access control strength in ORs, and flow control strength in IEs.

Assumption 1 (PII sensitivity): All PII parts are equally sensitive (e.g., there are no two items with an equal amount of PII, where one item contains more sensitive information than the other).

Assumption 2 (PII distribution - a): For every OR_i, except $\mathrm{OR}_{\mathtt{trg}}$, there is an immediately neighboring operational role, OR_{i^*}, whose PII payload is at least as large, $\hat{p}_{i^*} \geq \hat{p}_i$.
Comments on assumption 2: PII flow originates from $\mathrm{OR}_{\mathtt{trg}}$, and, hence, $\hat{p}_{\mathtt{trg}}$ is the largest in the system.

Assumption 3 (PII distribution - b): PII associated with IE is fully transferred into OR the IE inflows. In addition, every OR possesses (per unit of time) at least as much PII as the outflowing IEs carry from it.

Assumption 4 (PII distribution - c): We define the set of indices $\mathbb{I}^{[i]} \subset \mathbb{I}$ including all the IE_j items inflowing in or outflowing OR_i. The following property then holds:

$$\forall i \in \mathbb{O},\ \left(\hat{p}_i = \max_{j \in \mathbb{I}^{[i]}} \{\check{p}_j\}\right). \tag{1}$$

Comments on assumption 4: IEs disseminate PII across all ORs except `trg`, and therefore, we seek an expression explaining the relation between $\hat{p}_i$ and corresponding values $\check{p}_j,\ j \in \mathbb{I}^{[i]}$. For instance, any integral expression (e.g., defined through summation) would count on 2 or more items indexed in $\mathbb{I}^{[i]}$: the result may be inaccurate due to dependencies between these IEs. To avoid the latter while still aligning with assumption 3, we select a single maximum value $\check{p}_j$ whose index j belongs to $\mathbb{I}^{[i]}$.

Assumption 5 (Access strength): For every IE_j, we define the source and destination nodes of that item ($\hat{a}_{s(j)}$ and $\hat{a}_{d(j)}$) and $\mathbb{O}^{[j]} = \{\hat{a}_{s(j)}, \hat{a}_{d(j)}\}$. The following then holds:

$$\forall j \in \mathbb{I},\ \left(\check{a}_j = \min_{i \in \mathbb{O}^{[j]}} \{\hat{a}_i\}\right), \tag{2}$$

Comments on assumption 5: two major technical means for the flow control enforcement are authentication and encryption [1]. In eq. (2), we reflect that the mutual authentication strength and the strength of encryption for IE_j are defined by the capabilities of the weakest node indexed by $\mathbb{O}^{[j]}$.

Assumption 6 (Reference values): The reference values for $\hat{p}_i$ and $\check{a}_j$ are provided in tables I and II, respectively, and should be used for privacy assessments.
Comments on assumption 6: these values are consensus among the experts who took part in the study[2].

TABLE I
REFERENCE VALUES FOR $\check{p}_i$

Rate	**Fraction of PII***			
	N	*S*	*M*	*L*
Bit/sec	1	2	4	7
Kbit/sec	3	6	12	20
Mbit/sec	5	11	18	25
Gbit/sec	6	14	20	28
Tbit/sec	7	16	22	30

**N* - negligible; *S* - small; *M* - medium; *L* - large.

TABLE II
REFERENCE VALUES FOR $\hat{a}_j$

Authentication	**Encryption**	
	AES 128	AES 256
Cert. & Rem. Att.	8	10
Two-Factor	6	8
Certificate-Based	6	8
Token-Based	5	7
Pre-Shared Key	3	6
Password-based	1	2

3) Privacy risks: The adversary can compromise each of the OR_i and IE_j by exploiting corresponding weaknesses in these architectural components. As a result, valuable assets such as PII may be severely impacted.

Assumption 7 (Weaknesses): The weakness in OR_i is due to insufficient access control enforcement. The weakness in IE_j is due to insufficient information flow control enforcement.
Comment on assumption 7: Weakness becomes a vulnerability when its exploitation by a threat is detailed.

