Secrecy in Stochastic Discrete Event Systems

Mariam Ibrahim[1,2], Jun Chen[1], *Student Member, IEEE* and Ratnesh Kumar[1], *Fellow, IEEE*
[1] Dept. of Elec. & Comp. Eng., Iowa State Univ., Ames, IA 50011.
[2] German Jordanian University, Amman 11180 Jordan.

***Abstract*— In security critical system, keeping a property of system behaviors *secret* from an observer (or adversary, who has a partial observation of any executed behavior) is crucial. This paper proposes two notions of *secrecy* for stochastic discrete event systems. The notion of S_τ-*Secrecy* requires the set of system traces, that reveals the secret to an observer, occurs with probability smaller than τ. A stronger notion of *Increasing-S-Secrecy* captures system requirement that the secrecy level become increasingly tighter as the system evolves for longer periods. Algorithms for verifying both notions are provided. An illustrative example is examined to demonstrate the proposed notions.**

I. Introduction

Various aspects of secrecy have been explored in literature. References [1], [2], [3] defined the non-interference for input-output systems as a property in which the outputs that are observable to an *adversary* should not depend on any *secret* input so that the adversary does not deduce anything about the secret input by observing the output. Non-interference is a logical notion that is either satisfied or violated, and as such it does not allow the quantification of the degree to which a system may violate the property. To circumvent this limitation, the notion is enriched for probabilistic systems for which the amount of interference can be quantified in terms of the amount of information leaked by a system to an observer, where the amount of information leaked is measured in terms of the loss of uncertainty about the inputs due to the observation of the outputs, i.e., the difference between the prior and posterior entropies of the inputs, and equals the mutual information between the inputs and the outputs [1]. While such a quantification of information leakage is satisfactory for long periods of system operation (since entropy measures uncertainty in an average sense), it is of limited use for systems in which an adversary makes a single observation. To address this situation, the average case measure of entropy was replaced by its best case measure, namely *min-entropy* in the definition of mutual information [3].

In general, a secret can be a property of a sequence of inputs, as opposed to just a single input, and this general situation has also been examined in the literature. For example in the setting of discrete event systems (DESs), the definition of secrecy examined in [4] requires that the execution of behaviors constituting a secret must not be revealed to an observer by masking those behaviors through indistinguishable behaviors that are non-secret, known as *cover*. This is indeed analogous to the notion of non-interference, which by virtue of being logical has the same limitation that it cannot quantify the degree to which a system is interfering (or leaks information). For probabilistic DESs, where each discrete transition is associated with a certain occurrence probability, more powerful notions of secrecy can be defined. For example, [5] used Jensen-Shannon divergence between the distributions of a secret versus its cover as a way to quantify the secrecy, which it measures as the divergence of two distributions over the set of feasible observations, one being the probabilities of secret behaviors and the other being the probabilities of the behaviors in the cover. While this is indeed an interesting way of quantifying the level of secrecy (see Section III for further discussions), computability of the Jensen-Shannon divergence was not reported: Only an approximation algorithm for upper bounding was provided in [5]. Another attempt to generalize secrecy from logical to stochastic DESs includes [6], where, alike the setting of mutual information based characterization of information leakage, the authors consider the difference between the prior and posterior distributions of the secret states, and require it to be upper bounded. The corresponding verification problem turns out to be undecidable. In another paper [7], the same authors proposed the notion of *Step-Based Almost Current-State Opacity* requiring the probability of revealing the secret must be upper bounded at each time step. While this notion is decidable, it is more stringent than the one we propose below, which bounds the probability of revealing the secret over the set of *all* behaviors, as opposed to for each step.

In this paper we propose a new divergence based notion of secrecy, with the benefit that it remains computable. Our notion of S_τ-Secrecy requires that the set of behaviors revealing the secret to an adversary, must occur with probability upper bounded by τ. We show that S_τ-Secrecy can be viewed as a generalization of the logical secrecy defined in [4], and that it is a variant of the divergence used in [5]. Further, it is desirable that as a system operates for longer and longer periods, it reveals lesser and lesser amount of secret, i.e., the probability of unambiguous traces that can reveal the secret becomes increasingly smaller. Accordingly, we introduce a stronger notion, namely, *increasingly stochastic-secret*, or *I-S-Secret* to capture this additional property of system becoming progressively secret. Decidable algorithms for computing the aforementioned divergence as well as for verifying S_τ-Secrecy and I-S-Secrecy are also provided. The

The research was supported in part by the National Science Foundation under the grants, NSF-ECCS-0801763, NSF-CCF-0811541, and NSF-ECCS-0926029. Author emails: {mariami,junchen,rkumar}@iastate.edu.

978-1-4799-3106-4/14/$31.00

concepts and the algorithms are illustrated through a simple vehicle tracking application.

II. Notations and Preliminaries

For an event set Σ, define $\overline{\Sigma} := \Sigma \cup \{\epsilon\}$, where ϵ denotes "no-event". The set of all finite length event sequences over Σ, including ϵ is denoted as Σ^*, and $\Sigma^+ := \Sigma^* - \{\epsilon\}$. A *trace* is a member of Σ^* and a *language* is a subset of Σ^*. Use $s \leq t$ to denote if $s \in \Sigma^*$ is a prefix of $t \in \Sigma^*$, and $|s|$ to denote the length of s or the number of events in s. For $L \subseteq \Sigma^*$, its prefix-closure is defined as $pr(L) := \{s \in \Sigma^* | \exists t \in L : s \leq t\}$ and L is said to be prefix-closed (or simply closed) if $pr(L) = L$, i.e., whenever L contains a trace, it also contains all the prefixes of that trace. For $s \in \Sigma^*$ and $L \subseteq \Sigma^*$, $L\backslash s := \{t \in \Sigma^* | st \in L\}$ denotes the set of traces in L *after* s.

