Information Security as Strategic (In)effectivity

Wojciech Jamroga[1] and Masoud Tabatabaei[2]

[1] Institute of Computer Science, Polish Academy of Sciences
[2] Interdisciplinary Centre for Security and Trust, University of Luxembourg
w.jamroga@ipipan.waw.pl, masoud.tabatabaei@uni.lu

Abstract. Security of information flow is commonly understood as preventing *any* information leakage, regardless of how grave or harmless consequences the leakage can have. In this work, we suggest that information security is not a goal in itself, but rather a means of preventing potential attackers from compromising the correct behavior of the system. To formalize this, we first show how two information flows can be compared by looking at the adversary's ability to harm the system. Then, we propose that the information flow in a system is *effectively information-secure* if it does not allow for more harm than its idealized variant based on the classical notion of noninterference.

1 Introduction

In most approaches to information flow security, information defines the ultimate goal of the interaction between agents. Classical information security properties specify *what* information must not leak, and *how* it could possibly leak (i.e., what channels of information leakage are considered), but they do not give account of *why* the information should not leak to the intruder. For example, the property of *noninterference* [7] assumes that the "low clearance" users cannot learn anything about the activities of the "high clearance" users. In order to violate this, the "low" users can try to analyse their observations and/or execute a sequence of explorative actions of their own. *Nondeducibility on strategies* [30] makes the same assumption about *what* should not leak, but takes also into account covert channels that some "high" users can use to send signals to the "low" agents according to a previously agreed code. *Anonymity* in voting [2, 5] captures that an observer cannot learn what candidate a particular has voted for by looking at the voter's behavior, scanning the web bulletin board, coercing the voter to hand in the vote receipt, etc.

As a consequence, the classical properties of information security can only distinguish between relevant and irrelevant information leaks if the distinction is given explicitly as a parameter, e.g., by classifying available actions into sensitive and insensitive [7]. However, it is usually hard (if not impossible) to obtain such a distinction based on the internal characteristics of the actions. We illustrate the point below by means of a real-life example.

Example 1. In some phone banking services, the maiden name of the user's mother is used as a part of authentication.[3] Consider now a user posting an essay about some ancestor of hers on her blog, mentioning also the name of the ancestor. If the essay is about the user's mother, it reveals potentially dangerous information. On the other hand, if the post is about some other member of the user's family (father, grandmother, paternal grandfather, etc.) revealing the name of the person is probably harmless. Note that it is impossible to distinguish between the two pieces of information (say, the mother's maiden name vs. the grandmother's maiden name) based on their internal features. The only difference lies in the context: the first kind of information is used in some important social procedures, while the second one is not. □

In this paper, we claim that a broader perspective is needed to appropriately model and analyse such scenarios. Agents compete for information not for its own sake, but for reasons that go beyond purely epistemic advantages. More precisely, information is a commodity that the players compete for in an "information security game" but the game is played in the context of a "real" game where information is only a resource, enabling (some) players to achieve their non-epistemic goals. As players obtain new information, their uncertainty is reduced, and they increase their ability to choose a good strategy in the real game.

What would a *significant information leak* be in this view? To answer the question, we draw inspiration from the concept of the *value of information*: a piece of information is worth as much as it increases the expected payoff of the player. Similarly, an information leak is significant if it increases the ability of the attacker to construct a damaging attack strategy in the real game.

Contribution of the Paper. First, we use the concept of *surely winning strategies* from game theory to analyze the adversary's strategic ability to disrupt the correct behavior of the system. We will see the *effective security* of the system as the attacker's *in*ability to come up with such a strategy.

Secondly, we use the notion of *effective security* for comparing two systems by looking at the strategic ability of an adversary to harm the goal of the system.

Thirdly, a successful attack strategy can exist due to flawed design of either the control flow or the information flow in the system. Here, we are interested in the latter. That is, we want to distinguish between vulnerabilities coming from the control vs. the information flow, and single out systems where redesigning the flow of information alone can make the system more secure. To this end, we define the noninterferent idealized variant of the system, which has the same control flow as the original system, but with the information reduced so that the system satisfies noninterference. Then, we define the system to be *effectively information-secure* if it is as good as its noninterfering idealized variant. As the main technical result, we show that the concept is well defined, i.e., the maximal noninterferent variant exists for every state-transition model.

[3] This is a real-life example from the authors' personal experience. For similar security questions, used by various phone or web services, cf. e.g. [14].

Due to lack of space, we include only proof sketches for most results. The complete proofs, together with additional examples, can be found in the extended version of the paper, available at http://arxiv.org/abs/1608.02247 .

2 Related Work

Various formalizations of information flow security have been proposed and studied. The classical concept here is *noninterference* [7] and its variations: *nondeducibility* [28], *noninference* [22], *restrictiveness* [18], *nondeducibility on strategies* [30], and *strategic noninterference* [13]. Probabilistic noninterference and quantitative noninterference have been investigated, e.g., in [9, 30, 19, 23, 16, 27]. All the above concepts assume that the information flow in the system is secure only when no information ever flows from High to Low players. In this paper, we want to discard irrelevant information leaks, and only look at the significant ones (in the sense that the leaking information can be used to construct an attack on a higher-order correctness property).

The problem of how to weaken noninterference to successfully capture security guarantees for real systems has been also extensively studied. Most notably, postulates and policies for *declassification* (called also *information release*) were studied, cf. [26] for an introduction. This submission can be viewed as an attempt to determine *what information is acceptable to declassify*. In this sense, our results can useful in proposing new declassification policies and evaluating existing ones. We note, however, that the existing work on declassification are mainly concerned by the question *what* information can be released, *when*, *where*, and by *whom*. In contrast, we propose an argument for *why* it can be released. Moreover, declassification is typically about intentional release of information, whereas we do not distinguish between intentional and accidental information flow. Finally, the research on declassification assumes that security is defined by some given "secrets" to be protected. In our approach, no information is intrinsically secret, but the information flow is harmful if it enables the attacker to gain more strategic ability against the goals of the system.

Parameterized noninterference [6] can be seen as a theoretical counterpart of declassification, where security of information flow is parameterized by the analytic capabilities of the attacker. Again, that research does not answer why some information must be kept secret while some other needs not, and in particular it does not take strategic power of the attacker into account.

