

Modelling Privacy Compliance in Cross-border Data Transfers with Bigraphs

Ebtihal Althubiti 

Northern Border University, Arar, Saudi Arabia

ebtihal.althubiti@nbu.edu.sa

University of Glasgow, Glasgow, United Kingdom

e.althubiti.1@research.gla.ac.uk

Michele Sevegnani 

University of Glasgow, Glasgow, United Kingdom

michele.sevegnani@glasgow.ac.uk

Advancements in information technology have led to the sharing of users' data across borders, raising privacy concerns, particularly when destination countries lack adequate protection measures. Regulations like the European General Data Protection Regulation (GDPR) govern international data transfers, imposing significant fines on companies failing to comply. To achieve compliance, we propose a privacy framework based on Milner's Bigraphical Reactive Systems (BRSs), a formalism modelling spatial and non-spatial relationships between entities. BRSs evolve over time via user-specified rewriting rules, defined algebraically and diagrammatically. In this paper, we rely on diagrammatic notations, enabling adoption by end-users and privacy experts without formal modelling backgrounds. The framework comprises predefined privacy reaction rules modelling GDPR requirements for international data transfers, properties expressed in Computation Tree Logic (CTL) to automatically verify these requirements with a model checker and sorting schemes to statically ensure models are well-formed. We demonstrate the framework's applicability by modelling WhatsApp's privacy policies.

1 Introduction

Transferring data across various jurisdictions worldwide is one of the requirements for advancing digital systems [27]. For example, WhatsApp states that they need to share users' data with Meta's data centres and Facebook's branches around the world to improve their services [33]. Transferring data to different countries can potentially threaten users' privacy as the receiver country may not use sufficient mechanisms to protect the data, *e.g.* encryption techniques [10].

Several regulations have been imposed to restrict such transfers, for example, the European Union (EU) General Data Protection Regulation (GDPR) [15], and the Australian Privacy Principles (APPs) [5]. The main aim of these regulations is to guarantee that the recipient country provides an adequate data protection level compared to the sender country's protection level.

According to the GDPR, there are two main ways to restrict data transfer towards non-EU countries: transferring based on the *adequacy decision* or transferring based on *appropriate safeguards* [9]. The former permits the transfer to specific countries that have an adequate level of data protection as specified by the European Commission. The latter requires safeguards, *e.g.* *Standard Contractual Clauses (SCCs)* or a certification, if the adequacy decision does not cover the receiver country. Besides these two methods, the GDPR identifies some exceptional cases where transferring data is permitted, *e.g.* if the transfer is important for the public interest. We discuss all these ways further in Section 2.

Organisations must adhere to the above requirements to transfer data from the EU to third countries. Otherwise, they can be at risk of being fined by an EU authority. For example, Facebook has been fined 1.2 billion euro by the Irish Data Protection Authority (IE DPA) for not adhering to the GDPR

requirements by transferring personal data from EU countries to the US¹ [7]. Such breaches could occur for many reasons. One is that the regulations are text-based and subject to update, which can create challenges for developers to convert them into technical requirements [17]. Another reason is that tools used by developers or supervisory authorities to check compliance are not automated, *e.g.* Transfer Impact Assessment (TIA) [30, 17]. This motivates the need for an automated approach that enables organisations to ensure their systems comply with the privacy regulations and serve as evidence of their compliance for supervisory authorities.

Formal methods are mathematical techniques that have been used in various domains to solve these issues thanks to their ability to model systems and provide a rigorous analysis of properties of interest such as security and safety [35]. They can also be used as proof of fulfilling these specifications because they enable automated and comprehensive verification processes. To the best of our knowledge, they have yet to be used to prove the systems' compliance with the GDPR requirements for cross-border personal data transfers.

In this paper, we propose a framework based on Milner's Bigraphs [25] to model systems and prove for the first time their compliance to these aspects of privacy regulations. Bigraphs are a universal formalism for the modelling of interacting systems that evolve in their connectivity and space via rewriting rules called *reaction rules*. Compared to other formalisms, reaction rules are user-specified allowing for greater flexibility to model a wide variety of systems. Bigraphs and reaction rules can be specified algebraically or through an equivalent diagram notation. This feature enables system designers to collaborate with privacy experts, *e.g.* privacy lawyers, who might not be expert in formal modelling and verification. Another advantage of using bigraphs is that modelling the flow of data among different jurisdictions requires modelling the spatial relationships between entities, and bigraphs can natively express spatial properties such as containment relation. Additionally, the sorting discipline provides another key advantage by categorising entities and links by means of sorts, ensuring that bigraphs are well-formed and preventing invalid compositions. This adds a layer of structural validation, making the model more robust.

Our approach is described in Fig. 1. The privacy aspects of our framework are predefined while system-specific aspects have to be specified for each application. The framework also provides a set of privacy properties that should be verified, *e.g.* providing safeguards, the validity of them *etc.* These privacy properties are expressed using the Computation Tree Logic (CTL) and can be checked *automatically* using off-the-shelf verification tools such as PRISM [22].

We provide the following contributions:

- We define a bigraph framework to model the GDPR requirements for cross-border data transfers. The framework captures *data transfers based on adequacy decisions and appropriate safeguards (Standard Contractual Clauses (SCCs) and certification mechanism)*. To the best of our knowledge, this is the first framework that models the GDPR notion of *international data transfers*.
- We apply the framework to an example based on WhatsApp's privacy policies and showcase the capabilities of bigraphs to model complex systems diagrammatically.
- We define the GDPR requirements for international data transfers using CTL and show how they can be proven *formally and automatically* using PRISM.
- We define sorting schemes that ensure the model is well-formed by assigning types to entities and links, effectively preventing nonsensical or undesirable model configurations.

¹This is before considering the US as an adequate country.

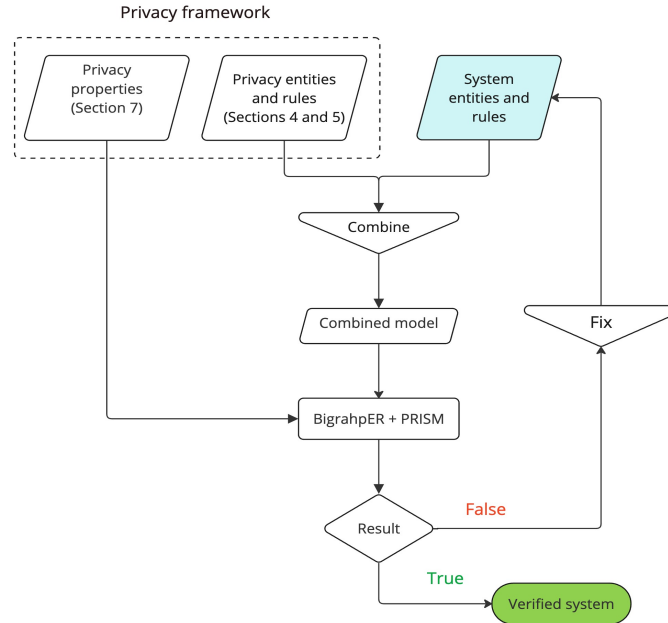


Figure 1: Overview of our bigraph-based privacy framework: The teal box represents the system entities and rules that end-users should specify. The process starts by combining the privacy framework with the system’s entities and rules (via links) into a unified model. This combined model is then analysed using BigraphER and the PRISM model checker to prove the privacy properties. If the verification result is true, it indicates that the system meets the GDPR requirements for cross-border data transfers. If false, the system model should be fixed to address the identified issues, after which it is reanalysed. This loop continues until the system passes verification.

The rest of this paper is structured as follows: Section 2 presents the GDPR requirements for international data transfers, while Section 3 gives an overview of Bigraphical Reactive Systems. Section 4 and Section 5 introduce the proposed privacy model, followed by Section 6, which demonstrates how we integrate the privacy framework with the specific system. We discuss the formal verification of the model in Section 7, while Section 8 displays the defined sorts of our model. We introduce an example of detecting privacy violations in Section 9. Section 10 and Section 11 present the related work and the conclusion, respectively.

2 Privacy Regulations for Cross-Border Data Transfers

Many governments have imposed legislation to regulate transferring data outside their countries or territories. They specify certain conditions to transfer data internationally as such transfers could affect their citizens’ privacy or national security (especially if the transferred data is sensitive) [10].