A stochastic DES can be modeled by a *stochastic automaton* $G = (X, \Sigma, \alpha, x_0, X_m)$, where X is the set of states, $X_m \subseteq X$ is the set of marked states, Σ is the finite set of events, $x_0 \in X$ is the initial state, and $\alpha : X \times \Sigma \times X \to [0, 1]$ is the probability transition function [8], and $\forall x \in X, \sum_{\sigma \in \Sigma} \sum_{x' \in X} \alpha(x, \sigma, x') = 1$. G is non-stochastic if $\alpha : X \times \Sigma \times X \to \{0, 1\}$, and a non-stochastic DES is deterministic if $\forall x \in X, \sigma \in \Sigma, \sum_{x' \in X} \alpha(x, \sigma, x') \leq 1$. The transition probability function α can be generalized to $\alpha : X \times \Sigma^* \times X$ in a natural way. Define the languages generated and marked by G as $L(G) := \{s \in \Sigma^* \mid \exists x \in X, \alpha(x_0, s, x) > 0\}$ and $L_m(G) := \{s \in \Sigma^* \mid \exists x \in X_m, \alpha(x_0, s, x) > 0\}$, respectively. For a given G, a *component* $C = (X_C, \alpha_C)$ of G is a "subgraph" of G, i.e., $X_C \subseteq X$ and $\forall x, x' \in X_C$ and $\sigma \in \Sigma$, $\alpha_C(x, \sigma, x') = \alpha(x, \sigma, x')$, whenever the latter is defined. C is said to be a *strongly connected component* (SCC) or *irreducible* if $\forall x, x' \in X_C$, $\exists s \in \Sigma^*$ such that $\alpha_C(x, s, x') > 0$. A SCC C is said to be *closed* if for each $x \in X_C$, $\sum_{\sigma \in \Sigma} \sum_{x' \in X_C} \alpha_C(x, \sigma, x') = 1$. The states which belong to a closed SCC are *recurrent states* and the remaining states (that do not belong to any closed SCC) are *transient states*. Identifying the set of recurrent states can be done by the polynomial algorithm presented in [9]. The following is a useful property of a finite-state Markov chain, which states that as the number of transitions increases, the probability of the Markov chain being in a transient state approaches zero.

Property 1 ([10]): Let X be the state space of a finite-state Markov chain and $X = X_R \cup X_T$, where X_R and X_T denote the set of recurrent and transient states, respectively. Let $x \in X$ be an arbitrary state of the chain and t be any transition sequence starting from x. Then

$$(\forall \tau > 0)(\exists n \in \mathbb{N}) \quad Pr(t : \exists x' \in X_T, \alpha(x, t, x') > 0, |t| \geq n) < \tau. \quad (1)$$

Example 1: Fig. 1 is an example of a stochastic automaton G. The set of states is $X = \{0, 1, \ldots, 7\}$ with initial state $x_0 = 0$ and marked states $X_m = \{1, 3, 6, 7\}$, event set $\Sigma = \{a, b, c, d\}$. A state is depicted as a node, whereas a transition is depicted as an edge between its origin and termination states, with its event name and probability value labeled on the edge. It can be seen that the marked language is $L_m(G) = ab^*d^* \cup ab^*cd^+ \cup cab^*cb^*d^*$. ■

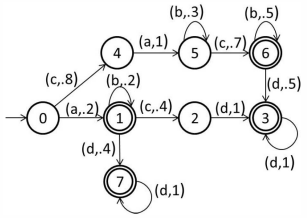

Fig. 1. Stochastic automaton G.

The events executed by a DES can be partially observed by an observer (i.e., an adversary). The limited observation capability of an observer can be represented as an observation mask, $M : \overline{\Sigma} \to \overline{\Delta}$, where Δ is the set of observed symbols and $M(\epsilon) = \epsilon$. An event σ is unobservable if $M(\sigma) = \epsilon$. The set of unobservable events is denoted as Σ_{uo} and the set of observable events is then given by $\Sigma - \Sigma_{uo}$. The observation mask can be generalized to $M : 2^{\Sigma^*} \to 2^{\Delta^*}$ in a natural way.

In [4], the logical secrecy for DESs was examined. Suppose $K \subseteq \Sigma^*$ models some property of interest, then the subset $L_m \cap K$ of marked traces is considered a secret, whereas the remaining marked traces in $L_m - K$ can be viewed as its cover. Then the following definition of logical secrecy requires that each trace in the secret must be masked by an indistinguishable cover, and vice-versa:

Definition 1: Given a DES G with marked language $L_m(G) = L_m$, a language $K \subseteq \Sigma^*$, and an observation mask M of an observer, (L_m, K) is said to be *Secret* if $\forall s \in L_m$,

- $s \in K \Rightarrow (M^{-1}M(s) \cap L_m) - K \neq \emptyset$, and
- $s \notin K \Rightarrow (M^{-1}M(s) \cap L_m) \cap K \neq \emptyset$.