Economic and strategic analysis of security properties is a growing field in general, cf. [21] for an introduction. A number of papers have applied game-theoretic concepts to define the security of information flow [17, 11, 12, 3, 4, 13]. However, most of those papers [17, 11, 12, 3] use games only in a narrow mathematical sense to provide a proof system (called the *game semantics*) for deciding security properties. We are aware of only a handful of papers that investigate the impact of participants' incentives and available strategies on the security of information flow. In [1, 10], economic interpretations of privacy-preserving behavior are proposed. [4] uses game-theoretic solution concepts (in particular,

Nash equilibrium) to prescribe the optimal defense strategy against attacks on information security. In contrast, our approach is analytic rather than prescriptive, as we do not propose how to manage information security. Moreover, in our view, privacy is not the goal but rather the means to achieve some higher-level objectives. Finally, [13] proposes a weaker variant of noninterference by allowing the High players to select an appropriate strategy, while here we look at the potential damage inflicted by adverse strategies of the Low users.

Our idea of looking at the unique most precise non-interfering variant of the system is related on the technical level to [6]. There, attackers displaying different analytical capabilities are defined by abstract interpretation, which leads to a lattice of noninterference variants with various strength. Attackers with weakened observational powers were also studied in [31].

3 Preliminaries

3.1 Simple Models of Interaction

Since we build our proposal around the standard notion of noninterference by Goguen and Meseguer [7], we will use similar models to represent interaction between actions of different agents. The *system* is modeled by a multi-agent asynchronous transition network $M = \langle St, s_0, \mathfrak{U}, \mathfrak{A}, Obs, obs, do \rangle$ where: St is the set of *states*, s_0 is the initial state, $\mathfrak{U}$ is the set of *agents* (or *users*), $\mathfrak{A}$ is the set of *actions*, Obs is the set of possible *observations* (or *outputs*); $obs : St \times \mathfrak{U} \to Obs$ is the observation function. $do : St \times \mathfrak{U} \times \mathfrak{A} \to St$ is the transition function that specifies the (deterministic) outcome $do(s, u, a)$ of action a executed by user u in state s. We will sometimes write $[s]_u$ instead of $obs(s, u)$. Also, we will call a pair *(user, action)* a *personalized action*. We construct the multi-step transition function $exec : St \times (\mathfrak{U} \times \mathfrak{A})^* \to St$ so that, for a finite string $\alpha \in (\mathfrak{U} \times \mathfrak{A})^*$ of personalized actions, $exec(s, \alpha)$ denotes the state resulting from execution of α from s on. We may sometimes write $s \xrightarrow{\alpha} t$ instead of $exec(s, \alpha) = t$, and $exec(\alpha)$ instead of $exec(s_0, \alpha)$.

Three remarks are in order. First, Goguen and Meseguer's models define agents' observations based on states only, whereas it is often convenient to also model the information flow due to observing each others' actions. Secondly, the models are fully asynchronous in the sense that if each user "submits" a sequence of actions to be executed then every interleaving of the submitted sequences can occur as the resulting behavior of the system. No synchronization is possible. Thirdly, the models are "total on input" (each action label is available to every user at every state), and hence no synchronization mechanism can be encoded via availability of actions. Especially the last two features imply that models of Goguen and Meseguer allow for representation of a very limited class of systems.

We start by using the purely asynchronous models of Goguen and Meseguer. Then, in Section 6, we extend our results to a broader class of models by allowing partial transition functions.

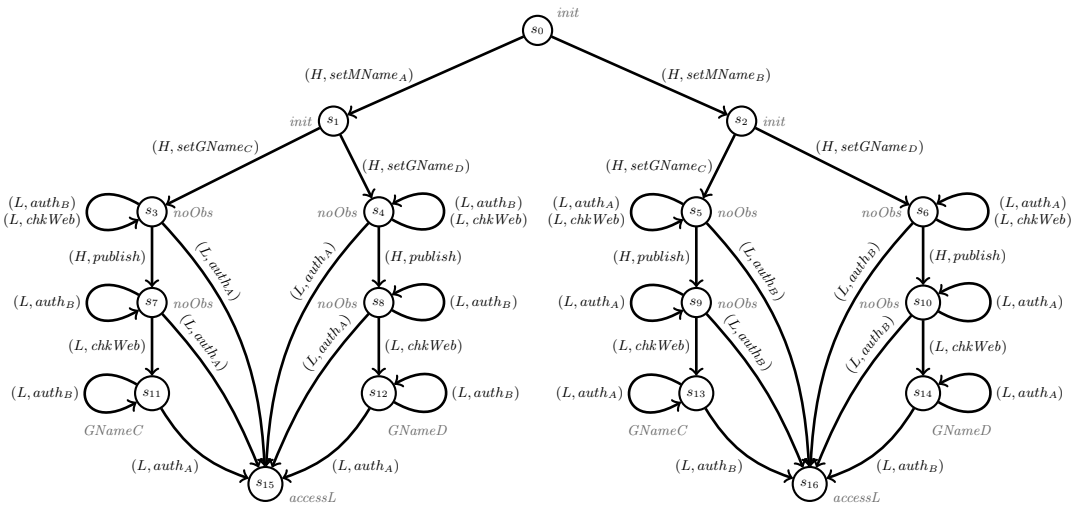

Fig. 1. Transition network M_a in which the High player publishes her grandmother's maiden name on her blog. Only the observations of L are shown

3.2 Noninterference

We now recall the standard notion of noninterference from [7]. Let $U \subseteq \mathfrak{U}$ and $\alpha \in (\mathfrak{U} \times \mathfrak{A})^*$. By $Purge_U(\alpha)$ we mean the subsequence of α obtained by eliminating all the pairs (u, a) with $u \in U$.

