# Investigating Stochastic Dependencies Between Critical Infrastructures

Sandra König
Austrian Institute of Technology GmbH, Center for Digital Safety & Security
Vienna, Austria
Email: sandra.koenig@ait.ac.at

Stefan Rass
Universität Klagenfurt, Institute of Applied Informatics, System Security Group
Klagenfurt, Austria
Email: stefan.rass@aau.at

***Abstract*—Critical infrastructures (CIs) are essential for the welfare and prosperity of a society, and failure of one infrastructure has a significant impact on our everyday life. However, a problem in one critical infrastructure is rarely local but often affects other infrastructures, e.g., limited availability of electricity affects hospitals, water providers and food suppliers. Even a partial failure of critical infrastructures has consequences that are hard to predict unless under stringent assumptions. Among other things, the damage on another infrastructure depends on the availability of substitutes. While such factors are mostly known, many external factors such as weather, temporary demand or load peaks are not precisely predictable so that a stochastic model is required to describe the state of an infrastructure. The state of each infrastructure is described by a random variable and changes its state according to a transition regime that depends on the state of other CIs but also the type of dependency. This yields a model of complex interdependencies with unknown dynamics where the state of a CI is determined by several Markov chains. Several ways exist to determine the actual state of the CI under several influences; the most conservative one is to assume the worst case (by applying the maximum principle). In this work, we provide a more general view that allows incorporating dependencies between input providers. Further, we discuss practical issues such as assessments from several experts and investigate chances for healing and total failure. An implementation of the model in R is used to illustrate how the model may be used in practice to estimate the states of a dependent CI due to limited availability of a provider. This paper describes a stochastic model of dependencies between CIs and discusses issues that arise when applying it.**

***Keywords-critical infrastructure; stochastic dependencies; Markov chain; risk propagation; copula.***

## I. Introduction

Many Critical infrastructures (CIs) are supply networks satisfying the basic needs of society, such as power, water, food, health care or transportation. These CIs naturally depend on one another, and recent developments such as the increased use of control systems increase these interdependencies. The type of dependency is manifold. For example, a hospital depends on water supply as it needs drinking water for staff and patients but also cooling water is necessary for smooth operation. A water provider needs electricity to keep its pumps running but also for the operation of a Supervisory Control and Data Acquisition (SCADA) system. These complex interdependencies are not exactly predictable a can thus be regarded as stochastic [1]. A core characteristic of today's CIs is the fact that a failure or limited availability of one CI often has a considerable impact on CIs depending on it. This has shown in recent years, for example, the disruption of electric power in California in 2001 [2] affected several other CIs, a significant power outage in Italy of about 12 hours [3] resulted in a financial damage of over one billion euros or the most recent hacking of the Ukrainian power grid caused a power outage of several hours [4]. Generally, such dependencies between CIs can be either *continuous*, as it is the case of electricity where a stable supply is required, or *instantaneous*, for example, if the CI's support is just required in an emergency (e.g., police or fire brigade).

In this work, we consider structures that mutually and *continuously* depend on input from several providers, such as water or electricity (see [5][6] for a more detailed discussion). Reduced or even missing supply from a critical provider may cause significant problems for an infrastructure. The actual damage naturally depends on the degree of failure of the provider but is also influenced by many other factors such as availability of substitutes (see [7] for work related to water supply). Especially consequences of reduced support are usually not precisely predictable, which is why we use a stochastic model to describe the condition (state) of a CI based on the states of its providers. These dependencies may be grouped depending on nature or importance of the relation. Such an abstract model can be applied to any infrastructure, as long as the dependency structure is known and can be classified qualitatively in terms of "how severe" a provider's outage is on a finite scale (say, from 1 to 5; see [8] for a discussion of this requirement in light of compliance, auditing, and monitoring). The model thus speaks about different "degrees of failure," where the particular meaning of such a "degree" is up to the specific characteristics of the CI (e.g., status 3 may represent different things or problems for a water provider than for a hospital). The basic model is not too complex by considering only dependencies between two infrastructures at a time and by grouping infrastructures into different classes with different

characteristics. However, it is also possible to take into account dependencies between providers of the same type that could be used as a substitute in case of limited availability. Making such a dependence among providers explicit is, for instance, doable with help of copulas that give the joint probability distribution as a function of the individual distributions. This method keeps the complexity still manageable while allowing for higher flexibility and more accuracy of the model.