III. Stochastic Secrecy of DESs

In this section we consider a generalization of the definition of logical secrecy presented in Section II to the stochastic setting so as to introduce a more flexible notion. The logical version is rigid in the sense that it is either satisfied or violated, and there is no way to assess the degree to which a system satisfies/violates the property of secrecy. To address this issue, the authors in [5] introduced a series of generalizations, each finer than its precursor, with the final one computing the divergence between the distribution of the behaviors in the secret versus their cover. As noted above, we can view traces in $L_m \cap K$ play the role of the secret, whereas traces in $L_m - K$ are treated as cover. When the system executes a trace $s \in L_m$, its observation $M(s)$ is received by an adversary, and based on this, it can compute the probabilities that the observation came from the secret versus its cover. Clearly, for perfect secrecy, the two probabilities must be equal, and this must hold for all observations, and in which case the divergence between the two distributions would be zero. But if the divergence is non-zero, it quantifies the degree to which the perfect secrecy is compromised. In certain applications, low

levels of compromise, as determined in terms of low levels of divergence may be acceptable. Thus, one may measure the degree or level with which secrecy is compromised by measuring the divergence between the two distributions defined over all observations, namely, the probabilities of the observation coming from the secret versus its cover. With this in mind, for each $o \in M(L_m)$, we define the two probabilities $p_s(o)$ and $p_c(o)$, the probability of secret versus of cover given the observation o, as below:

$$p_s(o) = \frac{Pr(s \in L_m \cap K : M(s) = o)}{Pr(L_m \cap K)}, \quad (2)$$

$$p_c(o) = \frac{Pr(s \in L_m - K : M(s) = o)}{Pr(L_m - K)}. \quad (3)$$

Their weighted Jensen-Shannon (JS) divergence $D^{\lambda}_{JS}(p_s, p_c)$ with weight $0 \leq \lambda \leq 1$, is given by:

$$\begin{aligned} D^{JS}_{\lambda}(p_s, p_c) &:= \lambda D(p_s, \lambda p_s + (1-\lambda)p_c) \\ &\quad +(1-\lambda)D(p_c, \lambda p_s + (1-\lambda)p_c), \end{aligned}$$

where $D(p_1, p_2)$ denotes the relative-entropy between the distributions p_1 and p_2 and is defined by,

$$D(p_1, p_2) := \sum_{o \in M(L_m)} p_1(o) \log_2 \left(\frac{p_1(o)}{p_2(o)} \right). \quad (4)$$

It should be noted that while the relative-entropy is asymmetric, the JS divergence is symmetric (and was introduced for symmetrization). Authors in [5] proposed the use of the JS divergence $D^{JS}_{\lambda}(p_s, p_c)$ with the weight $\lambda = \frac{Pr(L_m \cap K)}{Pr(L_m \cap K) + Pr(L_m - K)}$ (so $1 - \lambda = \frac{Pr(L_m - K)}{Pr(L_m \cap K) + Pr(L_m - K)}$), where the weight is simply used for normalization, as a means to measure the level with which secrecy is compromised: When $p_s = p_c$, there is zero divergence and perfect secrecy; any non-zero divergence then measures the degree of deviation from the perfect secrecy; lower the divergence, lower the deviation, lower the compromise in secrecy, and higher the degree of secrecy. While the JS divergence is indeed a meaningful way of measuring the level of secrecy, a difficulty is that in general, JS divernce is hard to compute. In fact, [5] only provides a computation that yields an upper bound approximation to the JS divergence. In order to alleviate this computational difficulty, we propose a new type of divergence, one that replaces the "$\log_2$" function in the definition of JS divergence with *another monotonic function*, namely, an indicator function of a lower-bounded set (so the indicator value is 1 if and only if the argument exceeds the lower bound). We first define the probability of unambiguous traces in L_m, where the ambiguity for a trace in $L_m \cap K$ is produced by an indistinguishable trace in $L_m - K$ and vice-versa. Following which, we show that the probability of unambiguous traces in L_m is indeed a new type of divergence, obtained by doing the aforementioned replacement of $\log_2$ function by an indicator function. This establishes that the notion of degree of secrecy that we examine in this paper is in fact a type of divergence, and as we show as follows, the advantage of using this new notion is that it can be computed precisely in a decidable fashion. Define,

$$Pr_{unamb}(L_m) := Pr\{s \in L_m : Pr_{amb}(s) = 0\}, \quad (5)$$

where $Pr_{amb} : L_m \to [0, 1]$ is the probability of s being ambiguous, and for $s \in L_m - K$ (resp., $s \in L_m \cap K$), $Pr_{amb}(s)$ is the probability of all indistinguishable traces in $L_m \cap K$ (resp., $L_m - K$) conditioned by the fact that the ambiguity is only caused by indistinguishable traces that are also feasible in L_m, and is given by:

$$Pr_{amb}(s) = \begin{cases} \frac{Pr(M^{-1}M(s) \cap L_m - K)}{Pr(M^{-1}M(s) \cap L_m)} & \text{if } s \in L_m \cap K \\ \frac{Pr(M^{-1}M(s) \cap L_m \cap K)}{Pr(M^{-1}M(s) \cap L_m)} & \text{if } s \in L_m - K. \end{cases} \quad (6)$$