Definition 1 (Noninterference [7]). *Let M be a transition network with sets of "high clearance" agents H and "low clearance" agents L, such that $H \cap L = \emptyset, H \cup L = \mathfrak{U}$. We say that H* is non-interfering with *L iff for all $\alpha \in (\mathfrak{U} \times \mathfrak{A})^*$ and all $u_l \in L$, $[exec(\alpha)]_{u_l} = [exec(Purge_H(\alpha))]_{u_l}$. We denote the property by $NI_M(H, L)$.*

Thus, $NI_M(H, L)$ expresses that L can neither observe nor deduce what actions of H have been executed

Example 2. Consider a simplified version of the phone banking scenario from Example 1. There are two users: H who has an account in the bank, and L who may try to impersonate H. H can access her account by correctly giving the maiden name of her mother. Moreover, H runs a blog, and can publish some of her personal information on it. We consider two alternative variants: one where H publishes her grandmother's maiden name on the blog (Figure 1), and one where she publishes her mother's maiden name (Figure 2). We assume that the possible names are A and B in the former case, and C and D in the latter.Each model begins by initialization of the relevant names. The observations of L are shown beside each state. The observations for H are omitted, as they will be irrelevant for our analysis.

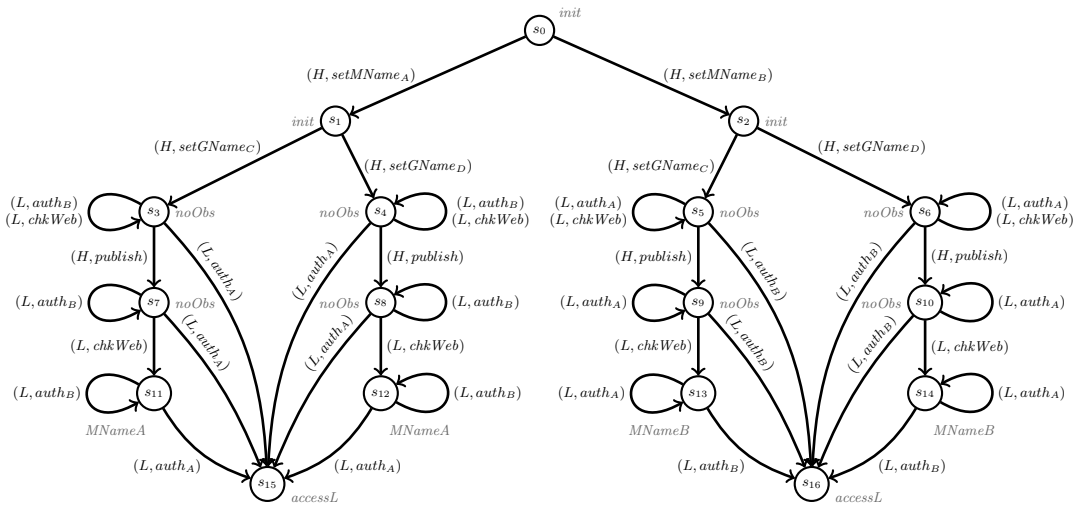

Fig. 2. Transition network M_b in which H publishes her mother's maiden name

Note that, for mathematical completeness, we must define the outcome of every user-action pair in every state. We assume that there are two "error states" s_{HErr}, s_{LErr} in models M_a and M_b (not shown in the graphs). Any action of H not depicted in the figure leads to s_{HErr}, and any action of L not depicted in the figure leads to s_{LErr}. We will later use the error states in the definition of the players' goals, in such a way that L will always want to avoid s_{LErr} and H will want to avoid s_{HErr}. This way we can (however imperfectly) simulate some synchronization in the restricted framework of Goguen and Meseguer.

Neither M_a nor M_b satisfies noninterference from H to L. For instance, in the model of Figure 1, if $\alpha = \langle (H, setMName_A), (H, setGName_D), (H, publish), (L, chkWeb)\rangle$, the observation of L after sequence α is $GNameD$, but the observation of L after $Purge_H(\alpha) = \langle (L, chkWeb)\rangle$ is $noObs$, which is clearly different. □

3.3 Strategies and Their Outcomes

Strategy is a game-theoretic concept which captures behavioral policies that an agent can consciously follow in order to realize some objective [15]. Let $T(M)$ be the *tree unfolding* of M. Also if $U \subseteq \mathfrak{U}$ is a subset of agents, let T' be a *U-trimming* of tree T iff T' is a subtree of T starting from the same root and obtained by removing an arbitrary subset of transitions labeled by actions of agents from U. For the moment, we assume that each subset of agents $U \subseteq \mathfrak{U}$ is assigned a set of available coalitional strategies Σ_U. The most important feature of a strategy $\sigma_U \in \Sigma_U$ is that *it constrains the possible behaviors of the system.* We represent it formally by the *outcome function* $out_M(\sigma_U)$ that removes the executions of the system that strategy σ_U would never choose. Therefore, for every $\sigma_U \in \Sigma_U$, its outcome $out_M(\sigma_U)$ is a U-trimming of $T(M)$.

Let h be a node in tree T corresponding to a particular finite history of interaction. We denote the sequence of personalized actions leading to h by $act^*(h)$. Furthermore, $act^*(T) = \{act^*(h) \mid h \in nodes(T)\}$ is the set of finite sequences of personalized actions that can occur in T.

Strategies are usually constructed as mappings from possible situations that the player can recognize in the game, to actions of the player (or subsets of actions if we allow for nondeterministic strategies). Formally, the set of *perfect recall strategies* of agent u is $\Sigma_u^{\mathfrak{Rec}} = \{\sigma_u : nodes(T(M)) \to \mathcal{P}(\mathfrak{A}) \setminus \{\emptyset\} \mid obs_u(h) = obs_u(h') \Rightarrow \sigma_u(h) = \sigma_u(h')\}$, where $obs_u(h)$ denotes the accumulate observations collected by agent u along history h. How to define obs_u for sequences of states? For asynchronous systems, this is typically defined as $obs_u(q) = [q]_u$, $obs_u(h \circ q) = obs_u(h)$ if $last(h) = q$, and $obs_u(h \circ q) = obs_u(h) \circ [q]_u$ otherwise (where $\circ$ denotes the concatenation operator). That is, what u has learned along h is equivalent to the sequence of observations she has seen, modulo removal of "stuttering" observations. Now, coalitional strategies of perfect recall for a group of agents $U \subseteq \mathfrak{A}$ are combinations of individual strategies, i.e., $\Sigma_U^{\mathfrak{Rec}} = \times_{u \in U}(\Sigma_u^{\mathfrak{Rec}})$. The outcome of $\sigma_U \in \Sigma_U^{\mathfrak{Rec}}$ in model M is the tree obtained from $T(M)$ by removing all the branches that begin from a node h with a personalized action $(u, a) \in U \times \mathfrak{A}$ such that $a \notin \sigma_U(h)$.