### *Paper Outline*

The remainder of this article is organized as follows. After a recap of selected related work in Section II, Section III introduces a stochastic model for dependencies between critical infrastructures and its use in practice. Section IV analyzes the chances of total failure or normal functionality based on this model, and Section V extends the basic model towards dependencies between providers of a CI. Section VI shows a small example, and Section VII provides concluding remarks.

## II. Related Work

Several models exist on dependencies among critical infrastructures. In [9], a framework for addressing infrastructure interdependencies is presented that distinguishes five different classes of critical infrastructure interdependencies (including dependencies of information and communication technologies). Recent models consider random failure and stochastic dependencies, for example, a multi-graph model that analyze random failures and their effects on critical infrastructures [10]. Other models explicitly look at interdependencies of higher order to identify and assess the effect of failures not only for "consumers" but also for subsequent infrastructures in the dependency chain [11][12]. Cascading effects have been investigated in [13] using an Input-Output Inoperability Model (IIM) based on financial data and Hierarchical Holographic Modeling (HHM) [14] has been used to describe the diverse nature of CI networks and to analyze failures therein. Interdependency graphs are another popular tool for development of methodologies to describe the propagation of failures and cascading effects. Examples include the Cross Impact Analysis (CIA) [15], [16] or the Cross Impact Analysis and Interpretative Structural Model (CIS-ISM) [17]. The Input-Output Inoperability Model (IIM) [18], [19], [20] also provides a detailed view on interdependencies between CIs where linear equations model the consequences. However, these effects are in general measured according to economic aspects, which is only one (and not always the most important) point of view, especially when looking at critical infrastructures. Models that include a more detailed description of the infrastructures are based on Bayesian networks [21] as, for example, the Hierarchical Coordinated Bayes Model (HCBM) [22], [23], [24] or other approaches (cf. [25] and references therein). These models take into account effects of extreme events as well as events with only spars data but can also focus on technical dependencies, see, e.g., [26]. Our basic model is related to various approaches by simulation and co-simulation [27][28][29][30][31]. Typically, these are applicable when the analyst is much more informed about the infrastructure in question since the simulation depicts the internal dynamics (even up to the level of actual network packets to be exchanged). Our perspective is much more high-level and assumes the absence of this detailed information but rather assumes categorical valuations of interdependencies (cf. [5][32][33][34] for more comprehensive overviews), as it is often done in risk management. The stochastic dependency model used here can also be understood as a part of a classical risk management analysis [35].

More formal models include coupled complex systems [36] that are more exposed to large-scale failures or models as described in [37] that describe the increase and the decrease of random failures. Most popular among the stochastic models are Markov chain models. The Interdependent Markov Chain (IDMC) model describes cascading failures in interdependent infrastructures [38]. Conditional Markov transition models are applied in electric power grids [39] and a model including higher orders by adding memory to the Markov chain is presented in [40].

## III. Stochastic Dependencies Between Critical Infrastructures

We first describe a basic probabilistic model of dependencies between critical infrastructures in Section III-A and then show some issues when applying it. Potential application to the problem of measuring the resilience of critical infrastructures is given in [41]. Here we show that this model is capable of incorporating assessments from several experts, see Section III-B, and how it can be of use when applying risk management best practices to increase security in Section III-C. Finally, we sketch how the model can be implemented in Section III-D.