Assumption 8 (Threat): Threat is due to active adversary who can exploit component's weakness and compromise it by violating/bypassing access and information flow controls. He can then modify, disclose or misuse the PII extracted from the compromised component.
Comment on assumption 8: We consider only one kind of threat, which is assumed to be equally applicable to all the components. However, the likelihood of the corresponding attack being successful and the impact (of the consequences) it can cause on the asset (e.g., PII) may differ depending on the component.

Assumption 9 (Impact and likelihood): For both OR_i and IE_j, the impacts of the components' compromises are in a direct relation to the values of $\hat{p}_i$ and $\check{p}_j$, respectively. The likelihoods of such compromises are in an inverse relation to the values $\hat{a}_i$ and $\check{a}_j$, respectively.
Comment on assumption 9: The impact is typically determined by the value of an asset, which (according to assumption 1) only depends on the amount of PII.

Assumption 10 (Total risk): The quantitative level of the whole system's privacy risk, $\bar{r}$, satisfies the following:

$$\sum_{i \in \mathbb{O}} \hat{r}_i + \sum_{j \in \mathbb{I}} \check{r}_j \geq \bar{r} \geq \max\left\{\max_{i \in \mathbb{O}} \hat{r}_i, \max_{j \in \mathbb{I}} \check{r}_j\right\}. \tag{3}$$

Comment on assumption 10: Usage of the lower bound of $\bar{r}$ is justifiable in cases of high dependency among assets belonging to different components.

C. Privacy model

The purpose of the model is to prioritize privacy risks of architectural components. The principle of comparison

[2]The dependency between $\check{p}_i$ and the **Rate** is non-linear.

discussed here is equally applicable to the characteristics of ORs and IEs on the diagram (see fig. 4). For OR_i and IE_j components, we introduce composite privacy characteristics $\hat{\pi}_i = (\hat{p}_i, \hat{a}_i)$ and $\check{\pi}_j = (\check{p}_j, \check{a}_j)$, respectively: these characteristics are sufficient to define components' privacy risks and are grouped into the following sets: $\hat{\Pi} = \{\hat{\pi}_i\}_{\forall i \in \mathbb{O}}$, $\check{\Pi} = \{\check{\pi}_j\}_{\forall j \in \mathbb{I}}$.

Defining the least preferred architectural component (e.g., with the highest privacy risk r^*) requires specifying ordinal or cardinal utility over Π. The ordinal utility allows the establishment of preferences, such as $\pi_l \succsim \pi_k$.

Based on assumption 9, we substantiate the following axiomatic proposition:

$$\forall l \neq k \Big((p_l \leq p_k) \wedge (a_l \geq a_k) \implies \pi_l \succsim \pi_k \Big) . \quad (4)$$

Condition $(p_l > p_k) \wedge (a_l \geq a_k)$ is not addressed in eq. (4). For such a case, we define the following propositions:

$$\forall l \neq k \Big((p_l > p_k) \wedge (a_l \geq a_k) \overset{P}{\implies} \pi_l \succsim \pi_k \Big) \quad (5)$$

$$\forall l \neq k \Big((p_l > p_k) \wedge (a_l \geq a_k) \overset{1-P}{\implies} \pi_k \succ \pi_l \Big) \quad (6)$$

Probability P in eqs. (5) and (6) depends on (π_k, π_l). Such a probabilistic concept describes situations where consensus among experts is hardly achievable. The latter causes uncertainty as for the least preferred privacy characteristics in Π which can be expressed using an entropy: lower entropy generally means that the opinions about the architectural component with the least preferred π are closer to unanimity. If agreement is achievable among the experts evaluating privacy, a set of indifference curves can be defined by them on $\mathcal{P} \times \mathcal{A}$, similarly to how it is done for security parameters in [4]. Useful approximations can be obtained based on these curves resulting in a simpler and less costly model for judgments not requiring further involvement of experts. The algebraic function, $f : \mathcal{P} \times \mathcal{A} \to \mathbb{R}$, obtained from such an approximation becomes the cardinal utility of privacy. It can be used to define ordinal preferences as follows:

$$\forall l \neq k \Big(f(\pi_l) \geq f(\pi_k) \implies \pi_l \succsim \pi_k \Big) , \quad (7)$$

and must be consistent with eq. (4). Besides finding the component with the least preferred π_ω, the entire system's privacy evaluation θ can be expressed using $f(\pi_\omega)$:

$$\theta = f(\pi_\omega) = \min \Big\{ \min_{i \in \mathbb{O}} f(\hat{\pi}_i), \min_{j \in \mathbb{I}} f(\check{\pi}_j) \Big\} , \quad (8)$$

$$\omega = \arg\min \Big\{ \min_{i \in \mathbb{O}} f(\hat{\pi}_i), \min_{j \in \mathbb{I}} f(\check{\pi}_j) \Big\} , \quad (9)$$

where θ in eq. (8) and the lower bound on $\bar{r}$ in eq. (3) are in inverse relation (see assumption 10). In analogy to [5], we use the following function to establish privacy preferences:

$$f^*(\pi) = \frac{a}{p} , \quad (10)$$

from which we conclude that

$$\forall l \neq k \Big((p_l \leq p_k) \wedge (a_l \geq a_k) \implies f^*(\pi_l) \geq f^*(\pi_k) \Big) , \quad (11)$$

making eq. (10) consistent with eq. (7) and eq. (4). Next, we will outline a rationale for determining values for characteristics constituting $\hat{\pi}_i$ and $\check{\pi}_j$ in the JCAS-LX use case.

D. Expert judgment

For the diagram in fig. 4, some of the privacy-related characteristics can be defined by privacy expert utilizing additional information about the use case. For instance, among the main factors limiting higher $\hat{a}_i$ are computational resources and technologies (such as Public Key Infrastructure) available to OR_i. It is important that the expert-defined values of $\hat{a}_i$ are consistent with the table II. For every IE_j, parameter $\check{p}_j$ can be defined based on the operational and functional specifics (e.g., data rate) of the architecture and the knowledge about the PII collected and processed in the use case. These expert-defined values of $\check{p}_j$ must be consistent with the table I.

All other values can be inferred without expert's help. For example, $\forall j,\ \check{a}_j$ can be inferred based on eq. (2) since $\forall i,\ \hat{a}_i$ are already defined. Similarly, $\forall i,\ \hat{p}_i$ can be inferred based on eq. (1) since $\forall j,\ \check{p}_j$ are already known. However, it is important to note that without corresponding consistency checks, parameters $\check{p}_j$, defined by the expert may lead to inferring parameters $\hat{p}_i$, contradicting assumption 2.

IV. Privacy enhancement

The diagram for privacy enhancement is presented in fig. 5. Component-by-component enhancement is a distinct property of the diagram and contrasts with enhancing the whole system at once [1]. We reason about such a solution by starting with the holistic constrained optimization task and progressing toward an iterative procedure: it requires determining sub-constraints for each iteration. Examples of privacy enhancements are given in the next section. Resource usage is not addressed in this paper.

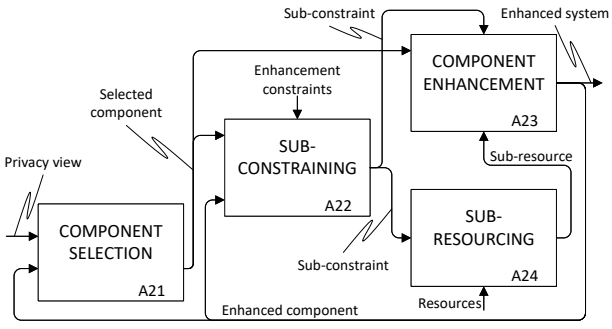

Fig. 5. IDEF0 diagram of 'PRIVACY ENHANCEMENT' (see A2 in fig. 2).