3.4 Temporal Goals and Winning Strategies

A goal is a property that some agents may attempt to enforce by selecting their behavior accordingly. We base our approach on the concepts of *paths* and *path properties*, used in temporal specification and verification of systems [20]. Let $paths(M)$ denote the set of infinite sequences of states that can be obtained by subsequent transitions in M. Additionally, we will use $paths_M(\sigma)$ as a shorthand for $paths(out_M(\sigma))$.

Definition 2 (Temporal goal [20]). *A* goal *in M is any $\Gamma \subseteq paths(M)$. Note that $paths(M) = paths(T(M))$, so a goal can be equivalently seen as a subset of paths in the tree unfolding of M.*

Most common examples of such goals are safety and reachability goals.

Definition 3 (Safety and reachability goals [20]). *Given a set of safe states $\mathbb{S} \subseteq St$, the* safety goal *$\Gamma_{\mathbb{S}}$ is defined as $\Gamma_{\mathbb{S}} = \{\lambda \in paths(M) \mid \forall i.\lambda[i] \in \mathbb{S}\}$. Moreover, given a set of target states $\mathbb{T} \subseteq St$, the* reachability goal *$\Gamma_{\mathbb{T}}$ can be defined as $\Gamma_{\mathbb{T}} = \{\lambda \in paths(M) \mid \exists i.\lambda[i] \in \mathbb{T}\}$.*

Definition 4 (Winning strategies). *Given a transition network M, a set of agents $U \subseteq \mathfrak{A}$ with goal Γ_U, and a set of strategies $\Sigma_U^{\mathfrak{Rec}}$, we say that U have a (surely winning) strategy to achieve Γ_U iff there exists a strategy $\sigma_U \in \Sigma_U^{\mathfrak{Rec}}$ such that $paths_M(\sigma_U) \subseteq \Gamma_U$.*

Example 3. Consider the models in Figure 1 and Figure 2, and suppose that L wants to access H's bank account. This can be expressed by the reachability goal

$\Gamma_{\mathbb{T}}$ with $\mathbb{T} = \{s_{15}, s_{16}\}$ as the target states. In fact, L also wins if H executes an out-of-place action (cf. Example 2 for detailed explanation). In consequence, the winning states for L are $\mathbb{T} = \{s_{15}, s_{16}, s_{HErr}\}$. Note that L has no strategy that guarantees $\Gamma_{\mathbb{T}}$ in model M_a (although information is theoretically leaking to L as the model does not satisfy noninterference). Even performing the action *chkWeb* does not help, because L cannot distinguish between states s_{11} and s_{13}, and there is no single action that succeeds for both s_{11}, s_{13}. Thus, L does not know whether to use $auth_A$ or $auth_B$ to get access to H's bank account.

On the other hand, L has a winning strategy for $\Gamma_{\mathbb{T}}$ in model M_b. The strategy is to execute *chkWeb* after H publishes her mother's maiden name, and afterwards do $auth_A$ in states s_{11}, s_{12} (after observing *MNameA*) or $auth_B$ if the system gets to s_{13}, s_{14} (i.e., after observing *MNameB*). □

4 Security as Strategic Property

The property of noninterference looks for *any* leakage of *any* information. If one can possibly happen in the system, then the system is deemed insecure. In many cases, this view is too strong. There are lots of information pieces that can leak out without bothering any interested party. Revealing the password to your web banking account can clearly have much more disastrous effects than revealing the price that you paid for metro tickets on your latest trip to Paris. Moreover, the relevance of an information leak cannot in general be determined by the type of the information. Think, again, of revealing the maiden name of your mother vs. the maiden name of your grandmother. The former case is potentially dangerous since the maiden name of one's mother is often used to grant access to manage banking services by telephone. Revealing the latter is quite harmless to most ends and purposes.

In this paper, we suggest that the relevance of information leakage should be judged by the extent of damage that the leak allows the attackers to inflict on the goal of the system. Thus, as the first step, we define the security of the system in terms of damaging abilities of the Low players.

In order to assess the relevance of information flow from High to Low, we will look at the resulting strategic abilities of Low. Moreover, we assume that the goal of L is to violate a given goal of the system. The goal can be a functionality or a security requirement, or a combination of both. Moreover, it can originate from a private goal of the High players, an objective ascribed to the system by its designer (e.g., the designer of a contract signing protocol), or a combination of requirements specified by the owner/main stakeholder in the system (for instance, a bank in case of a web banking infrastructure).

Definition 5 (Effective security). *Let M be a transition network with some Low players $L \subseteq \mathfrak{U}$, and let Γ be the goal of the system. We say that M is* effectively secure *for (L, Γ) iff L does not have a strategy to enforce $\overline{\Gamma}$, where $\overline{X}$ denotes the complement of set X. That is, the system is effectively secure iff the attackers do not have a strategy that ensures an execution violating the goal of the system. We will use $ES(M, L, \Gamma)$ to refer to this property.*

Besides judging the effective security of a system, we can also use the concept to compare the security level of two models.

Definition 6 (Comparative effective security). *Let M, M' be two models, and Γ be a goal in M, M' (i.e., $\Gamma \subseteq paths(M) \cup paths(M')$). We say that:*

- *M has* strictly less effective security *than M' for (L, Γ), denoted $M \prec_{L,\Gamma} M'$, iff $ES(M', L, \Gamma)$ but not $ES(M, L, \Gamma)$.*
- *M' is* at least as effectively secure as *M for (L, Γ), denoted $M \preceq_{L,\Gamma} M'$, iff $ES(M, L, \Gamma)$ implies $ES(M', L, \Gamma)$;*
- *M is* effectively equivalent to *M' for (L, Γ), denoted $M \simeq_{L,\Gamma} M'$, iff either both $ES(M, L, \Gamma)$ and $ES(M', L, \Gamma)$ hold, or both do not hold.*

Thus, if in one of the models L can construct a more harmful strategy then the model displays lower effective security than the other model. Conversely, if both models allow only for the same extent of damage then they have the same level of effective security. This way, we can order different alternative designs of the system according to the strategic power they give away to the attacker.