### *A. The Model*

Dependencies are often modeled through a directed graph where the nodes represent the various components, and a directed edge describes that the target node depends on the start node. This simple model can also be used to describe interdependencies between critical infrastructures. A high-level view on a system of CIs lets nodes represent the different CIs and a directed edge from CI 1 to CI 2 indicates that CI 2 depends on CI 1, e.g., a hospital depends on a water provider for drinking, but also waste management and fire extinguishing. A more detailed analysis lets nodes represent components of a CI, e.g., a pump or a well as parts of a water utility, and investigates how these depend on one another and other CIs. In either case, the visualization of dependencies provides a basis to understand how limitations in one provider affect dependent CIs and how this effect changes over time. A simulation tool for these cascading effects that is also able to distinguish between short- medium and longtime dependencies is presented in [42]. CIs that support other CIs are called *provider* in the following. In the basic model of stochastic dependencies among CIs [1] we represent the CIs as directed graphs whose nodes, also called *components*, can be in various

states that represents their degree of functionality. Whenever one component changes its change from 1 ("working properly") into a state that represents limited functionality (up to state $k$ that represents total failure), this may cause a state change in every CI depending on it.

More formally, suppose that a CI $S$ depends on a set of $n$ providers, enumerated as $P_1, \ldots, P_n$ (all being perhaps themselves CIs). We assume that $S$ will not endogenously experience any state changes since keeping $S$ up and running is a matter of $S$'s business continuity management. The model we describe thus centers on *exogenous* triggers for $S$ to change its state, namely upon problems with one of $S$'s providers $P_1, \ldots, P_n$, drawn as the lower layer of nodes in the bipartite right graph shown in Figure 1. Specifically, we let each of them maintain its own state of functionality, which is communicated (or generally observed) by $S$, and $S$ can react on a change. This change is governed probabilistically by the state of the provider, or more formally, let $S_i$ be the condition, i.e., state, that $P_i$ can cause for $S$. This is a random variable distributed over the state space $\{1, 2, \ldots, k\} \simeq \{\text{ok}, \ldots, \text{total failure}\}$, and described as a conditional distribution $\Pr(S_i = x | P_i = y)$, where $P_i = y \in \{1, \ldots, k\}$ is the current state of provider $P_i$, and $x \in \{1, 2, \ldots, k\}$ is the state that $P_i$ may drive $S$ into. The exact way in which $S$ now depends on the provider $P_i$ can then be specified by a value $p_{i,x,y}$, which can, for example, be set following considerations like these:

- $P_i$ is of vital importance for $S$, so if $i$ is in a bad condition, $S$ is highly likely to be in trouble as well. Thus, $\Pr(S_i = 5 | P_i = 5) \approx 1$, indicating that $S_i$ will become unavailable if $P_i$ becomes unavailable, since there may be no compensation for $P_i$'s service.
- $P_i$ may only be of minor importance, and an outage of $P_i$ can be bridged by backup resources that $S$ maintains. In that case, we could define $\Pr(S_i = 5, P_i = 5) \approx 0$, modeling that an outage of $S$ is very unlikely even if $P_i$ no longer provides its service.

More fine-grained considerations like the above examples are possible and in a practical instance depend on the application at hand. We leave the two examples here only for illustration and now turn back to the formal description of the model.

Since the dependence of $S$ on two providers may be quite individually different, we internally let $S$'s state be a vector of random states $(S_1, \ldots, S_n)$, each $S_i$ determined by the corresponding provider $P_i$. These nodes $S_1, \ldots, S_n$ correspond to the colored upper layer in the bipartite graph shown in Figure 1, and the actual state of $S$ that it communicates as a provider to other CIs is a single value in $\{1, 2, \ldots, k\}$ compiled (aggregated) from the vector, i.e., node states, $S_i$.