A. Iterative privacy improvement

Privacy criterion θ in eq. (8) needs to be improved under the set of constraints $\aleph$, $\aleph \subset \mathcal{P} \times \mathcal{A}$. The following constrained maximization task allows to obtain maximized $\dot{\theta}$:

$$\dot{\theta} = \max_{\pi \in \aleph} \min \left\{ \min_{i \in \mathbb{O}} f^*(\hat{\pi}_i), \min_{j \in \mathbb{I}} f^*(\check{\pi}_j) \right\}, \qquad (12)$$

Even though discrete function $f^*(\cdot,\cdot)$ is convex on $\mathcal{P} \times \mathcal{A}$, the optimization might be complicated since the whole goal function in eq. (12) is non-algebraic.

In contrast to the direct solving of eq. (12), iterative and component-wise enhancements are simpler, which explains their popularity in the fields of security and privacy [1]. To improve practicality of the privacy enhancing methodology, we propose the following modification of the task in eq. (12):

a) `start` iteration ι and ***define***

$$\omega^\iota = \arg\min \left\{ \min_{i \in \mathbb{O}} f^*(\hat{\pi}_i^\iota), \min_{j \in \mathbb{I}} f^*(\check{\pi}_j^\iota) \right\}; \qquad (13)$$

b) ***define*** sub-constraint $\aleph^\iota$ based on the whole system's constraints at time t_ι. If $\aleph^\iota = \varnothing$, `end`; else, ***enhance*** corresponding architectural component(s). As a result, the privacy characteristic of the component with index ω^ι undergoes a simplified convex optimization, $\pi^\iota_{\omega^\iota} \rightarrow \pi^{+\iota}_{\omega^\iota}$, where

$$f^*(\pi^{+\iota}_{\omega^\iota}) = \max_{\pi \in \aleph^\iota} f^*(\pi^\iota_{\omega^\iota}); \qquad (14)$$

c) increase $\iota := \iota + 1$, ***update*** $\forall i \in \mathbb{O}$, $\hat{\pi}_i^\iota$, and $\forall j \in \mathbb{I}$, $\check{\pi}_j^\iota$, ***adjust*** the total constraint, and `go to` item a).

B. Properties of the method

Several details of the method in section IV-A need to be further specified. First, there is a freedom of interpretation as to how $\aleph^\iota$ is defined during iteration ι, at step b). Assuming that constraints of the use case can be mapped into $\mathcal{P} \times \mathcal{A}$, the method should carefully constrain optimization for current ι, such that $\aleph^\iota \subseteq \aleph$. This is because enhancement of a single component at a time does not account for the situation when its privacy characteristic outperforms other components not being enhanced during that iteration. Improving $\pi^\iota_{\omega^\iota}$ beyond the worst privacy characteristic of the whole system contradicts eq. (13). This also does not increase $\dot{\theta}$ in eq. (12), which may lead to sub-optimal solutions. The (heuristic) remedy for this problem is to find a balance between heavily reducing sub-domain $\aleph^\iota$ and drastically increasing the total number of iterations.

Second, there is an implicit assumption that enhancing component's ω^ι privacy characteristic at iteration ι does not impede characteristics of other components. To demonstrate validity of this assumption, we need to analyze how characteristics are updated at step c). Our demonstration relies on the following (more refined) assumption. Reduction of PII in the ω^ι-th data flow does not increase PII in any of the other data flows:

$$\forall j \in \mathbb{I}(j \neq \omega^\iota) \Big(\big(\check{p}^{+\iota}_{\omega^\iota} \leq \check{p}^\iota_{\omega^\iota}\big) \implies \big(\check{p}^{+\iota}_j \leq \check{p}^\iota_j\big) \Big). \qquad (15)$$

Based on eq. (15) and eq. (1), we deduce that PII does not increase for operation roles as well:

$$\forall i \in \mathbb{O} \Big(\big(\check{p}^{+\iota}_{\omega^\iota} \leq \check{p}^\iota_{\omega^\iota}\big) \implies \big(\hat{p}^{+\iota}_i \leq \hat{p}^\iota_i\big) \Big). \qquad (16)$$