Example 4. Consider models M_a, M_b from Figure 1 and Figure 2, and let the goal Γ be to prevent L from accessing H's bank account. Thus, Γ is the safety goal $\Gamma_{\mathbb{S}}$ with $\mathbb{S} = St \setminus \{s_{15}, s_{16}, s_{HErr}\}$, and therefore $\overline{\Gamma} = \Gamma_{\mathbb{T}}$ with $\mathbb{T} = \{s_{15}, s_{16}, s_{HErr}\}$. As we saw in Example 3, L has no strategy to guarantee $\overline{\Gamma}$ in M_a, but she has a surely winning strategy for $\overline{\Gamma}$ in M_b. Thus, M_b is strictly less effectively secure than M_a, i.e., $M_b \prec_{L,\Gamma} M_a$.

5 Effective Information Security

We will now propose a scheme that allows to determine whether a given model of interaction leaks relevant information or not. We use the idea of refinement checking from process algebras, where a process is assumed correct if and only if it refines the ideal process [25]. A similar reasoning scheme is also used in analysis of multi-party computation protocols (a protocol is correct iff it is equivalent to the ideal model of the computation [8]).

5.1 Ability-Based Security of Information Flows

Definition 6 allows for comparing the effective security of two alternative information flows. However, we usually do not want to compare several alternative information flows. Rather, we want to determine if a single given model M reveals relevant information or not. A natural idea is to compare the effective security of M to an *ideal model*, i.e. a variant of M that leaks no relevant information by construction. Then, a model is effectively information-secure if it has the same level of effective security as its idealized variant:

Definition 7 (Effective information security). *Let M be a transition network with some Low players $L \subseteq \mathfrak{U}$, and let Γ be the goal of the system. Moreover, let $Ideal(M)$ be the idealized variant of M. We say that M is* effectively information-secure *for (L, Γ) iff $M \simeq_{L,\Gamma} Ideal(M)$.*

How do we construct the idealized variant of M? The idea is to "blur" observations of Low so that we obtain a variant of the system where the observational capabilities of the attackers are minimal.

5.2 Idealized Models Based on Noninterference

We begin by recalling the notion of *term unification* which is a fundamental concept in automated theorem proving and logic programming [24]. Given two terms t_1, t_2, their unification $(t_1 \equiv t_2)$ can be understood as a declaration that, from now on, both terms refer to exactly the same underlying object. In our case the terms are observation labels from the set Obs. A unification can be seen as an equivalence relation on observation labels, or equivalently as a partitioning of the labels into equivalence classes. The application of the unification to a model yields a similar model where the equivalent observations are "blurred".

Definition 8 (Unification of observations). *Given a set of observation labels Obs, a* unification on Obs *is any equivalence relation $\mathcal{U} \subseteq Obs \times Obs$.*

Given a model $M = \langle St, s_0, \mathfrak{U}, \mathfrak{A}, do, Obs, obs \rangle$ and a unification $\mathcal{U} \subseteq Obs \times Obs$, the application of $\mathcal{U}$ to M *is the model $\mathcal{U}(M) = \langle St, s_0, \mathfrak{U}, \mathfrak{A}, do, Obs', obs' \rangle$, where: $Obs' = \{[o]_{\mathcal{U}} \mid o \in Obs\}$ replaces Obs by the set of equivalence classes defined by $\mathcal{U}$, and $obs'(q, u) = [obs(q, u)]_{\mathcal{U}}$ replaces the original observation in q with its equivalence class for any $u \in \mathfrak{U}$.*

Our reference model for M will be the variant of M where noninterference is obtained by the minimal necessary "blurring" of L's observations.

Definition 9 (Noninterferent idealized model). *Having a transition network M and a set of "low" players L, we define the* noninterfering idealized variant of M *as $\mathcal{U}(M)$ such that:*

(i) $NI_{\mathcal{U}(M)}(H, L)$, and
(ii) for every $\mathcal{U}' \subsetneq \mathcal{U}$ it is not the case that $NI_{\mathcal{U}'(M)}(H, L)$.

We need to show that the concept of noninterferent idealized model is well defined. The proof is constructive, i.e., given a model M, we first show how one can build its idealized variant, and then show that it is unique.[4]

Theorem 1. *For every transition network M, there is always a unique unification $\mathcal{U}$ satisfying properties (i) and (ii) from Definition 9.*

[4] We will only sketch the proofs due to lack of space. The complete proofs can be found in the extended version of the paper, available at http://arxiv.org/abs/1608.02247 .

The proof of Theorem 1 needs some preliminary steps. First, we recall the concept of unwinding relations [29]. Unwinding relations are important because they characterize noninterference in purely structural terms. Moreover, existence of an unwinding relation is usually easier to verify than proving noninterference directly. We then use the concept of unwinding relation to define relation R^*_M on the states of a transition network M. We use this relation to construct and prove the uniqueness of the idealized variant of M.

Definition 10 (Unwinding for Noninterference [29]). $\sim_{NI_L} \subseteq St \times St$ *is an* unwinding relation *iff it is an equivalence relation satisfying the conditions of* output consistency (OC)*,* step consistency (SC)*, and* local respect (LR)*. That is, for all states* $s, t \in St$*:*

(OC) *If* $s \sim_{NI_L} t$ *then* $[s]_L = [t]_L$*;*
(SC) *If* $s \sim_{NI_L} t$*,* $u \in L$*, and* $a \in \mathfrak{A}$ *then* $do(s,u,a) \sim_{NI_L} do(t,u,a)$*;*
(LR) *If* $u \in H$ *and* $a \in \mathfrak{A}$ *then* $s \sim_{NI_L} do(s,u,a)$*.*

Proposition 1 ([29]). $NI_M(H,L)$ *iff there exist an unwinding relation* $\sim_{NI_L}$ *on the states of* M *that satisfies (OC), (SC) and (LR).*

Next we define R^*_M on the states of a transition network M. The definition goes as follows: first we relate any two states of M' if one of them can be reached from the other one by a sequence of High personalized actions. Then in each step we relate the pair of states that are reached by a similar Low personalized action from any two states that are already related. Also, we enforce transitivity on the set. We continue adding related states until the relation becomes stable. The mathematical definition of R^*_M is as follows:

Definition 11 (Relation R^*_M for a transition network M). *Given a model* $M = \langle St, s_0, \mathfrak{U}, \mathfrak{A}, do, Obs, obs\rangle$ *and sets of High players* H *and Low players* L*, we define the relation* $R^*_M \subseteq St \times St$ *as the least fixpoint of the following function* F*, transforming relations on* St*:*

$$F(R) = R_0 \cup \{(t_1,t_2) \mid \exists (s_1,s_2) \in R, l \in L, a \in \mathfrak{A}. do(s_1,l,a) = t_1, do(s_2,l,a) = t_2\} \cup \{(t_1,t_2) \mid \exists s \in St.(t_1,s) \in R \& (s,t_2) \in R\},$$

where $(s_1,s_2) \in R_0$ *iff for some sequence of personalized actions of High players* α*, either* $s_1 \xrightarrow{\alpha} s_2$*, or* $s_2 \xrightarrow{\alpha} s_1$*.*

It can be shown that it is sufficient to unify Low's observations in states connected by R^*_M in order to obtain a non-interferent model. In consequence, R^*_M generates the minimal unification that achieves the task. Now, by using relation R^*_M, we define the unification of observations $\mathcal{U}^*_M$ that will provide the noninterferent idealized variant of M.

Definition 12 (Unification for noninterference $\mathcal{U}^*_M$). *We define the unification of observations* $\mathcal{U}^*_M \subseteq Obs \times Obs$ *as follows. For any* $o_1, o_2 \in Obs$*, we have* $(o_1,o_2) \in \mathcal{U}^*_M$ *iff there exist* $s_1, s_2, t_1, t_2 \in St$ *and* $l \in L$ *such that:*
(a) $obs(s_1,l) = o_1$*, (b)* $obs(s_2,l) = o_2$*, (c)* $(s_1,t_1) \in R^*_M$*, (d)* $(s_2,t_2) \in R^*_M$*, and (e)* $obs(t_1,l) = obs(t_2,l)$*.*

It then holds that $\mathcal{U}^*_M(M)$ satisfies the noninterference property (Proposition 2) and no refinement of $\mathcal{U}^*_M$ achieves that (Proposition 3).

Proposition 2. *Given a model M, and $\mathcal{U}^*_M(M) = \langle St, s_0, \mathfrak{U}, \mathfrak{A}, do, Obs^*, obs^* \rangle$ defined as in Definition 12 on M, it holds that $NI_{\mathcal{U}^*_M(M)}(H, L)$.*

Proposition 3. *Given a model M, and sets of players H and L, for any unification of observations U where $U(M) = \langle St, s_0, \mathfrak{U}, \mathfrak{A}, do, Obs', obs' \rangle$, if $NI_{U(M)}(H, L)$ then $\mathcal{U}^*_M \subseteq U$.*

We can now complete the proof of Theorem 1.

Proof (of Theorem 1). We want to prove that, given a model M, set of players H and L, and any unification of observations $\mathcal{U}$, if $\mathcal{U}(M)$ is a noninterfering idealized variant of M, then $\mathcal{U} = \mathcal{U}^*_M$. Assume that $\mathcal{U}(M)$ is a noninterfering idealized variant of M. By property (i) of Definition 9 and Proposition 3 we infer that $\mathcal{U}^*_M \subseteq \mathcal{U}$. Also, by Proposition 2, we have that $NI_{\mathcal{U}^*_M(M)}(H, L)$. Therefore by property (ii) of Definition 9 it holds that $\mathcal{U} = \mathcal{U}^*_M$.

Example 5. Consider models M_a, M_b in Figure 1 and Figure 1. We recall that both models are not noninterferent. In the noninterferent idealized variant of M_a, observations $noObs$, $MnameC$, and $MNameD$ of L are unified and replaced by the equivalence class $\{noObs, MNameD, MNameD\}$. The idealized variant of M_b is constructed analogously by unification of $noObs$, $MnameA$, and $MNameB$. Clearly, L has no surely winning strategy to guarantee $\overline{\Gamma} = \Gamma_{\mathbb{T}}$ for $\mathbb{T} = \{s_{15}, s_{16}, s_{HErr}\}$ in both $Ideal(M_a)$ and $Ideal(M_b)$.

Recall from Example 4 that L has no winning strategy for $\overline{\Gamma}$ in M_a, but she has one in M_b. So, $M_a \simeq_{L,\Gamma} Ideal(M_a)$, but $M_b \not\simeq_{L,\Gamma} Ideal(M_b)$. Thus, M_a is effectively information-secure for (L, Γ), but M_b is not. □

It is important to notice that noninterferent variants are indeed idealizations:

Proposition 4. *For every M, L, and Γ, we have that $M \preceq_{L,\Gamma} Ideal(M)$.*

Proof. Note that because M and $Ideal(M)$ differ only in their observation functions. Also we have that for any pair of states $s_1, s_2 \in St$, if $[s_1]^M_L = [s_2]^M_L$ then $[s_1]^{Ideal(M)}_L = [s_2]^{Ideal(M)}_L$. Therefore all the strategies of L in $Ideal(M)$ are also L's strategies in M. Thus for any for any goal $\Gamma \subseteq paths(M)$, if L have a surely winning strategy to enforce $\overline{\Gamma}$ in $Ideal(M)$ then they also have a surely winning strategy for $\overline{\Gamma}$ in M, □

.

6 Extending the Results to a Broader Class of Models

As mentioned before, the models of Goguen and Meseguer are "total on input," i.e., each action label is available to every user at every state. This makes modeling actual systems very cumbersome. In this section, we consider a broader class

of models, and show how our results carry over to the more expressive setting. That is, we consider *partial transition networks (PTS)* $M = \langle St, s_0, \mathfrak{U}, \mathfrak{A}, Obs, obs, do \rangle$ which are defined as in Section 3.1, except that the transition function $do : St \times \mathfrak{U} \times \mathfrak{A} \rightharpoonup St$ can be a partial function. By $do(s, u, a) = undef$ we denote that action a is unavailable to user u in state s; additionally, we define $act(s, u) = \{a \in \mathfrak{A} \mid do(s, u, a) \neq undef\}$ as the set of actions available to u in s. Moreover, we assume that players are aware of their available actions, and hence can distinguish states with different repertoires of choices – formally, for any $u \in \mathfrak{U}, s_1, s_2 \in St$, if $obs(s_1, u) = obs(s_2, u)$ then $act(s_1, u) = act(s_2, u)$.