Letting $S := Pr(L_m \cap K)$ denote the probability of secrets and $C := Pr(L_m - K)$ denote the probability of covers, we can rewrite equations (2) and (3) as:

$$\begin{aligned} Pr(s \in L_m \cap K : M(s) = o) &= Sp_s(o), \\ Pr(s \in L_m - K : M(s) = o) &= CP_c(o), \end{aligned}$$

and so

$$Pr(s \in L_m : M(s) = o) = Sp_s(o) + Cp_c(o).$$

Then,

$$Pr_{amb}(s) = \begin{cases} \frac{Cp_c(M(s))}{Sp_s(M(s)) + Cp_c(M(s))} & \text{if } s \in L_m \cap K \\ \frac{Sp_s(M(s))}{Sp_s(M(s)) + Cp_c(M(s))} & \text{if } s \in L_m - K. \end{cases}$$

Proposition 1: Given (L_m, K), $Pr_{unamb}(L_m)$ is equal to the following variant of JS divergence up to a constant scale:

$$\begin{aligned} \widetilde{D}^{JS}_{\lambda}(p_s, p_c) &:= \lambda \widetilde{D}(p_s, \lambda p_s + (1-\lambda)p_c) \\ &\quad +(1-\lambda)\widetilde{D}(p_c, \lambda p_s + (1-\lambda)p_c), \end{aligned}$$

where $\lambda = \frac{Pr(L_m \cap K)}{Pr(L_m \cap K) + Pr(L_m - K)} = \frac{S}{S+C}$, and $\widetilde{D}$ unction is obtained by replacing the "$\log_2$" function in (4) with an indicator function.

Since $\widetilde{D}^{JS}_{\lambda}$ is same as D^{JS}_{λ} with $\log_2$ replaced by indicator function in (4), the probability of unambiguous marked traces, defined in (5), can be viewed as a type of divergence.

Having defined a new notion of divergence, next we use it to define a new notion of secrecy: When the divergence, as measured in terms of probability of unambiguous traces, is smaller than $0 \leq \tau \leq 1$, we say that the system is stochastically-secret with level τ, denoted *S_τ-Secrect*. Further it is desirable that as the system operates for longer and longer duration, it becomes stronger and stronger in terms of level of secrecy, i.e., the probability of unambiguous traces becomes increasingly smaller. To capture this additional property, we introduce a stronger notion, namely, *increasingly stochastic-secret*, or *I-S-Secret*.

Definition 2: Given a DES G with marked language $L_m(G) = L_m$, a language $K \subseteq \Sigma^*$, and an observation mask M of an observer, (L_m, K) is said to be stochastically-secret with level τ, or simply *S_τ-Secrect* if

$$Pr_{unamb}(L_m) = Pr(s \in L_m : Pr_{amb}(s) = 0) < \tau.$$

Further, (L_m, K) is said to be *increasingly stochastic-secret*,

or simply *I-S-Secret* if $\forall \tau > 0$, $\exists n \in \mathbb{N}$,

$$Pr(s \in L_m : |s| \geq n, Pr_{amb}(s) = 0) < \tau,$$

where $Pr_{amb}(s)$ is defined by (6).

Remark 1: By the definition of $Pr_{amb}(s)$ in (6), it is trivial to see that,

- $\forall s \in L_m \cap K$, $[(M^{-1}M(s) \cap L_m) - K \neq \emptyset] \Leftrightarrow [Pr_{amb}(s) > 0]$, and
- $\forall s \in L_m - K$, $[(M^{-1}M(s) \cap L_m) \cap K \neq \emptyset] \Leftrightarrow [Pr_{amb}(s) > 0]$.

Therefore Definition 1 of logical secrecy requires the probability of unambiguous traces to be *zero*, i.e., it is equivalent to S_0-Secrecy and hence is a special case of S_τ-Secrecy. It should further be noted that as the level τ is reduced, the requirement becomes tighter (since the probability of unambiguous traces must now be smaller), and so S_τ-secrecy becomes stronger. In particular, S_0-secrecy is the strongest.

IV. Divergence Computation and Secrecy Verification

In this section we first present an algorithm to compute the new notion of divergence that we introduced in Section III, namely, the probability of unambiguous traces in L_m. This can be used to determine the smallest parameter τ with which a given system is S_τ-Secret. Then we provide an algorithmic test for verifying I-S-Secrecy.

Let the stochastic automaton $G = (X, \Sigma, \alpha, x_0, X_m)$ with marked language $L_m(G) = L_m$ be the system model, and the *deterministic* automaton $R = (Y, \Sigma, \beta, y_0, Y_m)$ be a *trim* acceptor of the language K, i.e., $L_m(R) = K$ and $L(R) = pr(K)$. Then a refinement of G with respect to R, denoted G^R, can be used to capture the property-satisfying/violating traces in form of the reachability of certain marked states (see below) and is given by $G^R := (X \times \overline{Y}, \Sigma, \gamma, (x_0, y_0), X_m \times \overline{Y})$, where $\overline{Y} = Y \cup \{D\}$, and $\forall (x, \overline{y}), (x', \overline{y}') \in X \times \overline{Y}, \sigma \in \Sigma, \gamma((x, \overline{y}), \sigma, (x', \overline{y}')) = \alpha(x, \sigma, x')$ if the following holds:

$$(\overline{y}, \overline{y}' \in Y \wedge \beta(\overline{y}, \sigma, \overline{y}') > 0) \vee (\overline{y} = \overline{y}' = D) \vee (\overline{y}' = D \wedge \sum_{y \in Y} \beta(\overline{y}, \sigma, y) = 0),$$

and otherwise $\gamma((x, \overline{y}), \sigma, (x', \overline{y}')) = 0$. Then it can be seen that the refined plant G^R has the following properties: (1) $L(G^R) = L(G)$ and $L_m(G^R) = L_m$; (2) any property-satisfying trace in $s \in L_m \cap K$ transitions the refinement G^R to a state that is marked in both coordinates, and any property-violating trace $s \in L_m - K$ transitions the refinement G^R to a state marked in the first coordinate but not marked in its second coordinate; (3) for each $s \in L(G) = L(G^R)$, $\sum_{x \in X} \alpha(x_0, s, x) = \sum_{(x, \overline{y}) \in X \times \overline{Y}} \gamma((x_0, y_0), s, (x, \overline{y}))$, i.e., the occurrence probability of each trace in G^R is the same as that in G.