When ω^ι-th component is operational role, we note that

$$\forall i \in \mathbb{O}(i \neq \omega^\iota) \Big(\big(\hat{a}^{+\iota}_{\omega^\iota} \geq \hat{a}^\iota_{\omega^\iota}\big) \implies \big(\hat{a}^{+\iota}_i = \hat{a}^\iota_i\big) \Big) \qquad (17)$$

because access controls for different operational roles are independent. Based on eq. (17) and eq. (2) we deduce that the strength of access for information flows does not decrease:

$$\forall j \in \mathbb{I} \Big(\big(\hat{a}^{+\iota}_{\omega^\iota} \geq \hat{a}^\iota_{\omega^\iota}\big) \implies \big(\check{a}^{+\iota}_j \geq \check{a}^\iota_j\big) \Big). \qquad (18)$$

Third, previous reasoning implies that enhancement at the step b) of the method can be accomplished through privacy improvements (e.g., modernization, upgrades, data minimization) in the components other than the component with index ω_ι. Nevertheless, rationality of such an indirect enhancement should be carefully considered by the system's privacy engineer.

V. Demonstration

Here, we demonstrate the application of the proposed methodology. The demonstration is based on OV-2 diagram (see fig. 4), expert judgments about ORs and IEs, properties described by assumptions 2 to 6, and the enhancement steps from section IV.

A. Obtaining privacy characteristics

Expert judgments about comparative amounts of PII, $\check{p}_j$, for all the IEs, are provided in table III. Every cell in **IEs** column contains a pair of IE indices where the index of the IE with higher value of $\check{p}_j$ is highlighted with red: entries in all other cells of the row correspond to that IE only[3]. The same IE may be used in two different modes, A and B (see fig. 4). Judgments about the values of $\check{p}_j$ depend on the data **Rate** and fraction of PII in that specific mode of JCAS-LX operation and must be consistent with the information in table I.

Expert judgments about the comparative strength of access, $\hat{a}_i$, for all the ORs are provided in table IV: these characteristics must be consistent with the information in table II.

B. Inferring privacy characteristics

Based on the information in tables III and IV, we deduce $\forall i$, $\hat{p}_i$, and $\forall j$, $\check{a}_j$, using eqs. (1) and (2), respectively. For example, we deduce that $\hat{p}_{\texttt{oprt}} = 14$ and combine it with $\hat{a}_{\texttt{oprt}} = 2$ (assuming operator authenticates using passwords) from table IV. As a result, for that role, we obtain complete privacy characteristic, $\hat{\pi}_{\texttt{oprt}} = (14, 2)$: this can be compared with the characteristics of other components in fig. 4.

C. First iterations

We continue elaborating on the previous example for mode `B`. It can be seen that $f^*(\hat{\pi}_{\texttt{oprt}}) = 1/7$ has the lowest value meaning that $\text{OR}_{\texttt{oprt}}$ needs to be enhanced under the constraint $\aleph^1$. Based on eq. (10), this can be done with the help of security and privacy controls that reduce $\hat{p}_{\texttt{oprt}}$

[3]Such a representation avoids unnecessary information: paired IEs have the same $\check{a}$ (see eq. (2)). Hence, according to eqs. (8) and (10), only the IE with higher $\check{p}$ from the pair should be considered.