We begin by a suitable update of the definition of noninterference:

Definition 13 (Noninterference for partial transition networks). *Given a PTS M and sets of agents H, L, such that $H \cup L = \mathfrak{U}, H \cap L = \emptyset$, we say that H* is non-interfering with *L iff for all $\alpha \in (\mathfrak{U} \times \mathfrak{A})^*$ and all $u_l \in L$, if $exec(\alpha) \neq undef$ then $[exec(\alpha)]_{u_l} = [exec(Purge_H(\alpha))]_{u_l}$. We denote the property also by $NI_M(H, L)$, thus slightly overloading the notation.*

Note that Definition 1 is a special case of Definition 13. We now define the noninterferent idealized variant based on the *total extension* of a PTS.

Definition 14 (U-total extension). *Given a PTS $M = \langle St, s_0, \mathfrak{U}, \mathfrak{A}, Obs, obs, do \rangle$ and a subset of users $U \subseteq \mathfrak{U}$, we define the U-total variant of M as $total_U(M) = \langle St, s_0, \mathfrak{U}, \mathfrak{A}, Obs, obs, do' \rangle$ where the transition function $do'(.)$ is defined as follows: for every $s \in St$, $v \in \mathfrak{U}$ and $a \in \mathfrak{A}$, $do'(s, v, a) = s$ if for some $u \in U$ we have $v = u$ and $do(s, u, a) = undef$, otherwise $do'(s, v, a) = do(s, v, a)$.*

Definition 15 (Noninterferent idealized model for PTN). *Given a partial transition network M and a set of "low" players L, we define the* noninterferent idealized variant of *M as $\mathcal{U}(total_L(M))$ such that:*

(i) $NI_{\mathcal{U}(total_L(M))}(H, L)$, and
(ii) for every $\mathcal{U}' \subsetneq \mathcal{U}$ it is not the case that $NI_{\mathcal{U}'(total_L(M))}(H, L)$.

Theorem 2. *For every partial transition network M, there is always a unique unification $\mathcal{U}$ satisfying properties (i) and (ii) from Definition 15.*

The proof is similar to the proof of Theorem 1, with the difference that we use $R^*_{total_L(M)}$ instead of R^*_M for constructing the idealized variant. However, as we use the concept of unwiding relation as the basis for using the R^* relation for constructing the idealized variant, we first need to modify the definition of the unwinding relation in Definition 10 and its corresponding proposition, Proposition 1 to adapt them to the new model:

Definition 16 (Unwinding for Noninterference in PTN). *$\sim_{NI_L} \subseteq St \times St$ is an* unwinding relation *iff it is an equivalence relation satisfying the conditions of* output consistency (OC)*,* step consistency (SC)*, and* local respect (LR)*. That is, for all states $s, t \in St$:*

(OC) *If $s \sim_{NI_L} t$ then $[s]_L = [t]_L$;*

(SC) *If* $s \sim_{NI_L} t$, $u \in L$, *and* $a \in \mathfrak{A}$ *then* $a \in act(s,u)$ *implies* $do(s,u,a) \sim_{NI_L} do(t,u,a)$*;*
(LR) *If* $u \in H$ *and* $a \in \mathfrak{A}$ *then* $a \in act(s,u)$ *implies* $s \sim_{NI_L} do(s,u,a)$*.*

Proposition 5. $NI_M(H,L)$ *iff there exist an unwinding relation* $\sim_{NI_L}$ *on the states of* M *that satisfies (OC), (SC) and (LR).*

The rest of the proof of Theorem 2 follows analogously.

Example 6. With PTS, the scenario from Example 2 can be modeled directly, without spurious states that ruled out illegal transitions. Thus, our models M_a, M_b for the two variants of the scenario are now exactly depicted in Figures 1 and 2.

The noninterferent idealized variants of M_a (resp. M_b) is again obtained by the unification of observations $noObs$, $MnameC$, and $MNameD$ (resp. $noObs$, $MnameA$, and $MNameB$). Clearly, L has no surely winning strategy to guarantee $\overline{\Gamma} = \Gamma_{\mathbb{T}}$ for $\mathbb{T} = \{s_{15}, s_{16}\}$ in M_a, $Ideal(M_a)$, and $Ideal(M_b)$. Moreover, he has a surely winning strategy in M_b. In consequence, M_a is effectively information-secure for (L,Γ), but M_b is not. □

The noninterferent variant was indeed an idealization in simple transition networks of Goguen and Mesguer. Is it still the case in partial transition networks? That is, is it always the case that L has no more abilities in $Ideal(M)$ than in M? In general, no. On one hand, L's observational capabilities are more limited in $Ideal(M)$, and in consequence some strategies in M are no longer uniform in $Ideal(M)$. On the other hand, unification U^* possibly adds new transitions to M, that can be used by L in $Ideal(M)$ to construct new strategies. However, under some reasonable assumptions, $Ideal(M)$ does provide idealization, as shown in the two propositions below. The proofs are relatively simple, and we omit them due to lack of space.

Proposition 6. *Let* M *be a PTN such that for every state* s *in* M *there is at least one player* $u \notin L$ *with* $act(s,u) \neq \emptyset$*. Then, for any* Γ*, we have that* $M \preceq_{L,\Gamma} Ideal(M)$*.*

Proposition 7. *For any PTN* M *and safety goal* Γ*, we have* $M \preceq_{L,\Gamma} Ideal(M)$*.*

7 Conclusions

In this paper, we introduce the novel concept of *effective information security*. The idea is aimed at assessing the relevance of information leakage in a system, based on how much the leakage enables an adversary to harm the correct behavior of the system. This contrasts with the common approach to information flow security where revealing any information is seen as being intrinsically harmful. We say that two information flows are *effectively equivalent* if the strategic ability of the adversary is similar in both of them. Moreover, one of them is *less effectively secure* than the other one if the amount of information leaked to the adversary in it increases the damaging ability of the adversary.

In order to determine how critical the information leakage is in a given system, we compare the damaging ability of the adversary to his ability in the idealized variant of the model. We define idealized models based on noninterference, and show that the construction is well defined. We prove this first for the deterministic, fully asynchronous transition networks of Goguen and Meseguer, and then extend the results to structures that allow for a more flexible modeling of interaction. The construction includes an algorithm that computes the idealized variant of each model in polynomial time wrt the size of the model.