Next we obtain a *deterministic* automaton $M(G^R)$ that accepts the masked traces $M(L(G^R))$ as follows. Replace each (σ, p) transition label in G^R by $M(\sigma)$, and next determinize the resulting automaton to obtain $M(G^R) = (Z, \Delta, \delta_M, Reach(\epsilon), Z_m)$, where $Z := 2^{X \times \overline{Y}} - \{\emptyset\}$, $Z_m := \{\hat{Z} \subseteq X \times \overline{Y} | \hat{Z} \cap (X_m \times \overline{Y}) \neq \emptyset\}$, and the function $Reach : M(\Sigma^*) \to 2^{X \times \overline{Y}}$ is defined as $Reach(o) := \{(x \times \overline{y}) \in X \times \overline{Y} : \exists s \in M^{-1}(o) \cap L(G), \gamma((x_0, y_0), s, (x, \overline{y})) > 0\}$. Next we construct a*deterministic* automaton that accepts all traces that are indistinguishable from system traces, $M^{-1}M(G^R) := (Z, \Sigma, \delta'_M, Reach(\epsilon), Z_m)$, where the transition function δ'_M is defined as $\delta'_M(z, \sigma, z') = \delta_M(z, M(\sigma), z')$ for each $z, z' \in Z$ and $\sigma \in \Sigma$. We also add self-loops at each state on all unobservable events, i.e., $\delta'_M(z, \sigma, z) = 1$ for each $z \in Z$ and $\sigma \in \Sigma_{uo}$.

To verify the secrecy properties, we construct a testing automaton $T = G^R \times M^{-1}M(G^R)$ which pairs each system trace with runs of all indistinguishable traces, and is given by $T = (X \times \overline{Y} \times Z, \Sigma, \delta, \{(x_0, y_0)\} \times Reach(\epsilon), X_m \times \overline{Y} \times Z_m)$. For each $(x, \overline{y}, z), (x', \overline{y}', z') \in T$ and $\sigma \in \Sigma$, $\delta((x, \overline{y}, z), \sigma, (x', \overline{y}', z')) = \gamma((x, \overline{y}), \sigma, (x', \overline{y}'))$ if $\delta'_M(z, \sigma, z') > 0$ and 0 otherwise. Define a state $(x, \overline{y}, z)$ of T as non-secret if $x \in X_m$ and

- $\overline{y} \in Y_m \Rightarrow \forall (x', \overline{y}') \in z, x' \notin X_m \vee \overline{y}' \in Y_m$, or
- $\overline{y} \notin Y_m \Rightarrow \forall (x', \overline{y}') \in z, x' \notin X_m \vee \overline{y}' \notin Y_m$.

The set of non-secret states of T is denoted as $\mathcal{N}_T$; the remaining states of T are secret, denoted as $\mathcal{S}_T$. The following proposition states that a marked trace is unambiguous (i.e., with zero probability of ambiguity) if and only if it can reach a non-secret state in T after the execution of s, and follows from the definition of $Pr_{amb}(s)$ in (6) and the construction of T.

Proposition 2: For any trace $s \in L_m$, $Pr_{amb}(s) = 0 \Leftrightarrow \exists v \in \mathcal{N}_T, \delta(v_0, s, v) > 0$, where v_0 is the initial state of T.

Example 2: For G shown in Fig. 1, suppose the acceptor R is given in Fig. 2(a), i.e., $K = L_m(R) = c(a^*b^*)^*c(b^*d^*)^* \cup ab^* \cup ab^*cd(d^*b^*)^*$. Suppose the observation mask M for the observer is such that $M(c) = \epsilon$ and for all other events $\sigma \in \{a, b, d\}$, $M(\sigma) = \sigma$. Then the refinement G^R, automaton $M^{-1}M(G^R)$ and testing automaton T are shown in Fig. 2(b)-(c), where $z_0 = \{(0,0), (4,4)\}, z_1 = \{(5,4), (6,3), (1,1), (2,2)\}, z_2 = \{(3,3), (7,D)\}$. T possesses a non-secret state $(1, 1, z_1)$, so following Proposition 2, there exists s with $Pr_{amb}(s) = 0$. Hence (L_m, K) is not S_0-Secret, and so not logically secret. ■

A. Computation of level τ in S_τ-Secrecy

The secrecy level τ in the definition of S_τ-Secrecy equals the probability of unambiguous marked traces, $Pr_{unamb}(L_m)$, and it can be computed by computing the probability of reachability of non-secret states as a target set of states in the testing automaton T. In a trivial case where there are no non-secret states ($\mathcal{N}_T = \emptyset$), then obviously $Pr_{unamb}(L_m) = 0$. When $\mathcal{N}_T \neq \emptyset$, then $Pr_{unamb}(L_m)$ equals to the probability of reaching $\mathcal{N}_T$, i.e., the probability of the set of traces that first hit $\mathcal{N}_T$.