This aggregation follows the maximum principle of system security, defining the risk in a system by the highest individual risk therein. Likewise, we compute the state of $S$ as $\max\{S_1, \ldots, S_n\}$, corresponding to $S$ being in trouble if at least one of its providers reports a bad condition (by having a state close or equal to $k$). In terms of Boolean or fuzzy logic, we would thus define $S$'s state as the (logical) OR of its provider's states. Adopting this view but changing the perspective, any more complex aggregation function to define the state of $S$ from the variables $S_1, \ldots, S_n$ is imaginable, including the use of copulas [43] or triangular norms (from multivalued- and fuzzy logic [44]); we postpone this discussion until Section V. The simulation model studied in this work uses the max-aggregation hereafter.

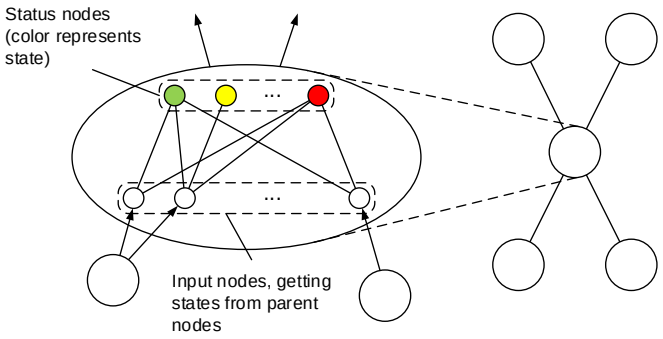


Fig. 1. Model of the inner structure of a critical infrastructure [1]

With $k$ states for each of the $n$ providers, and $k$ states of $S$ itself, the overall specification of the state transition is a set of $n$ stochastic $(k \times k)$ transition matrices with entries being the conditional likelihoods as described above. According to these transition matrices, provider $i$ yields a state $S_i$ of the dependent CI. In case different providers yield different states, the final state $S$ is determined by $S = \max\{S_1, \ldots, S_n\}$, i.e., we consider the worst case.

While a stochastic model is convenient (actually natural) to describe uncertain consequences, it is often challenging to put it in practice. The main issue in this regard is estimation of transmission probabilities since experts often feel uncomfortable or feel unable to provide concrete and reliable values. Several aids exist, however:

1) allow an expert to give qualitative values, e.g., on a 5-tier scale
2) ask an expert to tell a level of confidence with every estimate (also qualitative, such as "very sure", "somewhat unsure" or "'just guessing')
3) ask several experts for their individual (subjective) assessments

The last point may allow experts assess only some transitions (depending on their expertise) but also raises the question of how to deal with several expert opinions. We will focus on this below. The second point seems to blow up the data needed but can be put into practice in a way that actually in most cases reduced the amount of input data. For each state of the provider, it is necessary to find an entire distribution over all possible states of the dependent CI. This can be done by determining the most likely value as a subjective prediction, and interpreting the level of confidence in the told opinion as some kind of variance. Practically, that means that for confidence "very high" we choose a distribution that puts mass 1 on the predicted value and 0 elsewhere, while we choose a uniform distribution over all state for a confidence "just guessing". Intermediate assessments such as "somewhat unsure" put some probability mass on states close to the

specified one but no positive probability on values that are too far from the opinion. For the case of three different states of a CI this mapping is illustrated in [45], including a short discussion on the interpretation of "totally sure" assessments. For the case of five possible states (as used in our example later on) we get the mapping given in Table I, where the resulting vectors make up one row of the corresponding transitions matrix.