Note that the concept of effective security is orthogonal to noninterference. The latter can be in principle replaced in our construction by an arbitrary property of information flow. The same reasoning scheme could be applied to noninference, nondeducibility, strategic noninterference, and so on. The pattern does not change: given a property $\mathcal{P}$, we define the idealized variant of M through the minimal unification U such that $U(M)$ satisfies $\mathcal{P}$. Then, M is effectively information-secure in the context of property $\mathcal{P}$ iff it is strategically equivalent to $U(M)$. We leave the investigation of which information security properties have unique minimal unifications for future work. Moreover, we are currently working on a more refined version of effective information security based on coalitional effectivity functions, in which the strategic ability of the adversary is not only compared at the initial state of the system, but across the whole state space.

Acknowledgements. Wojciech Jamroga acknowledges the support of the 7th Framework Programme of the European Union under the Marie Curie IEF project ReVINK (PIEF-GA-2012-626398).

References

1. A. Acquisti and J. Grossklags. Privacy attitudes and privacy behavior - losses, gains, and hyperbolic discounting. In *Economics of Information Security*, volume 12 of *Advances in Information Security*, pages 165–178. Springer, 2004.
2. D. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24:84–90, 1981.
3. A.S. Dimovski. Ensuring secure non-interference of programs by game semantics. In *Security and Trust Management*, pages 81–96. Springer, 2014.
4. A. Fielder, E. Panaousis, P. Malacaria, C. Hankin, and F. Smeraldi. Game theory meets information security management. *IFIP Advances in Information and Communication Technology*, 428:15–29, 2014.
5. A. Fujioka, T. Okamoto, and K. Ohta. A practical secret voting scheme for large scale elections. In *Proceedings of AUSCRYPT*, pages 244–251, 1992.
6. R. Giacobazzi and I. Mastroeni. Abstract non-interference: parameterizing non-interference by abstract interpretation. In *Proceedings of POPL*, pages 186–197. ACM, 2004.
7. Joseph A Goguen and José Meseguer. Security policies and security models. In *Proceedings of S&P*, pages 11–20. IEEE Computer Society, 1982.
8. O. Goldreich, S. Micali, and A. Wigderson. How to play ANY mental game. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing STOC '87*, pages 218–229. ACM, 1987.

9. James W Gray III. Probabilistic interference. In *Proceedings of S&P*, pages 170–179. IEEE, 1990.
10. J. Grossklags, N. Christin, and J. Chuang. Secure or insure? A game-theoretic analysis of information security games. In *Proceedings of WWW*, pages 209–218. ACM, 2008.
11. C. Hankin, R. Nagarajan, and P. Sampath. Flow analysis: Games and nets. In *The Essence of Computation, Complexity, Analysis, Transformation. Essays Dedicated to Neil D. Jones [on occasion of his 60th birthday]*, volume 2566 of *Lecture Notes in Computer Science*, pages 135–156. Springer, 2002.
12. W.R. Harris, S. Jha, T.W. Reps, J. Anderson, and R.N.M. Watson. Declarative, temporal, and practical programming with capabilities. In *Proceedings of SP*, pages 18–32. IEEE Computer Society, 2013.
13. W. Jamroga and M. Tabatabaei. Strategic noninterference. In *Proceedings of the 30th International Conference on ICT Systems Security and Privacy Protection IFIP SEC 2015*, volume 455 of *IFIP Advances in Information and Communication Technology*, pages 67–81. Springer, 2015.
14. J. Levin. In what city did you honeymoon? and other monstrously stupid bank security questions. *Slate*, 2008.
15. K. Leyton-Brown and Y. Shoham. *Essentials of Game Theory: A Concise, Multidisciplinary Introduction*. Morgan & Claypool, 2008.
16. Peng Li and Steve Zdancewic. Downgrading policies and relaxed noninterference. In *ACM SIGPLAN Notices*, volume 40, pages 158–170. ACM, 2005.
17. P. Malacaria and C. Hankin. Non-deterministic games and program analysis: An application to security. In *Proceedings of LICS*, pages 443–452. IEEE Computer Society, 1999.
18. Daryl McCullough. Noninterference and the composability of security properties. In *Proceedings of S&P*, pages 177–186. IEEE, 1988.
19. Annabelle McIver and Carroll Morgan. A probabilistic approach to information hiding. *Programming Methodology*, pages 441–460, 2003.
20. R. McNaughton. Testing and generating infinite sequences by a finite automaton. *Information and Control*, 9:521–530, 1966.
21. T. Moore and R. Anderson. Economics and internet security: a survey of recent analytical, empirical and behavioral research. Technical Report TR-03-11, Computer Science Group, Harvard University, 2011.
22. Colin O'Halloran. A calculus of information flow. In *Proceedings of ESORICS*, pages 147–159, 1990.
23. Alessandra Di Pierro, Chris Hankin, and Herbert Wiklicky. Approximate noninterference. *Journal of Computer Security*, 12(1):37–81, 2004.
24. J. A. Robinson. A machine-oriented logic based on the resolution principle. *Journal of the ACM*, 12(1):23–41, 1965.
25. A. W. Roscoe, C. A. R. Hoare, and R. Bird. *The Theory and Practice of Concurrency*. Prentice Hall PTR, 1997.
26. A. Sabelfeld and D. Sands. Dimensions and principles of declassification. In *Proceedings of CSFW-18*, pages 255–269. IEEE Computer Society, 2005.
27. Geoffrey Smith. On the foundations of quantitative information flow. In *Foundations of Software Science and Computational Structures*, pages 288–302. Springer, 2009.
28. David Sutherland. A model of information. In *Proc. 9th National Computer Security Conference*, pages 175–183, 1986.
29. R. van der Meyden and C. Zhang. A comparison of semantic models for noninterference. *Theoretical Computer Science*, 411(47):4123–4147, 2010.

30. J.T. Wittbold and D.M. Johnson. Information flow in nondeterministic systems. In *IEEE Symposium on Security and Privacy*, pages 144–144, 1990.
31. S. Zdancewic and A.C. Myers. Observational determinism for concurrent program security. In *Proceedings of CSFW-16*, pages 29–43. IEEE Computer Society, 2003.