For example, the GDPR only allows international transfers in the following three cases:

1. **Adequacy decisions:** this requirement ensures that the recipient country provides a suitable level of data protection. The European Commission has specified a set of countries that provide an adequate level of protection, *e.g.* US, Canada, Japan *etc.* ².

²To view the full list of countries, see <https://t.ly/adequacy-decisions>.

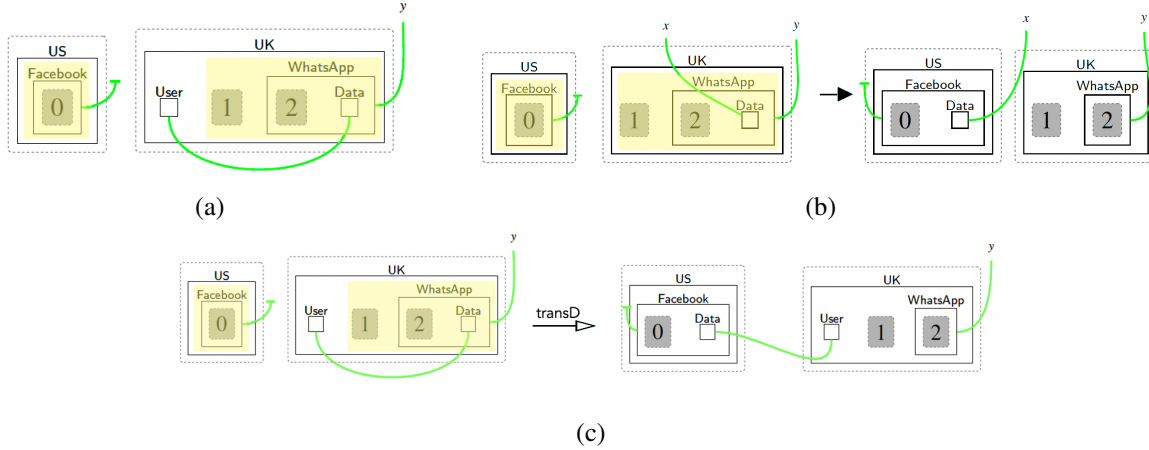


Figure 2: (a) A bigraph representing the initial state of a WhatsApp system that transfers Data to Facebook; (b) Rule transD transfers the Data from WhatsApp to Facebook; (c) The result of applying rule transD to the initial state, where Facebook acquires the Data.

2. **Appropriate safeguards:** if the adequacy decision does not cover the receiver country, then it should provide appropriate safeguards. The GDPR specifies many safeguards. One is adopting *Standard contractual clauses (SCCs)*, a set of predefined standards and rules identified by the European Commission to protect users' data when it is transferred outside EU countries. These rules should be incorporated into the contract between the sender and receiver organisation [11]. Another safeguard is using the *certification mechanism*. The certification should meet certain criteria (scheme) related to transparent processing and users' rights, *e.g.* clarifying the storage mechanisms of the data, data subjects rights to access *etc.* The certification also should be approved by certification bodies that the GDPR specifies, *e.g.* the European Data Protection Board (EDPB) [8].
3. **Derogations:** the GDPR identifies some special situations in which data transfer is permitted, *e.g.* when transferring data relates to the public interest³.

This paper models the transfer based on the adequacy decision and the appropriate safeguards. We focus on modelling the *SCCs and certification mechanism* as the SCCs are commonly used among organisations, and the certification should meet some criteria that need to be checked [9].

3 Bigraphical Reactive Systems

An initial bigraph and a set of reaction rules specifying the systems' evolution over time define Bigraphical Reactive Systems (BRSs). Unlike other formalisms, BRSs model systems diagrammatically and algebraically. In this paper, we only use the diagrammatic representation, but the equivalent algebraic specification is available online [2].

An example bigraph is in Fig. 2a: a WhatsApp system that transfers User's Data from the UK to Facebook in the US. Shapes represent entities, *e.g.* US, UK, *etc.* The shaded rectangles are *sites* that represent components that have been abstracted away as they are irrelevant to the context, *e.g.* site 0 could be a data centre of Facebook.

³For further details, see https://www.edpb.europa.eu/sme-data-protection-guide/international-data-transfers_en.

The dashed rectangles are called *regions*. Each parallel region can be considered as a modelling perspective that is used to split between different concerns [6, 29], *e.g.* the US is modelled in one perspective and the UK in another.

The green *links* are used to connect entities with each other. Links can be open to indicate the possibility of being linked to other unspecified entities *e.g.* link *y*, or closed to express for example exclusive ownership, *e.g.* the link that connects the User with the Data. Another way to close the links is making them one-to-zero hyperedge, *e.g.* the link that is attached to Facebook. We sometimes use coloured links for readability and visual clarity.

Bigraphs can evolve over time by using reaction rules. From now on, we will use rules to indicate reaction rules. Each rule ($L \rightarrow R$) contains two parts: the left-hand side (L) and the right-hand side (R). The left-hand side represents the pattern that will be changed, whereas the right-hand side represents the changed pattern. We use the notation $B \xrightarrow{r} B'$ to denote that B rewrites to B' when $r : L \rightarrow R$ is applied.

For example, rule `transD` in Fig. 2b models a transfer of the user Data to the US. By applying this rule to bigraph B shown in Fig. 2a, the part of the bigraph that matches the left-hand side of the rule (highlighted in yellow) is changed to match the right-hand side ⁴. Fig. 2c shows the result of applying rule `transD` (Fig. 2b) to bigraph B (Fig. 2a). The match is performed based on the structure of the bigraphs, not the links' names. This means the links' names are not considered for matching the left-hand side of the rule, *e.g.* changing link *y* in rule `transD` (Fig. 2b) to *z* does not affect the application of the rule.

Instantiation maps are a feature of BRS that allows us to copy, swap, or delete sites when the rules are applied. In this paper, we number sites to represent the instantiation maps. For instance, the sites are copied if shown in the left-hand side and the right-hand side of the rule as presented in Fig. 2b. Conversely, they are deleted if they appear in the left-hand side but do not exist in the right-hand side.

We specify and analyse our models using BigraphER [28], an open-source tool for modelling, rewriting, and visualising bigraphs. BigraphER offers two useful features: (1) enforcement of ordering over the rules via priority classes, *i.e.* each class is a set of rules and any rule in a class with lower priority can only be applied when no rules in any higher priority class can be applied, and (2) parameterized entities and rules, *e.g.* `Data(x)` and `transD(x)`, where x can be Name, ID, or Age, to allow sets of entities and rules (each entity or rule using one specific value for x).

4 Modelling Privacy Visually

We present our approach by applying it to an example based on WhatsApp's privacy policies. WhatsApp has two *data controllers* [34], one is located in Ireland to provide services to users in the EU countries, while the other is located in the US to serve users from other countries. These two controllers need to share users' data with the parent company Meta, its data centres, and Facebook's branches worldwide. For space, we present the data centre that is located in Ireland and Facebook's branches in Dubai and Mexico. A partial model of the example is shown in Fig. 3, while the full model is in [2]. The model consists of three perspectives: `WhatsSystem`, `SRTYPE` (for sender and receiver), and `Locations`. The specific system is modelled in the `WhatsSystem` perspective. It consists of the data controllers in the US and Ireland (`Whats`), the Facebook branches in Dubai and Mexico (`FB`), Meta company in the US, and the data centre in Ireland (`DC`). These entities are linked to their types in `SRTYPE` perspective, *e.g.*

⁴As this is an illustrative example, we transfer (or optionally duplicate) the data to the US without verifying the GDPR requirements for international data transfers.

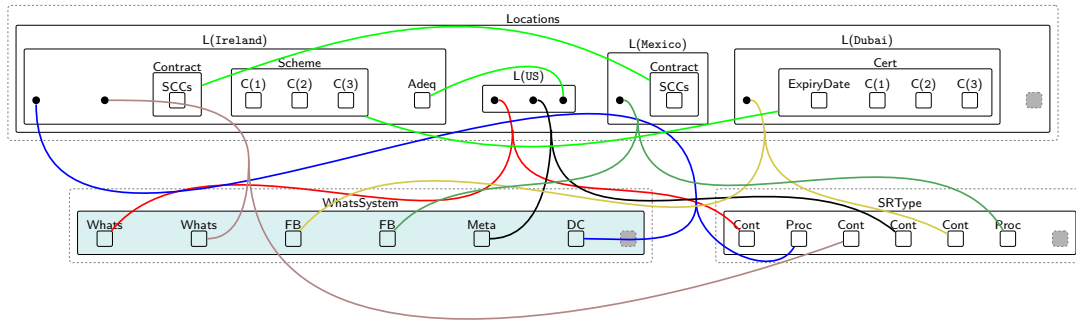


Figure 3: A partial initial state for the WhatsApp example. System-specific entities are shaded in teal. The predefined privacy model appears within the uncoloured regions.

data controller (Cont) or processor (Proc), and to their locations in Locations perspective as discussed in Section 4.1 and Section 4.2.