TABLE III
EXPERT-BASED ESTIMATION OF $\check{p}_j$ VALUES IN IES

IEs	Mode	Rate	PII*	$\check{p}_j$	Explanation
1, 2	B	1 Gbit/s	*S*	14	Point clouds
	A	100 bit/s	*N*	3	Objects detected ('Yes/No') within region, confidence levels
3, 4	B	1 Gbit/s	*S*	14	Point clouds
	A	10 Mbit/s	*N*	6	Object lists
5, 6	A-B	1 Mbit/s	*N*	5	Commands and msgs. to/from JCAS sec. serv.
7, 8	B	1 Gbit/s	*S*	14	Point clouds
	A	10 Mbit/s	*S*	12	Object lists
9, 10	B	1 Gbit/sec	*S*	14	Point clouds
	A	10 Mbit/s	*S*	12	Object lists
11, 12	B	1 Gbit/s	*S*	14	Pre-processed data, point clouds
	A	100 Mbit/s	*S*	13	Reduced resolution/measurement frequency
13, 14	A-B	100 Kbit/s	*N*	6	Sensing configuration according to policies
15, 16	A-B	100 Kbit/s	*N*	6	Data authorizations according to policies
17, 18	B	1 Gbit/s	*S*	14	Pre-processed data, point clouds
	A	100 Mbit/s	*S*	13	Reduced resolution/measurement frequency
19, 20	B	10 Gbit/s	*S*	15	Raw I/Q data (high frequency, resolution)
	A	1 Gbit/s	*S*	14	Raw I/Q data (reduced resolution/msrt. freq.)
21, 22	B	100 Gbit/s	*S*	15	Low level - phy. layer sensing - (transmn. of radio sign. with HF and massive MIMO) (high freq. and resolution)
	A	10 Gbit/s	*S*	14	Reduced resolution/measurement frequency
27, 28	A	1 Kbit/s	*N*	3	Commands and messages to/from LCPF
29, 30	B	1 Gbit/s	*S*	14	Pre-processed data, point clouds
31, 32	A	1 Kbit/s	*N*	3	Commands and messages to/from interlocking
33, 34	A-B	100 bit/s	*N*	2	Commands and messages to/from light sign.
35, 36	B	1 Kbit/s	*N*	3	Commands and messages to/from interlocking
37, 38	A-B	1 Kbit/s	*N*	3	Commands and messages to/from interlocking security service
39, 40	A-B	1 Kbit/s	*N*	3	Commands and messages to/from interlocking security service
41, 42	A-B	1 Kbit/s	*N*	3	Commands and messages to/from interlocking
43, 44	B	1 Mbit/s	*N*	5	Commands and msgs. to/from TMS sec. serv.
45, 46	B	1 Mbit/s	*N*	5	Commands and msgs. to/from TMS sec. serv.
47, 48	B	100 Mbit/s	*S*	12	Commands and msgs. to/from TMS data. proc.
49, 50	B	1 Mbit/s	*N*	5	Commands and msgs. to/from TMS sec. serv.
51, 52	B	1 Gbit/s	*S*	14	Rich sens. data to/from TMS data. proc.
53, 54	B	100 Kbit/s	*M*	14	Interpretable sens. inf. to/from TMS oprt.
55, 56	A-B	1 Mbit/s	*N*	5	Commands and msgs. to/from JCAS sec. serv.

TABLE IV
EXPERT-BASED ESTIMATION OF $\hat{a}_i$ VALUES IN ORS

OR	$\hat{a}_i$	Explanation	OR	$\hat{a}_i$	Explanation
brk	10	CB & RA, AES 256	japp	8	Cert. Based, AES 256
ntwx	8	Cert. Based, AES 256	cnss	8	Cert. Based, AES 256
sprc	8	Cert. Based, AES 256	sstr	8	Cert. Based, AES 256
sctr	8	Cert. Based, AES 256	spctm	8	Cert. Based, AES 256
gnbg	6	Cert. Based, AES 128	ru	6	Cert. Based, AES 128
oprt	2/6/8	Pass., AES 256 / 2FA, AES 128 / 2FA, AES 256	lcpf	8	Cert. Based, AES 256
inlk	8	Cert. Based, AES 256	lsgn	8	Cert. Based, AES 256
tctr	8	Cert. Based, AES 256	inss	8	Cert. Based, AES 256
indm	8	Cert. Based, AES 256	tss	8	Cert. Based, AES 256
tdm	8	Cert. Based, AES 256	dprc	8	Cert. Based, AES 256