Define for each $v = (x, \overline{y}, z) \in \mathcal{N}_T$,

$$Pr_{unamb}(v) := Pr(s \in L_m : [\delta(v_0, s, v) > 0] \wedge [\forall u \in pr(s) \cap L_m, Pr_{amb}(u) > 0]),$$

Then it follows that $Pr_{unamb}(L_m) = \sum_{v \in \mathcal{N}_T} Pr_{unamb}(v)$. To recursively compute $Pr_{unamb}(v)$, we define for each $i \in \mathcal{S}_T$ and $j \in \mathcal{N}_T$,

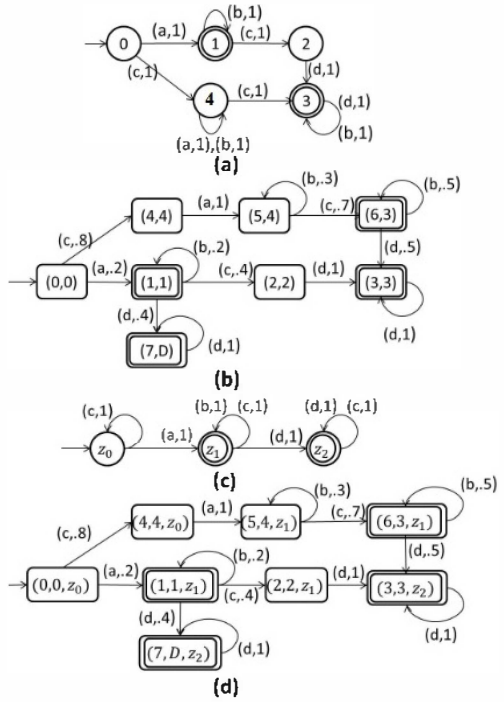

Fig. 2. (a) Acceptor R; (b) refinement G^R; (c) automaton $M^{-1}M(G^R)$; (d) testing automaton T.

$$p(i,j) := Pr(s \in L_m : [\delta(i,s,j) > 0] \wedge [\forall u < s, \forall k \in \mathcal{N}_T, \delta(i,u,k) = 0]),$$

which is the probability of the set of traces that transition the testing automaton T from a secret state i to a non-secret state j, without visiting any other non-secret state in $\mathcal{N}_T$.

Then it can be seen that for any $i \in \mathcal{S}_T$ and $j \in \mathcal{N}_T$, the following recursion holds: $p(i,j) = \sum_{k \in \mathcal{S}_T} \Omega(i,k)p(k,j) + \Omega(i,j)$, where $\Omega(m,n) := \sum_{\sigma \in \Sigma} \delta(m,\sigma,n), \forall m,n \in \mathcal{S}_T \cup \mathcal{N}_T$, and the first right hand side (RHS) term corresponds to transitioning to a non-secret state in more than one steps while the second RHS term corresponds to transitioning in exactly one step. Then all the probabilities $\{p(i,j) | i \in \mathcal{S}_T, j \in \mathcal{N}_T\}$ can be found by solving the following matrix equation (see for example [11], [12] for a similar matrix equation):

$$\boldsymbol{p} = \Omega_1 \boldsymbol{p} + \Omega_2, \tag{7}$$

where $\boldsymbol{p}$ is a $|\mathcal{S}_T| \times |\mathcal{N}_T|$ matrix with the ijth element given by $p(i,j)$, Ω_1 is a $|\mathcal{S}_T| \times |\mathcal{S}_T|$ matrix formed by the entries $\{\Omega(m,n) | m,n \in \mathcal{S}_T\}$ and Ω_2 is a $|\mathcal{S}_T| \times |\mathcal{N}_T|$ matrix formed by the entries $\{\Omega(m,n) | m \in \mathcal{S}_T, n \in \mathcal{N}_T\}$.

Then it follows that $Pr_{unamb}(v) = p(v_0, v)$. Note that in the trivial case when $v_0 \in \mathcal{N}_T$, we have $Pr_{unamb}(v_0) = 1$.

Remark 2: To find $\boldsymbol{p}$ using Equation (7), we need to solve $\boldsymbol{p} = (I - \Omega_1)^{-1}\Omega_2$. The complexity of matrix inverse is $O(|\mathcal{S}_T|^3)$ and the complexity of matrix multiplication is $O(|\mathcal{S}_T|^2 \times |\mathcal{N}_T|)$, and so overall complexity is $O((|\mathcal{S}_T|^2) \times (|\mathcal{S}_T| + |\mathcal{N}_T|))$. Since the number of secret states $\mathcal{S}_T$ and number of non-secret states $\mathcal{N}_T$ is upper bounded by the number of states in T, which is $O(|X| \times |Y| \times 2^{|X| \times |Y|})$, the complexity of finding $\boldsymbol{p}$ using Equation (7) is bounded by $O(|X|^3 \times |Y|^3 \times 2^{|X| \times |Y|})$.