4.1 Specifying Sender and Receiver Types

SRType perspective is used to specify the types of each system-specific sender and receiver by means of linking, as shown in Fig. 3. This perspective can only contain entities of types *data controller* (Cont) and *processor* (Proc). According to the GDPR, the *data controller* determines the privacy policies and the purposes of processing users' data, whereas the processor processes the data based on the controller's instructions [18]. This perspective can easily be extended to support other types, *e.g.* sub-processor, joint controller *etc.* In Fig. 3, the system's entities that have the controller role (Whats, FB in Dubai, and Meta) are linked to their own Cont. Similarly, the system's processors (DC and FB in Mexico) are linked to their own Proc.

4.2 Specifying Sender and Receiver Locations

After assigning types to the sender and receiver entities, we need to specify their locations in the Locations perspective. We do so by linking them to an entity of type P (called *pointer*) nested within a location L(x). This is shown in Fig. 3 where, for instance, DC is linked to a P within L(Ireland). This approach allows for a great modelling flexibility as the number of entities in each location does not need to be specified and fixed beforehand. Also, by toggling between types P, P' and Ps, it is easy to define rules to check the adequacy decision and the safeguards when the sender and receiver are in different regions as we will explain in Section 5.1.

Table 1 describes the entities of this perspective. The parameterised entity L(x) allows end-users to add new countries, *e.g.* L(Japan). The entity SCCs is nested within Contract. The Contract and Cert are nested within L(x) to indicate that the safeguard provided by the organisation in country x is either the Standard contractual clauses or certification, respectively. Consider the example in Fig. 3. The entity L(Mexico) contains SCCs, which is nested within the Contract, meaning that the safeguard provided by Facebook in Mexico is the SCCs. Cert is nested within L(Dubai) to indicate that Facebook's branch

Table 1: Locations perspective entities.

Entity	Description
$L(x)$	Locations as parameterised entities where x is a country's name, <i>e.g.</i> UK.
Adeq	Linked to adequate countries.
Contract	The contract between the sender and the receiver.
SCCs	The Standard Contractual Clauses.
Scheme	The scheme specified by the GDPR, which the certification should comply with.
ExpiryDate	The expiry date of the certification.
$C(x)$	Criteria that should be checked to determine the valid certification. Parameter x is a criterion's identifier, <i>e.g.</i> 1, 2, Transparency, Right-to-access.
Cert	The certification provided by the receiver.
P, P', Ps	Pointers for each entities within a location. Shown as solid black, blue and yellow bullets, respectively.

in Dubai uses the certification safeguard. We do not allow the case when a company provides both safeguards as this is not included in the GDPR. Since, Ireland is the sender country in our example, it includes Adeq, Contract, and Scheme. The scheme contains three criteria, denoted as $C(x)$, where $x \in \{1, 2, 3\}$.

5 Checking privacy requirements

The modelling strategy we have described so far allows to only represent the status of a system at a given time. To model the temporal evolution of a system, we have also to define a set of reaction rules. In this section, we introduce the reaction rules encoding the processes required to check various privacy requirements. Unlike system-specific rules (see Section 6), they can be utilised across systems without changes.

5.1 Checking Regions

To check if data transfers are restricted, we need to check the region of the sender and the receiver. If both are located in the same region, the transfer is considered safe. Otherwise, it is a restricted transfer, and we need to check the adequacy decision and the safeguards.

The mechanism for checking the region is to tag the pointers linked to the sender and recipient **type**. As these pointers are nested within the entity $L(x)$, tagging them enables us to specify if the sender and receiver are in the same or different regions as explained in Section 5.2.

Rule `checkSType` (Fig. 4a) specifies the region of the sender, while rule `checkRType` (Fig. 4b) specifies the receiver's location. The entities `SType`, `RType`, and `CheckReg` are generated via system rule(s) in the system perspective, as outlined in Section 6. `SType` and `RType` represent the types of the sender and receiver, respectively (*e.g.*, site 0 can be either a `Cont` or `Proc`), and they are linked to the corresponding location of each. The entity `CheckReg` is generated by a system rule to trigger the privacy model.

Consider the controller `Whats`, located in `Ireland`, which needs to transfer data to the processor `DC`, also located in `Ireland`. To model this, we define system rules that specify the type of the sender (`Whats`) by generating `SType` around `Cont` and linking it to the hyper-edge that connects `Cont` with its

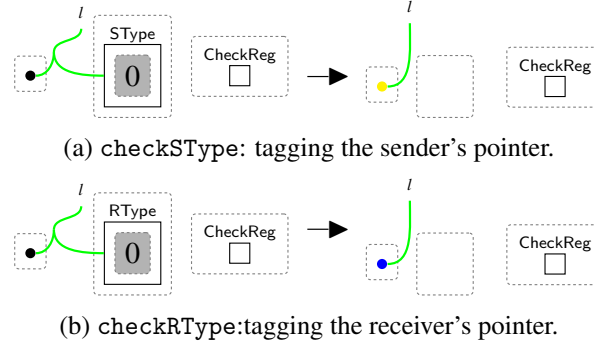


Figure 4: The process of checking the sender's and receiver's regions by tagging the pointers linked to their specified types.



Figure 5: `sameReg`: the sender and the receiver are in the same region.

pointer in $L(Ireland)$. Similarly, we specify the type for DC, generating $RType$ around $Proc$, as DC functions as a processor. Additionally, we generate the entity $CheckReg$ to initiate the checking process.

As we generate the $SType$ and $RType$ entities and link them to the locations of the specified types, we apply rule `checkSType` and `checkRType` (Fig. 4) to tag the pointers linked to $SType$ and $RType$, represented by yellow and blue bullets, respectively.

5.2 Entities are in the Same Region

If the tagged pointers (the yellow and blue pointers) are nested within the same region, then the transfer is safe. For instance, the tagged pointers in Section 5.1 are nested within $L(Ireland)$, so we can verify that the sender and the receiver are both in Ireland. The rule that models the result of checking the regions is `sameReg` (Fig. 5). The entity $CheckReg$ is replaced with $SameRegion$ to indicate the result of checking the region, *i.e.* the sender and the recipient are in the same region, and terminate the checking process. We also untag the pointers to recheck the regions if needed.

5.3 Checking Adequacy

Suppose the data sender is *Whats* in Ireland, and the receiver is *Meta* in the US. By checking their regions as explained in Section 5.1, we tag the pointers that are nested within the entities $L(Ireland)$ and $L(US)$, respectively. In such a case, the transfer is restricted because the tagged pointers are nested within different regions, so we must check the adequacy.

Rule `checkingAdeq` in Fig. 6 checks the adequacy requirement. This rule checks if there is a pointer in the relevant country linked to $Adeq$. If such a pointer exists, then the country is adequate. The entity $Adequate$ is generated to show the result of checking the adequacy requirement. The sender's and the receiver's pointers (the yellow and blue bullets, respectively) are untagged to allow reuse of the rule if needed. Importantly, failing to find the match of the left-hand side of this rule means the country is



Figure 6: checkingAdeq: checking the adequacy of the data importer country.

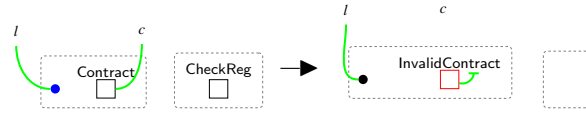


Figure 7: checkingSCCs: tagging the contract as invalid if it does not include the SCCs.

inadequate, so we should start checking the safeguards as shown in Section 5.4 and Section 5.5 which have higher priority than rule checkingAdeq (Fig. 6).