TABLE I. DISTRIBUTION OVER STATES OF DEPENDENT CI BASED ON EXPERT ASSESSMENT (PREDICTION, CONFIDENCE)

| prediction | very sure | somewhat unsure | just guessing |
|---|---|---|---|
| 1 | (1,0,0,0,0) | (2/3, 1/3, 0, 0, 0) | (1/5,1/5,1/5,1/5,1/5) |
| 2 | (0,1,0,0,0) | (1/4, 2/4, 1/4, 0, 0) | (1/5,1/5,1/5,1/5,1/5) |
| 3 | (0,0,1,0,0) | (0, 1/4, 2/4, 1/4, 0) | (1/5,1/5,1/5,1/5,1/5) |
| 4 | (0,0,0,1,0) | (0, 0, 1/4, 2/4, 1/4) | (1/5,1/5,1/5,1/5,1/5) |
| 5 | (0,0,0,0,1) | (0, 0, 0, 1/3, 2/3) | (1/5,1/5,1/5,1/5,1/5) |

Assessments of this kind yield an entire row of the transmission matrix (i.e., a discrete distribution) and we will denote it by $F$ to represent this fact.

### B. Several Expert Opinions

So far we have assumed that each dependency has been assessed by one expert only. We stress that, however, not necessarily the same person is required to rate all dependencies. In case we have more than one expert assessment for a connection, all opinions should be taken into account to increase the data quality underlying the subjective assessments, as well as to reduce the pressure on each expert to be responsible solely for the given assessment. In our setting, a number of $K$ experts opinions yield to multiple distributions over the possible states of the dependent CI. We denoted these as $F_1, \ldots, F_K$ and aggregate the distribution to find the estimate

$$F = \alpha_1 \cdot F_1 + \ldots + \alpha_K \cdot F_K$$

for one edge in the bipartite graph in Figure 1, where $\alpha_1, \ldots, \alpha_K$ satisfy

$$\alpha_1 + \ldots + \alpha_K = 1.$$

The parameter $\alpha_i$ can be interpreted as the weight (influence) of expert $i$'s opinion in the overall assessment. In case we do not distinguish between different expertise, we choose uniform weights, i.e., $\alpha_i = 1/K$ for all $i$.

### C. Impact Estimation as a Part of Risk Management

One step in risk analysis and risk management [46][35] is to estimate the impact due to a security incident. There exist several ways to do that, and the choice of a specific method depends on the situation at hand, see, e.g., [47] for co-simulation applied to power distribution grids, [48] for an application of agent-based simulation or [49] for the spreading of ransomware. This impact estimation is not only needed for an analysis of effects of an incident (such as a malware attack) but can also help to test the use of countermeasures (such as patching a computer). Whenever actions are taken to reduce the damage on a CI due to a problem in a provider, this changes the dependency between the two, in particular, it reduced the probability that the dependent CI changes into a severe state (if the countermeasure is effective). The network itself does not change but the effect of reduced availability of a provider on the dependent CI changes, which yields a different transitions matrix. The simulation is then be rerun with a different set of transition matrices to see if the resulting losses reduce. These estimates of damages for various scenarios build up a generalized payoff matrix that allows finding an optimal way to protect the system at hand. The important point is that the assessment as we outline here is exactly the same as what is done along a conventional risk management process anyway: following standard frameworks like the IT Grundschutz (by the German federal office for information security (BSI) [50]), or the ISO 27k standard, a typical step is an assessment of how assets or components depend on one another. The pure information of a dependence then naturally defines the dependency graph topology. Our method then goes further in asking what would happen to one component if another component fails. That is, we propose a mere "additional use" of the artifacts from risk management in the here proposed simulation framework to aid the impact assessment to gain some "objectivity."

### D. Simulation

The stochastic dependency model between critical infrastructures has a straightforward implementation in software such as the freeware `R`. This makes it handy to use in any field due to the high interoperability of `R` with other systems.

The simulation starts with an incident affecting one node, which subsequently (directly and indirectly) triggers descendant CIs to change their status according to the likelihoods in their inner bipartite graphs. In that way, the simulation reveals how far an incident will propagate through the network of CIs (within the runtime of the simulation), and can thus be used to estimate the effect a problem in one component has on a specific critical infrastructure or generally on other components. Further, it allows an empirical estimation of the number of components that are in a critical state (i.e., reach the highest status $k$) or the relative frequency of one specific CI being in a critical state.