Example 3: For the testing automaton shown in Fig. 2, states can be indexed as $v_0 = (0,0,z_0)$, $v_1 = (4,4,z_0)$, $v_2 = (5,4,z_1)$, $v_3 = (6,3,z_1)$, $v_4 = (1,1,z_1)$, $v_5 = (2,2,z_1)$, $v_6 = (3,3,z_2)$ and $v_7 = (7,D,z_2)$. The set of non-secrete states is given as $\mathcal{N}(T) = \{v_3, v_4\}$. Then the matrices Ω_1 and Ω_2 are given as: $\Omega_1(1,2) = 0.8, \Omega_1(2,3) = 1, \Omega_1(3,3) = 0.3, \Omega_1(4,5) = 1, \Omega_1(5,5) = 1, \Omega_1(6,6) = 1$ and 0 for other entries of Ω_1; $\Omega_2(1,2) = 0.2, \Omega_2(3,1) = 0.7$ and 0 for other entries of Ω_2. By solving (7), we get: $\boldsymbol{p}(1,1) = 0.8$, $\boldsymbol{p}(1,2) = 0.2$, $\boldsymbol{p}(2,1) = 1$, $\boldsymbol{p}(3,1) = 1$, and 0 for other entries of $\boldsymbol{p}$. Hence, $Pr_{unamb}(L_m) = \sum_{v \in \mathcal{N}_T} Pr_{unamb}(v) = \sum_{v \in \mathcal{N}_T} p(v_0, v) = 1$, i.e., (L_m, K) is not S_τ-Secret with any $\tau \leq 1$. ■

B. Verification of I-S-Secrecy

Next theorem gives a necessary and sufficient condition for verification of I-S-Secrecy.

Theorem 1: Given a DES G with marked language $L_m(G) = L_m$, a language $K \subseteq \Sigma^*$, and an observation mask M of an observer, (L_m, K) is *I-S-Secret* if and only if $\mathcal{R}_T \cap \mathcal{N}_T = \emptyset$, where $\mathcal{R}_T$ is the set of recurrent states of testing automaton T.

Proof: (Sufficiency) If $\mathcal{R}_T \cap \mathcal{N}_T = \emptyset$, then for any $s \in L_m$, $Pr_{amb}(s) = 0$ implies that s transitions T to a transient state and so

$$Pr(s \in L_m : |s| \geq n, Pr_{amb}(s) = 0) \leq Pr(s \in L : |s| \geq n, \exists v \notin \mathcal{R}_T, \delta(v_0, s, v) > 0). \tag{8}$$

Combining (1) and (8), we have $\forall \tau > 0, \exists n \in \mathbb{N}, Pr(s \in L_m : |s| \geq n, Pr_{amb}(s) = 0) < \tau$. Therefore the system is *I-S-Secret*, and the sufficiency follows.

(Necessity) When there exists an non-secret state v which is recurrent, let it be reached by the execution of trace s. Let C be the closed SCC that v belongs to, and π_C be the stationary distribution of C. Let $t \in L_m \backslash s$ be such that st steers C into its stationary distribution. Then all extensions of st, that transitions C back to state v, occurs with probability $\pi_C(v)$, i.e., $Pr(u \in L_m \backslash st : \delta(v_0, stu, v) > 0) = \pi_C(v)$. Also since $v \in \mathcal{N}_T$, we have for any $u \in L_m \backslash st$, $\delta(v_0, stu, v) > 0 \Rightarrow Pr_{amb}(stu) = 0$. Therefore for any $n \in \mathbb{N}$, $Pr(s \in L_m : |s| \geq n, Pr_{amb}(s) = 0) \geq Pr(st) Pr(u \in L_m \backslash st : \delta(v_0, stu, v) > 0) = Pr(st)\pi_C(v)$. Therefore, by choosing $\tau < Pr(st)\pi_C(v)$, one can conclude that the system is not *I-S-Secret*. ■

Remark 3: Note the testing automaton $T = G^R \times M^{-1}M(G^R)$ has $O(|X| \times |Y| \times 2^{|X| \times |Y|})$ states and $O(|\Sigma| \times |X|)$ transitions per state since only the G part is nondeterministic, whereas the complexity for identifying all the non-secret recurrent states in T is cubic in the number of states in T and linearly in the number of transitions in T, respectively [12]. So the complexity of checking I-S-Secrecy using *Theorem 1* is $O(|X|^3 \times |Y|^3 \times 2^{|X| \times |Y|} + |\Sigma| \times |X|^2 \times |Y| \times 2^{|X| \times |Y|})$.

Example 4: For the testing automaton shown in Fig. 2, $\mathcal{N}(T) = \{(1,1,z_1), (6,3,z_1)\}$ and the set of recurrent states

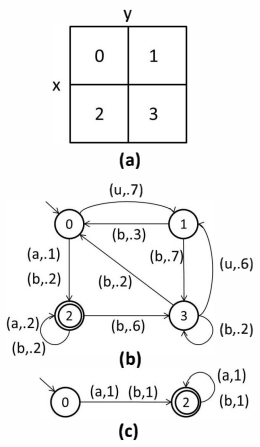

Fig. 3. (a) 2-dimensional grid in which a vehicle can move; (b) automaton G modeling the vehicle kinematic and the sensor readings; (c) acceptor R.

is given by $\mathcal{R}_T = \{(7, D, z_2), (3, 3, z_2)\}$. Since $\mathcal{N}(T) \cap \mathcal{R}_T = \emptyset$, (L_m, K) is *I-S-Secret*. The same system is shown to be not S_τ-Secret for any τ, however it turns out to be *I-S-Secret*, meaning that the longer the system operates, the stronger the system becomes in terms of the level of secrecy, and for any $\tau > 0$, there exists a length bound, beyond which the system is S_τ-Secret. ■