5.4 Checking SCCs

As mentioned previously, we must check the safeguards when rule checkingAdeq (Fig. 6) is not applied. In case the provided safeguard is SCCs, we use rule checkingSCCs in Fig. 7 to check whether the receiver country incorporates the SCCs in their Contract.

For example, the Contract of FB in Mexico includes SCCs and is linked to the entity Contract nested within L(Ireland), as shown in Fig. 3, indicating that Mexico agrees to adhere to the SCCs specified by the sender, thereby confirming the validity of the Contract.

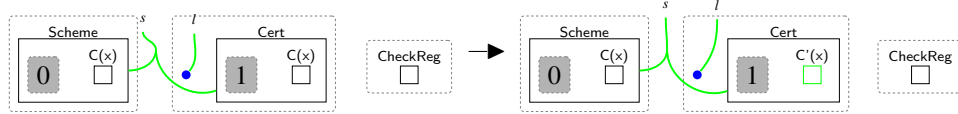
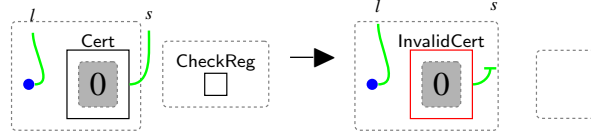
Conversely, if the Contract in the receiver country does not include SCCs, rule checkingSCCs (Fig. 7) is applied to tag the Contract as InvalidContract. This rule also results in the closure of the link connected to the entity Contract in Ireland, *i.e.* link *c*.

5.5 Checking Certification

If the provided safeguard is a certification (Cert), it must be checked to prove that it meets the Scheme specified by the GDPR. To do so, we use rule tagCriteria(*x*) shown in Fig. 8 to check each criterion (*x*). Suppose that the Scheme has three criteria: C(1), C(2) and C(3). We replace the parameter *x* in rule tagCriteria(*x*) (Fig. 8) with the number of criterion, *e.g.* tagCriteria(1), tagCriteria(2) *etc.*

Each rule checks the existence of the criterion in the Scheme and the provided Cert, *e.g.* checking that C(1) is nested within the Scheme and the Cert, by tagging the existing criterion (the green entity). Although the GDPR defines more than three criteria, we model only three as the modeller can extend the model by adding *x* number of criteria, *e.g.* C(4), C(5) \dots C(*x*).

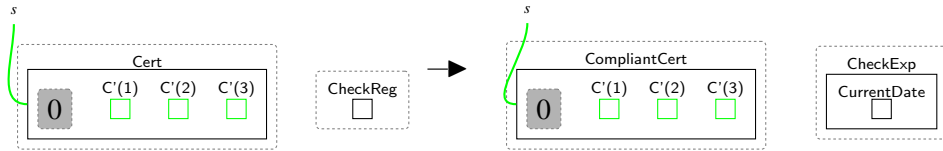
If **one** of the defined family rules is not applied, the Cert does not meet the criteria. For example, if rule tagCriteria(1) is not applied but rule tagCriteria(2) is applied, it means the Cert does not fulfil the first criterion (C(1)). In such a case, we use rule tagInvalidCert (Fig. 9) to tag the Cert as invalid (InvalidCert) and close link *s* that is connected to the Scheme. We omit CheckReg to end the checking process.

Figure 8: `tagCriteria(x)`: tagging each criterion that must be satisfied.Figure 9: `tagInvalidCert`: tagging the certification as invalid if at least one criterion is not satisfied.

5.6 Result of Checking Certification

Rule `checkingCertResult` (Fig. 10) shows the result of checking the criteria. If the Cert meets the criteria, *i.e.* $C(1)$, $C(2)$ and $C(3)$ are tagged, the Cert is tagged as `CompliantCert`.

Based on the GDPR, the certification (Cert) is valid for only **three years** [14]. This requires checking its validation date as well [8]. Rule `checkingCertResult` (Fig. 10) initialises the process of checking the expiration by replacing the entity `CheckReg` with `CheckExp` and specifying the current date (`CurrentDate`) to compare it with the expiry date of the Cert as shown in Section 5.7

Figure 10: `checkingCertResult`: result of checking the certification and initialising the process of checking its expiry date.

5.7 Checking the Expiry Date

As shown in Fig. 11, we can check the validation date of the `CompliantCert` by comparing the current date (`CurrentDate`) with the expiry date (`ExpiryDate`). If the `ExpiryDate` is **greater than** (`Greater`) the `CurrentDate`, it means the `CompliantCert` is expired and cannot be used as a safeguard. Otherwise, it is valid and can be used to transfer the data ⁵.

Rule `ExpiredDate` (Fig. 11) tags the `CompliantCert` as `InvalidCert` and removes link s connecting the `CompliantCert` to the `Scheme` if the `ExpiryDate` is `Greater` than the `CurrentDate`. The entities `CheckExp` and `CurrentDate` are omitted to terminate the checking process.

5.8 Withdrawing Certification

Based on the GDPR, the provider of the compliant certification (`CompliantCert`) has the right to withdraw it [8]. Rule `processWithdReq` (Fig. 12) models the provider's withdrawal request by closing link

⁵To generalise the model, we do not use mathematical computation to perform the comparison. However, we model the two cases: when the certification is expired and when it is valid. See the full model [2].

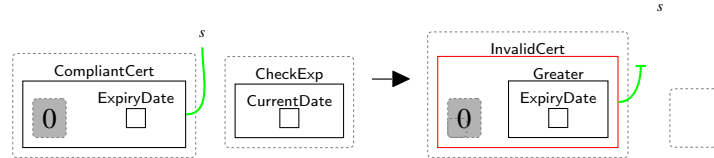


Figure 11: ExpiredDate: tagging the expired certification.

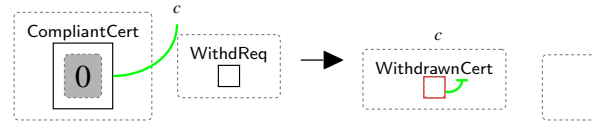


Figure 12: processWithdReq: processing the withdrawal request for the certification.

c connecting the CompliantCert to the Scheme and replacing CompliantCert with WithdrawnCert. The entity WithdReq is generated through a system rule (see rule ReqWithdCert in Fig. 25).

6 Integration of Privacy Models and Specific System

We introduce an example to show the privacy model’s usability and how we can merge it with the specific system. Suppose that the data centre of WhatsApp in Ireland (DC) needs to transfer users’ data to WhatsApp’s data centre in Singapore (DC) and to the Facebook branch in China (FB). Fig. 13 shows a partial initial state for this example. We consider the DC in Singapore to be a processor (linked to Proc) that has a Contract with the DC in Ireland. FB is considered as a data controller (linked to Cont), and the safeguard that it provides is a certification (Cert).

We need to trigger the privacy model to check the sender and receiver regions and their safeguards. To do so, we should specify the type of the sender (SType) and the receivers (RType) and generate the entity CheckReg via system rules. Rule dcIrType (Fig. 14) specifies the type of DC in Ireland. As DC is the sender and is linked to Proc, SType is generated around Proc to indicate that the sender is a processor. We also link SType to the sender’s pointer in L(Ireland). The entity StartTransfer is generated by a system rule (shown in Fig. 24 Appendix A). It is nested within DC as DC is the entity

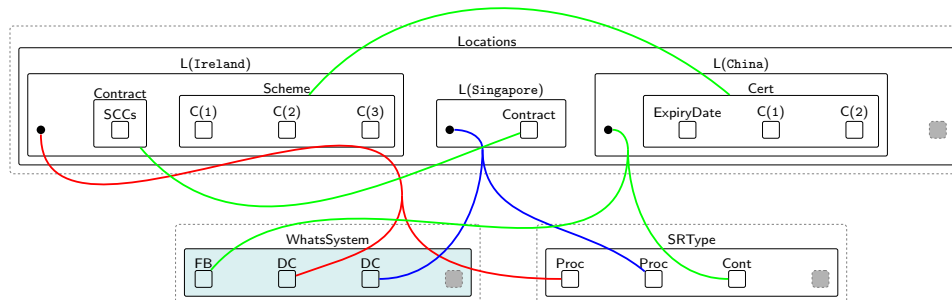


Figure 13: A Partial initial state for the example of transferring data to the data centre in Singapore and to Facebook in China.