and/or increase $\hat{a}_{\texttt{oprt}}$. For instance, few options are available for the latter (see table II): the technologies supporting the operator's access can be upgraded to two-factor authentication with AES 128 encryption for secure tunneling (e.g., VPN). Such an enhancement is described as $\hat{\pi}^1_{\texttt{oprt}} \rightarrow \hat{\pi}^{+1}_{\texttt{oprt}}$ where the new enhanced characteristic is $\hat{\pi}^{+1}_{\texttt{oprt}} = (14, 6)$ and $f^*\left(\hat{\pi}^{+1}_{\texttt{oprt}}\right) = 3/7$. If $\aleph^1$ permits, further enhancement (e.g., 2FA, AES 256) will provide $\hat{\pi}^{+1}_{\texttt{oprt}} = (14, 8)$ resulting in $f^*\left(\hat{\pi}^{+1}_{\texttt{oprt}}\right) = 4/7$. However, efficient sub-constraining during iterations should be carefully considered: another component may quickly become the 'weakest link' whose privacy needs to be enhanced (e.g., $\hat{\pi}_{\texttt{gnbg}} = (15, 6)$, $f^*(\hat{\pi}_{\texttt{gnbg}}) = 2/5$).

VI. DISCUSSION

As a result of answering questions *i*) - *iii*) (see section I), we developed a generalizable methodology based on a simple and interpretable characteristic $\pi = (p, a)$, allowing the expression of privacy risks for the whole system. The methodology includes steps for systematic privacy enhancement guided by a constraint optimization task. The applicability of the methodology has been demonstrated for the JCAS-LX use case, which is of immense importance in the context of the EU railway infrastructure.

Risk-based characteristic $\pi = (p, a)$ is straightforward and generalizable. It can be defined based on: a) Operational Views (e.g., UAF, DoDAF) popular among system engineers; b) knowledge about the authentication and encryption technologies supported by the system; c) reference tables (such as tables I and II). The set of privacy assumptions (see section III-B) and iterative privacy enhancement steps (see section IV-A) constitute the methodology's core. The assumptions explain and formalize the impacts and likelihoods of adversarial actions, while iterative steps aim at component-by-component enhancement in optimization sub-tasks. In contrast to the holistic risk treatment involving the entire system, only a subset of privacy-enhancing actions is considered at each iteration, significantly reducing complexity [1].

The developed methodology is applicable to both modes (A and B) of the JCAS-LX use case presented here. Nevertheless, the specifics of these modes are contrasting. For example, both A and B may benefit from enhancing the privacy of component $\text{OR}_{\texttt{gnbg}}$. On the one hand, in mode B, enhancement of $\text{OR}_{\texttt{gnbg}}$ is preconditioned by the characteristics of $\text{OR}_{\texttt{oprt}}$: if the operator's authentication is not upgraded from password-based to 2FA, enhancement of $\text{OR}_{\texttt{gnbg}}$ is inefficient. On the other hand, in mode B, enhancing authentication beyond 2FA, AES 128 is also inefficient since $\text{OR}_{\texttt{gnbg}}$ becomes the 'weakest link'. In contrast, mode A shows that $\text{OR}_{\texttt{gnbg}}$ and $\text{OR}_{\texttt{ru}}$ remain weaker than the rest of the components until substantial resource is spent on their enhancement. The latter demonstrates that the sub-constraining to simplify optimization is the key to timely switching between the enhancement tasks and needs to be better studied to avoid inefficiencies.

REFERENCES

[1] "ISO/IEC International Standard -Information security, cybersecurity and privacy protection – Privacy operationalisation model and method for engineering (POMME)," *ISO/IEC 27561:2024(en)*, pp. 1–36, 2024.

[2] G. Theeg and S. Vlasenko, *Railway Signalling and Interlocking International Compendium*, 3rd Ed. Edition Eurailpress, 2019.

[3] "IEEE/ISO/IEC International Standard for Software, systems and enterprise–Architecture description," *ISO/IEC/IEEE 42010:2022(E)*, pp. 1–74, 2022.

[4] A. J. Lohn, "Defense in Depth: The Basics of Blockade and Delay," 30, 2019. (visited on 09/26/2024).

[5] J. Gennari and D. Garlan, "Measuring attack surface in software architecture (CMU-ISR-11-121)," Carnegie Mellon University.