V. An Illustrative Example

Consider a vehicle that can navigate among 4 cells as shown in Fig. 3(a), whose movements among cells may be tracked by an adversary using a set of sensors to infer whether the vehicle may have performed certain secret navigation. (The example is inspired from [13], and modified to suit our setting.) Owing to the vehicle dynamics, transition among all cell pairs is not possible, and the model in Fig. 3(b) shows the allowable transitions. While traversing from one cell to a next possible cell, the vehicle may pass within the range of either sensor a, generating event a, sensor b, generating event b, or no sensor, generating no event – such transitions are labeled by u. These transitions, along with their event labels, and their occurrence probabilities are shown in Fig. 3(b). For example, from the initial cell, the vehicle can unobservably move to cell 1 with probability 0.7, and with the remainder probability 0.3 to cell 2, either producing observation a with probability 0.1 or b with probability 0.2. Cell 0 is the initial cell (shown as node with an entering arrow) and cell 2 is the final cell (shown as marked node). The set of behaviors starting at the initial cell (cell 0), ending at final cell (cell 2), and visiting only these two cells one or more times is considered *secret*, that should be hidden from an adversary. Then the remaining behaviors act as a *cover*. Letting L_m denote the marked behaviors, and K denote the set of secrets, the acceptor G of L_m is shown in Fig. 3(b), and the acceptor R of K is obtained by restricting G to the left two states 0 and 2, which is shown in Fig. 3(c). The testing automaton T, which consists of 36 states and 99 transitions, is omitted here for space considerations. In this example, the probability of unambiguous traces, $Pr_{unamb}(L_m) = 0.3$, and hence (L_m, K) is S_τ-Secret for any $\tau > 0.3$. It can also be verified that all the non-secret states of T are transient states, which satisfies the condition in Theorem 1. Therefore, (L_m, K) satisfies the I-S-Secrecy property.

VI. Conclusion

In this work we have studied the secrecy in stochastic discrete event systems. A new type of divergence based on the probability of unambiguous traces was introduced which alleviates the computational difficulty of the current divergence-based secrecy notion in literature. Algorithm for computing the proposed divergence was also presented, which employs the computation of the probability of reaching a target set of states. This yields the minimum level τ with which a given system is S_τ-Secret. We also proposed a notion of I-S-Secrecy, which requires the divergence of the longer behaviors of the system to be upper bounded by tighter bounds. Checking the I-S-Secrecy was shown to be equivalent to checking a certain recurrence property of a testing automaton. Future work will consider stronger notion where traces with ambiguity below a tolerance level, that is not necessarily zero, will be counted towards secrecy violation. Finally, while the property of secrecy is needed to hide information, the dual property of diagnosability is needed to reveal defects. Examining the joint property of secrecy and diagnosability (see [12]) for systems will be another future direction.

References

[1] G. Smith, "On the foundations of quantitative information flow," in *Proc. Int. Conf. Foundations of Software Science and Computation Structures (FoSSaCS 09)*, Miami, FL, 2009, pp. 288–302.

[2] B. K. Michael Backes and A. Rybalchenko, "Automatic discovery and quantification of information leaks," in *Proc. 30th IEEE Symp. Security and Privacy*, Washington, DC, May 2009, pp. 141–153.

[3] B. Espinoza and G. Smith, "Min-entropy as a resource," *Information and Computation*, vol. 226, pp. 57–75, Apr. 2013.

[4] S. Takai and R. Kumar, "Verification and synthesis for secrecy in discrete-event systems," in *Proc. 2009 Amer. Control Conf.*, St. Louis, MO, Jun. 2009, pp. 4741–4746.

[5] J. Bryans, M. Koutny, and C. Mu, "Towards quantitative analysis of opacity," *Technical Reports Series, Newcastle University*, Nov. 2011.

[6] A. Saboori and C. N. Hadjicostis, "Probabilistic current-state opacity is undecidable," in *Proc. of 19th Int. Symp. Math. Theory Netw. and Syst. (MTNS '2010)*, Budapest, Hungary, Jul. 2010, pp. 477–483.

[7] ——, "Opacity verification in stochastic discrete event systems," in *Proc. 49th IEEE Conf. Decision Control*, Atlanta, GA, 2010, pp. 6759–6764.

[8] V. K. Garg, R. Kumar, and S. I. Marcus, "A probabilistic language formalism for stochastic discrete-event systems," *IEEE Trans. Autom. Control*, vol. 44, no. 2, pp. 280–293, Feb. 1999.

[9] A. Xie and P. A. Beerel, "Efficient state classification of finite-state Markov chains," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 17, no. 12, pp. 1334–1339, Dec. 1998.

[10] P. Brémaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation and Queues*. New York: Springer-Verlag, 1999.

[11] X. Wang and A. Ray, "A language measure for performance evaluation of discrete-event supervisory control systems," *Applied Math. Modelling*, vol. 28, no. 9, pp. 817–833, Sep. 2004.

[12] J. Chen and R. Kumar, "Polynomial test for stochastic diagnosability of discrete event systems," *IEEE Trans. Auto. Sci. and Eng.*, vol. 10, no. 4, pp. 969–979, Oct. 2013.

[13] A. Saboori and C. N. Hadjicostis, "Verification of initial-state opacity in security applications of discrete event systems," *Information Sciences*, vol. 246, pp. 115–132, October 2013.