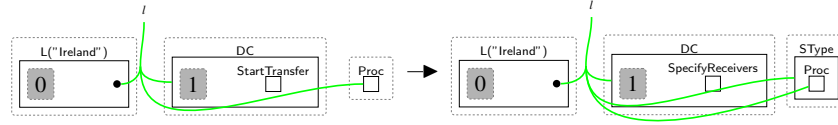


Figure 14: dcIrType: specifying the type of the sender (DC in Ireland).

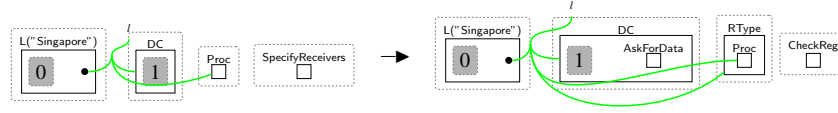


Figure 15: dcSingType: specifying the type of the receiver (DC in Singapore).

initiating the transfer. StartTransfer is replaced with SpecifyReceivers to specify the receivers types before transferring the data.

Rules dcSingType (Fig. 15) and FacebookCType (Fig. 16) specify the type of the receivers in Singapore (DC) and China (FB), respectively. Since DC is linked to Proc, the RType is generated around Proc and linked to DC's location. The entity AskForData represents the receiver's request for the data. FB is a data controller, so RType is generated around the token Cont and linked to China. We also replace SpecifyReceivers with CheckReg.

After specifying the sender and receivers types and generating CheckReg, we can start checking the regions by tagging the pointers linked to the specified types. We specify the region of the sender DC by using rule checkSType (Fig. 4a). We also use rule checkRType (Fig. 4b) to check the region of the receivers DC and FB. By applying these rules, the pointers that will be tagged are the pointers that are located in $L(\text{China})$, $L(\text{Singapore})$ and $L(\text{Ireland})$ as shown in Fig. 17. We can observe that the tagged pointers are located in different countries, so the transfer is restricted.

Based on Fig. 13, Singapore and China are not specified as adequate countries, *i.e.* no pointers in Singapore or China linked to Adeq, as is the case with the US in Fig. 3, meaning that the safeguards must be checked. The entity Contract that is nested within $L(\text{Singapore})$ does not include SCCs. This indicates that DC in Singapore does not incorporate the SCCs in their contract, *e.g.* does not agree to adhere to the SCCs. Rule checkingSCCs (Fig. 7) is applied to tag the Contract as InvalidContract and close its link connected to the entity Contract in Ireland. The entity InvalidContract is used as a flag to prevent the transfer and block the DC in Singapore through the system rule preventSing (Fig. 18). Similarly, FB uses the certification mechanism that has only two criteria (C(1) and C(2)) as shown in Fig. 13. By applying rule tagCriteria(x) (Fig. 8), only two criteria are tagged. In this case, rule InvalidCert (Fig. 9) tags the Cert as InvalidCert. To block FB, we define a system rule that matches explicitly on InvalidCert as shown in Fig. 19. Because we prevented the transfer to DC and FB, we can safely discard the entities InvalidContract and InvalidCert.

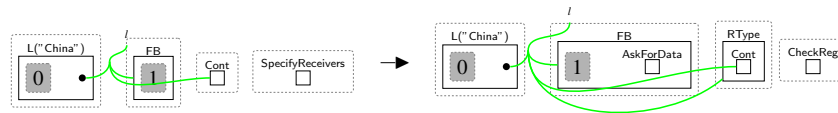


Figure 16: FacebookCType: specifying FB type.

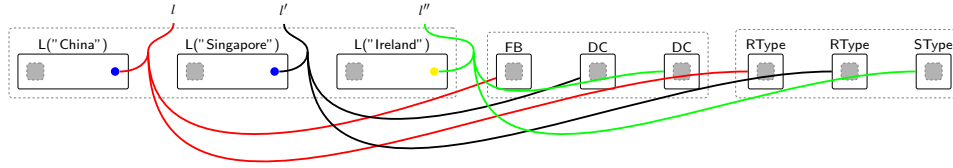


Figure 17: The tagged pointers after specifying DC in Ireland and Singapore and FB locations.

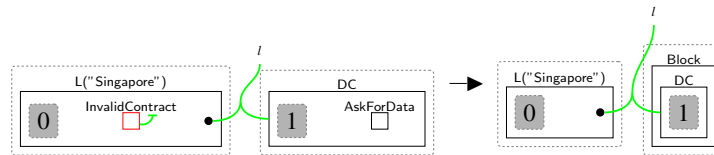


Figure 18: preventSing: preventing the transfer to the DC in Singapore.

7 Verification

The aim of defining the privacy model is to generate a transition system used to conduct the *verification* step, *i.e.* proving the GDPR requirements for international data transfer within the system. BigraphER utilises the initial state, *e.g.* Fig. 3, and the reaction rules (privacy rules and system rules) to *automatically* generate the transition system. The produced transition system consists of states (bigraphs) that describe the system configurations. Each state is a result of applying a rule (transition). Manually checking privacy properties, *i.e.* the GDPR requirements, within the transition system is challenging as the number of states can be considerable. For example, in the WhatsApp example, the number of the generated states is 193, while the number of the transitions is 267. This complexity would increase if we model a more complex system with additional reaction rules or locations.

Model checkers are tools used to *automatically* prove properties within the transition system. We use the PRISM [22] model checker as BigraphER supports it. To use PRISM, the properties that we aim to check should be encoded using a logic language. We use a non-probabilistic temporal logic language as the requirements we aim to check are specified based on *regulations*. In particular, we use Computation Tree Logic (CTL) [12] to encode the GDPR requirements for international data transfer because CTL quantifies the properties over all paths. This allows us to prove that the GDPR requirements hold across all possible execution paths. We also define a set of predicates using BigraphER to label the states. These predicates represent the right-hand sides of the reaction rules, and PRISM utilises them to parse the transition system. For instance, we define predicate `invalidContra` (Fig. 20) to label the states that

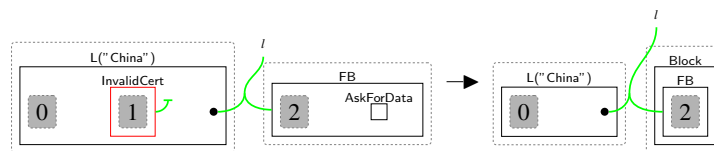
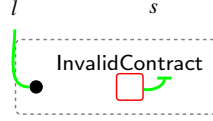


Figure 19: preventChina: preventing the transfer to the FB in China.

Figure 20: Bigraph pattern of the `invalidContra` predicate.

have the match of the right-hand side of rule `checkingSCCs` (Fig. 7).

The first requirement that we need to check is the adequacy requirement: *no data transfer if the country is inadequate*. The CTL syntax of this requirement is:

$$\mathbf{A}[\mathbf{G}(\neg \text{adequateCountry} \implies \neg \text{dataTransfer})] \quad (1)$$

`adequateCountry` is a predicate that labels all states that contain the match of the right-hand side of rule `checkingAdeq` (Fig. 6). The second predicate is `dataTransfer`, which labels the states that transfer the data to the receiver via system rules. \mathbf{A} is a path quantifier, meaning *for all paths*. \mathbf{G} is a state quantifier and it means *for all states* (globally). \mathbf{AG} means the property should hold for all paths globally throughout the entire transition system. Here, we check that: for all paths (\mathbf{A}), if `adequateCountry` does not hold (\neg), `dataTransfer` should not hold (\neg) in all states along that path (\mathbf{G}).

The second requirement that we need to check is: *no data transfer if the SCCs are invalid*:

$$\mathbf{A}[\mathbf{G}(\text{invalidContra} \implies \neg \text{dataTransfer})] \quad (2)$$

As previously mentioned, `invalidContra` is a predicate that holds if the `Contract` is invalid, so in this case we check: for all paths of the transition system, if `invalidContra` holds in a state, `dataTransfer` should not hold in the same state throughout the path.

The third requirement is: *no transfer should be performed if the certification is invalid*:

$$\mathbf{A}[\mathbf{G}(\text{invalidCert} \implies \neg \text{dataTransfer})] \quad (3)$$

This property proves that: it is always the case that if the certification is invalid, *e.g.* expired or does not meet the criteria, the transfer is never performed.

The last requirement is: *no transfer should be performed if the certification is withdrawn*:

$$\mathbf{A}[\mathbf{G}(\text{withdrawCert} \implies (\mathbf{X} \neg \text{dataTransfer}))] \quad (4)$$

Here, we prove that for all paths, if `withdrawCert` is true, then `dataTransfer` should be false in all subsequent states on that path. Since `withdrawCert` only becomes true when a system rule generates `WithdReq`, as discussed in Section 5.8, we apply the next-state operator \mathbf{X} to ensure that `dataTransfer` is false starting from the state immediately following the state that holds `withdrawCert`. This use of \mathbf{X} reflects the system's sequential processing order.

8 Sorting

Sorting schemes ensure that bigraphs are well-formed by imposing constraints that prevent undesired or nonsensical models [25], *e.g.* preventing scenarios where `Adeq` is incorrectly nested within `Scheme`, or where `Cert` is linked to `FB` instead of `Scheme`⁶.

⁶Currently, no automated tool supports these sorting schemes.

Table 2: Description of sort notations.

Notation	Example	Description
s^*	$A s^*$	The entity A has <i>zero</i> or more children of sort s .
$s \times s^*$	$A s \times s^*$	The entity A has at least <i>one</i> child of sort s .
$s \times i$	$A\{a \rightarrow s \times i\}$	The entity A has a port of sort a , which is connected to ports of sorts s AND i .
$s + i$	$A\{a \rightarrow s + i\}$	The entity A has a port of sort a , which is connected to a port of sort s OR i .
$s + 1$	$A s + 1$	The entity A has children of sort s OR it does not have any (<i>i.e.</i> it is empty).

We use the notation and grammar presented in [4] to define our sorts. Sorts can take two forms: sort y or sort $b = A$. The former specifies the sort of link ports, while the latter indicates that A is an atomic entity with sort b . If the entity A is linked to other entities and has children, its sort pattern is defined as:

$$\text{sort } b = A\{y \rightarrow z^*\} x + w$$

The sort y represents the sort of the port on A that connects to zero or more ports of sort z . Meanwhile, x and w refer to the sorts of A 's children, *i.e.* A has children of sort x *or* w . Table 2 describes through some examples the sort notations used in this paper.

To accurately model privacy and system entities, along with their associated behaviours, we define the necessary sorts for maintaining the integrity of the model and preventing ill-structured configurations across all perspectives.

SRType Perspective. We define the following sorts to ensure a well-formed perspective:

$$\begin{aligned} &\text{sort } a, \text{ sort } p, \text{ sort } et, \text{ sort } tp, \text{ sort } sy \\ &\text{sort } sr = \text{Cont}\{a \rightarrow (p + tp) \times sy + et\} \\ &\quad | \text{Proc}\{a \rightarrow (p + tp) \times sy + et\} \\ &\text{sort } ent = \text{SType}\{et \rightarrow (p + tp) \times a \times sy\} sr \mid \text{RType}\{et \rightarrow (p + tp) \times a \times sy\} sr \\ &\text{sort } srt = \text{SRType } sr \times sr^* \end{aligned}$$

The entities Cont and Proc , both having the sort sr , each have a port of sort a . This port is linked to:

- Either a port of sort p (for an untagged pointer) or tp (for a tagged pointer) along with a system entity via a port of sort sy ,
- Or to a port of sort et on the entity SType or RType .

The sort ent represents the sort of SType and RType . As these entities are used to specify the type of the sender and receiver (whether Cont or Proc), each must include at least one child of sort sr . Both SType and RType are linked to three ports: one on a tagged or untagged pointer, one on the entity type, and one on the system entity. The entity SRType , with sort srt , acts as the parent of Cont and Proc .

This imposes a structural constraint where SRTYPE contains one or more entities of sort sr , as illustrated in Figures 3 and 13.

By defining and applying these sorts, we ensure that site 0, nested within SType and RType in rule checkSType and checkRType (Fig. 4), either represents the entity Cont or Proc. We also ensure that SType and RType are linked to a pointer, with the other end of the link connected to a port on the system entity.

Locations Perspective. Similar to the SRTYPE perspective, we define the following sorts for the entities of the Locations perspectives:

```

sort c, sort i, sort t, sort p, sort st, sort d
sort cr = C(1) | C(2) | C(3)
sort tcr = C'(1) | C'(2) | C'(3)
sort e = ExpiryDate
sort g = Greater e
sort s = SCCs
sort ad = Adeq{d → p × p*}
sort pnt = P{p → (sy × a) + d + et} | P'{tp → (sy × a) + et} | Ps{tp → (sy × a) + et}
sort scm = Scheme{sc → c × c*} cr × cr*
sort certf = Cert{c → sc} (cr* + tcr*) × e
            | InvalidCert{c → 1} tcr* × e
            | CompliantCert{c → sc} (tcr × tcr*) × e
sort ctr = Contract{t → t × t*} s + 1
sort inctr = InvalidContract{i → 1}
sort l = L (Ireland) (pnt* × pnt) × (ad + scm + ctr)
        | L (US) (pnt* × pnt)
        | L (Mexico) (pnt* × pnt) × (ctr + inctr)
        | L (Singapore) (pnt* × pnt) × (ctr + inctr)
        | L (Dubai) (pnt* × pnt) × certf
        | L (China) (pnt* × pnt) × certf
sort loc = Locations l* × l

```

We assign the sort pnt to the pointers P , P' , and Ps , which represent the solid black, blue, and yellow bullets, respectively. Pointer P has a port of sort p , while pointers P' and Ps have ports of sort tp . These pointers are linked to two ports: one on a system entity (with sort sy) and one on a corresponding entity type within the SRTYPE perspective via a port of sort a . Additionally, pointer P can be linked to a port of sort d on the entity Adeq, or to SType or RType via a port of sort et , while pointers P' and Ps should not be linked to a port of sort d .

The entity Scheme is linked to one or more certifications (Cert) and must always include at least one criterion of sort cr or tcr . The certifications, whether invalid or compliant, are assigned the sort $certf$. Both Cert and CompliantCert should be linked to a port on the entity Scheme, and both have a child of sort e . However, Cert may contain zero or more *untagged* or *tagged* criteria, while CompliantCert must always contain at least one *tagged* criterion to satisfy the criteria specified by the Scheme. In contrast, InvalidCert can include zero or more tagged criteria. As shown in Fig. 9 and Fig. 11, InvalidCert has a closed link, denoted by 1 in the defined sorts, indicating the absence of a connection to Scheme.

The sort of Contract is ctr . Since Contract is linked to one or more contracts, the output and input ports share the same sort (t). Contract may either have a child of sort s , or it may be empty, in which

case it can be tagged as `InvalidContract` by rule `checkingSCCs` (Fig. 7), which also closes its link.

The parameterised entity `L` is defined with the sort `l`. The sorts of its children vary based on its parameters. For example, `L(Ireland)` is the sender country and must contain a child of sort `ad`, `sc`, or `ctr` to allow the GDPR compliance check.

Receiver countries that use SSCs, *e.g.* `L(Mexico)` and `L(Singapore)`, must include the entity `Contract` (sort `ctr`) or `InvalidContract` (sort `inctr`) if the contract is empty, *i.e.* it does not contain SSCs. In contrast, countries like `L(Dubai)` and `L(China)` use a certification mechanism, so their children must be of sort `certf`. If the receiver country is adequate, it does not contain any safeguards, meaning it includes only tagged or untagged pointers.

Regardless of the value of the parameter for entity `L`, it must always contain at least one pointer (tagged or untagged) that links to system entities and their types. The parent of `L` is the entity `Locations` that has sort `loc` and it should contain at least one child of sort `l`.

WhatsSystem Perspective. The system consists of various entities such as `DC`, `FB`, and others, each of which has specific roles and children entities based on their functions in the data exchange process:

```

sort snd = SpecifyReceivers | CheckReg | SameRegion | Adequate
sort req = AskForData | WithdReq
sort info = CopyInfo
sort cd = CurrentDate
sort chex = CheckExp cd
sort sys = DC{sy → (p + tp) × a + et} snd + req + info + chex + 1
           | FB{sy → (p + tp) × a + et} req + info + 1
sort b = Block sys

```

`DC` and `FB` are connected by a hyper-edge to a tagged or untagged pointer, as well as to their corresponding types, *e.g.* `Cont`. When the entities `SType` and `RType` are generated, they can be linked to `DC` or `FB` via a port of sort `et`. In our model, `DC` can act as both a sender and a receiver. When it is a sender, it contains the following entities: `SpecifyReceivers`, `CheckReg` or `SameRegion`. When `DC` is a receiver, it can contain `AskForData` or `WithdReq`. `FB` and Other companies in the model, *e.g.* `Whats`, `Meta` *etc.*, contains children such as `AskForData` or `CopyInfo`, which are specifically related to their role in requesting and receiving data. The entity `Block` is used to block companies from accessing the data, thereby always containing one child of sort `sys`.

By applying these sorts to any initial state modelling the starting configuration of a system and all the reaction rules, we can formally verify the same properties proven by PRISM as discussed in Section 7. For example, we ensure that data transfer occurs only if the `Contract` is valid, specifically when it includes SSCs. Without the use of these sorts, we risk the `Contract` containing unintended entities instead of SSCs, which would mean it fails to be marked as `InvalidContract` and could wrongly allow the transfer to proceed.

9 Privacy Violations Detection

Defining system rules can inadvertently lead to privacy violations. As discussed in Section 6, data must not be transferred to `FB` in `China` because the provided certification is invalid. However, modellers may define a system rule that allows data to be transferred to `FB`, as shown in Fig. 21. This violation occurs because the rule is defined without explicitly matching on `InvalidCert`. The PRISM model checker detects this violation as Property (3) in Section 7 is not met (*i.e.* the certification is invalid, but the transfer

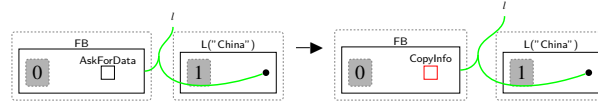


Figure 21: `privacyViolation`: system rule that leads to a privacy violation by transferring the data to FB in China .

occurs regardless). By inspecting the generated transition system, it is straightforward to identify the rule introducing the violation, *i.e.* the rule generating a state in which simultaneously `invalidCert` and `dataTransfer` hold. To correct this rule, we redefine it as presented in Fig. 19.

Malformed initial states or reaction rules can introduce privacy violations. For instance, according to the defined sorting constraints, `L(Singapore)` should contain only a single child of either sort `ctr` or `inctr`. However, if a malformed initial state or rule assigns `L(Singapore)` both sorts, this could lead to data being transferred to `L(Singapore)` during the execution of the transition system, violating the intended semantics.

Another example that highlights the vital role of sorts in ensuring the correctness of the model is when `Cont` is linked to a port of sort `sc` instead of `p`. In this scenario, the process of checking the region cannot be performed, resulting in a bug in the model. Such an issue can lead to significant consequences, especially when data transfer is critical. To maintain the integrity of the model, `Cont` must be linked to ports of sort `p` and `sy`, as specified in the sorts.

10 Related Work

There is a lack of using formal methods to prove systems' compliance with the GDPR requirements for international data transfer. However, multiple works have been proposed to prove the systems' compliance with other GDPR notions. For example, Karami *et al.* [20] define data protection language (DPL) as an object-oriented language that handles the GDPR notions of purpose and storage limitation, the right to be forgotten, and providing/withdrawing consent. They use multiset rewrite rules to formally model the operational semantics of DPL. The resulting model is checked using the model checker Maude [23], but the checking process was *partially automated* as they also provide a pen-and-paper proof. Another tool that checks these notions is the Model-Based approach to Identify Privacy Violations in software requirements (MBIPV) [36]. It is a *fully automatic* tool that detects GDPR violations by converting UML class diagrams into a Kripke model. The NuSMV [1] model checker is used to detect the violations in the defined Kripke model.

Kammüller [19] uses the theorem prover Isabelle (HOL) to prove the compliance of IoT healthcare systems with the GDPR requirements for users' right to access their data, their right to delete their data and the compliance of system functions with these rights. However, the proving process is not entirely automated.

Milner's π -calculus [24] is extended [21] by defining a set of privacy calculus syntax. Later, the privacy calculus is improved to model the GDPR notion of providing and withdrawing consent [32]. Another work based on π -calculus involves using multiparty session types to develop D'algoP tool that proves the GDPR notion of processing purposes within systems [31].

All the mentioned approaches use formal techniques to evaluate the systems' compliance with specific aspects of the GDPR. However, they are limited in verifying the GDPR requirements for international data transfers because they need to support *spatial properties* as bigraphs. [25].

Recent *informal* approaches are proposed to prove the GDPR requirements for international transfer. Guamán *et al.* [16] define a method that systematically analyses the compliance of Android mobile apps with the GDPR notion of international data transfer. Later, the method is converted to an automated tool [17] but does not handle the certification mechanisms as our model does. Pascual *et al.* [26] provide *Hunter*, an automated tool for tracking anycast communications protocol that routes the data to the nearest server regardless of location. The proposed approach mainly relies on the adequacy decision but does not cover other requirements for international data transfer, *e.g.* the use of safeguards. Unlike our approach, which *formally* proves the GDPR requirements for cross-border data transfers, these methods use *informal* techniques, potentially resulting in a less rigorous analysis of these requirements.

11 Conclusion

We provide a bigraphical framework that captures the GDPR requirements for cross-border data transfer. We encode these requirements using CTL and prove them by PRISM. We also define sorting schemes to ensure that our framework is well-formed.

Although we use only one example to define the framework, we believe it can capture different systems. The multi-perspectives approach enables us to model the GDPR requirements for international data transfers independently from specific systems. This supports the framework’s ability to model several systems. Using parameters also allows various developers to model different countries and criteria.

Using our framework requires collaboration between software engineers with a background in bi-graphs theory and privacy experts. The engineers build the model and explain the data flow to the privacy experts. The privacy experts advise on the certification criteria and the validity of the SCC, or they even amend the model if the regulations are updated. Organisations can use our framework to prove their adherence to the international transfer requirements.

In future, we aim to extend the framework to capture other safeguards, *e.g.* Binding Corporate Rules (BCRs) and Codes of conduct. We also aim to model the case of transferring data based on the user’s consent. The model need to be applied to more examples and real-world systems to show its generality. An automated tool is also needed to define and check sorting schemes automatically.

Acknowledgements

This work is partially supported by an Amazon Research Award on Automated Reasoning.

References

- [1] Alessandro Cimatti, Edmund Clarke, Fausto Giunchiglia & Marco Roveri (2000): *NUSMV: A New Symbolic Model Checker*. *International Journal on Software Tools for Technology Transfer* 2(4), pp. 410–425, doi:10.1007/s100090050046.
- [2] Ebtihal Althubiti & Michele Sevegnani (2024): *Modelling Privacy Compliance in Cross-border Data Transfers with Bigraphs (Latest Version)*, doi:10.5281/zenodo.14052899.
- [3] Blair Archibald, Muffy Calder & Michele Sevegnani (2020): *Conditional bigraphs*. In: *International Conference on Graph Transformation*, Springer, pp. 3–19, doi:10.1007/978-3-030-51372-6_1.
- [4] Blair Archibald & Michele Sevegnani (2024): *A Bigraphs Paper of Sorts*. In: *Graph Transformation*, Springer Nature Switzerland, Cham, pp. 21–38, doi:10.1007/978-3-031-64285-2_2.

- [5] Office of the Australian Information Commissioner (OAIC) (2019): *Sending personal information overseas*. Available at <https://www.t.ly/oIcI7>.
- [6] Steve Benford, Muffy Calder, Tom Rodden & Michele Sevegnani (2016): *On Lions, Impala, and Bigraphs: Modelling Interactions in Physical/Virtual Spaces*. *ACM Trans. Comput.-Hum. Interact.* 23(2), pp. 9:1–9:56, doi:10.1145/2882784.
- [7] European Data Protection Board (2023): *1.2 billion euro fine for Facebook as a result of EDPB binding decision*. Available at <https://www.t.ly/HRqTR>.
- [8] European Data Protection Board (2023): *Guidelines 07/2022 on certification as a tool for transfers*. Available at <https://www.t.ly/0PhXd>.
- [9] European Data Protection Board (n.d.): *International data transfers*. Available at <https://www.t.ly/HDX00>.
- [10] Francesca Casalini & Javier López González (2019): *Trade and cross-border data flows*. *OECD Trade Policy Papers*, doi:10.1787/b2023a47-en.
- [11] Francesca Casalini, Javier López González & Taku Nemoto (2021): *Mapping commonalities in regulatory approaches to cross-border data transfers*. *OECD Trade Policy Papers*, doi:10.1787/ca9f974e-en.
- [12] Edmund M Clarke & E Allen Emerson (1981): *Design and synthesis of synchronization skeletons using branching time temporal logic*. In: *Workshop on logic of programs*, Springer, pp. 52–71, doi:10.1007/BFb0025774.
- [13] European Commission (n.d.): *Standard Contractual Clauses (SCC)*. Available at <https://www.t.ly/7HnF6>.
- [14] Intersoft Consulting (n.d.): *Certification*. Available at <https://gdpr-info.eu/art-42-gdpr/>.
- [15] Intersoft Consulting (n.d.): *Transfers of personal data to third countries or international organisations*. Available at <https://gdpr-info.eu/chapter-5/>.
- [16] Danny S Guamán, Jose M Del Alamo & Julio C Caiza (2021): *GDPR compliance assessment for cross-border personal data transfers in android apps*. *IEEE Access* 9, pp. 15961–15982, doi:10.1109/ACCESS.2021.3053130.
- [17] Danny S Guamán, David Rodriguez, Jose M del Alamo & Jose Such (2023): *Automated GDPR compliance assessment for cross-border personal data transfers in android applications*. *Computers & Security* 130, p. 103262, doi:10.1016/j.cose.2023.103262.
- [18] Information Commissioner’s Office (ICO) (2023): *How do you determine whether you are a controller or processor?* Available at [t.ly/MPvgu](https://www.t.ly/MPvgu).
- [19] Florian Kammueeller (2018): *Formal modeling and analysis of data protection for gdpr compliance of iot healthcare systems*. In: *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, IEEE, pp. 3319–3324, doi:10.1109/SMC.2018.00562.
- [20] Farzane Karami, David Basin & Einar Broch Johnsen (2022): *DPL: A Language for GDPR Enforcement*. In: *2022 IEEE 35th Computer Security Foundations Symposium (CSF)*, IEEE, pp. 112–129, doi:10.1109/CSF54842.2022.9919687.
- [21] Dimitrios Kouzapas & Anna Philippou (2017): *Privacy by typing in the π -calculus*. *Log. Methods Comput. Sci.* 13(4), doi:10.23638/LMCS-13(4:27)2017.
- [22] M. Kwiatkowska, G. Norman & D. Parker (2011): *PRISM 4.0: Verification of Probabilistic Real-time Systems*. In G. Gopalakrishnan & S. Qadeer, editors: *Proc. 23rd International Conference on Computer Aided Verification (CAV’11)*, LNCS 6806, Springer, pp. 585–591, doi:10.1007/978-3-642-22110-1_47.
- [23] José Meseguer (2012): *Twenty years of rewriting logic*. *J. Log. Algebr. Program.* 81(7-8), pp. 721–781, doi:10.1016/j.jlap.2012.06.003.
- [24] Robin Milner (1999): *Communicating and mobile systems: the pi calculus*. Cambridge university press.
- [25] Robin Milner (2009): *The space and motion of communicating agents*. Cambridge University Press, doi:10.1017/CBO9780511626661.

- [26] Hugo Pascual, Jose M Del Alamo, David Rodriguez & Juan C Dueñas (2024): *Hunter: Tracing anycast communications to uncover cross-border personal data transfers*. *Computers & Security* 141, p. 103823, doi:10.1016/j.cose.2024.103823.
- [27] Abbas Razaghpanah, Rishab Nithyanand, Narseo Vallina-Rodriguez, Srikanth Sundaresan, Mark Allman, Christian Kreibich & Phillipa Gill (2018): *Apps, trackers, privacy, and regulators: A global study of the mobile tracking ecosystem*. In: *The 25th Annual Network and Distributed System Security Symposium (NDSS 2018)*, doi:10.14722/ndss.2018.23353.
- [28] Michele Sevegnani & Muffy Calder (2016): *BigraphER: rewriting and analysis engine for bigraphs*. In: *International Conference on Computer Aided Verification*, Springer, pp. 494–501, doi:10.1007/978-3-319-41540-6_27.
- [29] Michele Sevegnani, Milan Kabác, Muffy Calder & Julie A. McCann (2018): *Modelling and Verification of Large-Scale Sensor Network Infrastructures*. In: *23rd International Conference on Engineering of Complex Computer Systems, ICECCS 2018, Melbourne, Australia, December 12-14, 2018*, pp. 71–81, doi:10.1109/ICECCS2018.2018.00016.
- [30] European Data Protection Supervisor (n.d.): *International transfers*. Available at <https://www.t.ly/GPuii>.
- [31] Evangelia Vanezi, Georgia M Kapitsaki, Dimitrios Kouzapas, Anna Philippou & George A Papadopoulos (2020): *DiálogoP-A Language and a Graphical Tool for Formally Defining GDPR Purposes*. In: *International Conference on Research Challenges in Information Science*, Springer, pp. 569–575, doi:10.1007/978-3-030-50316-1_40.
- [32] Evangelia Vanezi, Dimitrios Kouzapas, Georgia M Kapitsaki & Anna Philippou (2019): *Towards GDPR Compliant Software Design: A Formal Framework for Analyzing System Models*. In: *International Conference on Evaluation of Novel Approaches to Software Engineering*, Springer, pp. 135–162, doi:10.1007/978-3-030-40223-5_7.
- [33] WhatsApp (2024): *WhatsApp Privacy Policy*. Available at <https://www.t.ly/uGjmN>.
- [34] WhatsApp (2024): *Who is providing your WhatsApp services*. Available at <https://www.t.ly/NqTfZ>.
- [35] Jim Woodcock, Peter Gorm Larsen, Juan Bicarregui & John Fitzgerald (2009): *Formal methods: Practice and experience*. *ACM computing surveys (CSUR)* 41(4), pp. 1–36, doi:10.1145/1592434.1592436.
- [36] Tong Ye, Yi Zhuang & Gongzhe Qiao (2023): *MBIPV: a model-based approach for identifying privacy violations from software requirements*. *Software and Systems Modeling* 22(4), pp. 1251–1280, doi:10.1007/s10270-022-01072-3.

A System Rules

We present some specific system rules for the WhatsApp example:

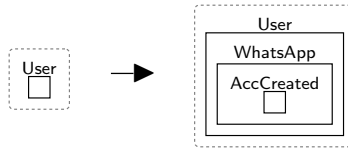


Figure 22: `creatingAccount`: the system's user starts using WhatsApp to create an account.

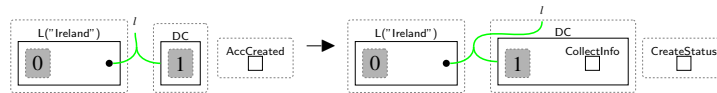


Figure 23: `creatingStatus`: the user creates a status, so the entity `AccCreated` is replaced with `CreateStatus`. We assume that DC collects information about the user's status, *e.g.* status content type, location information, information about the device used to post the status *etc.* We model that by generating the entity `CollectInfo`. We do not specify exactly what the collected data is, as our model focuses on checking the adequacy decision and the safeguards regardless of the type of transferred data.

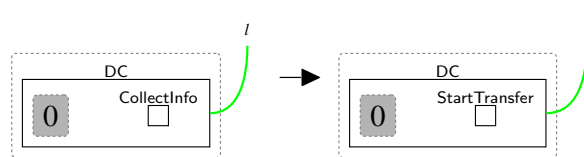


Figure 24: `initialisingTrnas`: DC initialises the transferring process by replacing the entity `CollectInfo` with `StartTransfer` to start applying rule `dcIrType` in Fig. 14

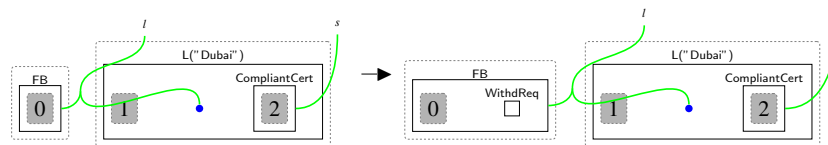


Figure 25: `ReqWithdCert`: FB in Dubai asks to withdraw its certification. The rule explicitly matches on the `CompliantCert` as we need to ensure that the entity requesting the withdrawal is using the certification, not the SCCs.