More explicitly, we model the network of infrastructures as a graph $G = (V, E)$ with vertices $v \in V$ that represent the infrastructures and edges $e \in E$ representing the dependencies between them. A common difficulty in specifying such probabilistic models is the issue of where to get the conditional probabilities from (that we already mentioned in Section III-A, along with hints on how to think about these values). To relieve this practical challenge, we let the conditional likelihood specification be discrete and replace the poll for probabilities by the question to specify, resp. assign, a certain *edge class* $c$ instead (the edge again being one in the bipartite inner graph modeling a CI; cf. Figure 1). An edge classification is hereby chosen from a set of candidates $\{1, 2, \ldots, C\}$. Each edge class represents a fixed type of inner or mutual dependency which carries the sought probabilities with them. The information in

the edge class can also include different levels of importance of a CI for its successor CI (provider consumer dependency), and other explanatory information or data useful for the simulation. Each edge $v \to w$ is then associated with a representative number for its class $c$ that carries an attribute being the probability for the simulation.

This allows the model parameterization to be done upfront and independently of the concrete CI, and eases matters of model parameterization in the absence of empirical data to estimate conditional probabilities. Depending on this class $c$ the state $i$ of $v$ influences the state of $w$ through a multinomial distribution MN($p_{i,c}$). That is, the $j$-th component of the vector $p_{i,c}$ gives the probability that $w$ will be in state $j$ in this situation. Figure 2 shows an algorithm in pseudo-code, which simulates $T$ time steps.

```
1: t ← 0
2: while t < T
3:   for each node v, set N(v) = {w ∈ V : (v, w) ∈ E}
4:     for each neighboring node w ∈ N(v)
5:       let c be the class of v → w,
6:       let i be the current state of node v,
7:       draw the status of w from MN(p_{i,c})
8:       t ← t + 1.
9:     endfor
10:  endfor
11:endwhile
```

Fig. 2. Simulation algorithm

Just as the input, the result of this simulation is a network of connected critical infrastructures where each CI is in one specific state. For a better understanding of the results, visualization with use color codes (e.g., ranging from green to indicate a working state to red, alerting about a critical condition) is helpful. Numerically, the results of the simulation can be summarized as a table that lists how many components are on average in any of the possible states.

An implementation of the dependency model in the event simulation tool `OMNeT++` is presented in [42]. The prototype described therein enables modeling the network of CIs as a directed graph whose nodes are colored to represent its state (ranging from green to red to represent several levels of functionality). External events can trigger the simulation by changing the states of one or more components. The propagation through the network is implemented as a message exchange over a fictitious communication channel. After a predefined running time, the tool yields a chronological record of the state changes for each component of the network. An illustrative example is included.

## IV. Chances of Healing and Total Failure

Technically speaking, the changes between the states of the CIs based on the state of its provider is described by a Markov chain, whose states correspond to the states of the CI. We adopt an ordered numeric representation for the nominal scale of health, ranging from "good" $\simeq$ state 1, up to state "failure" $\simeq$ state $k$. The rich theory related to Markov chains then enables us to compute the chances that a CI fails completely (i.e., is in state $k$) or remaining in good shape (i.e., is in state 1) for a certain period. We will denote the $i$-th unit vector by $u_i$ to represent the situation where an asset is in state $i$ with likelihood 1.

As before, let $S_i$ denote the state of the dependent CI due to the state of its $i$-th provider. In the classical model from Section III-A, we assumed a worst case scenario, i.e., the overall state $S$ of the dependent CI is $S = \max\{S_1, \ldots, S_n\}$. Under this assumption, a CI is in state 1 only if every provider causes a switch to state 1 (or a stay in state 1). That is:

$$\Pr(S = 1) = F_{S_1,\ldots,S_n}(1, \ldots, 1), \tag{1}$$

where $F$ is the joint distribution over all CI states. Unfortunately, this only simplifies in the case of i.i.d. variables (where we get a product). However, it can generally be decomposed into the individual, i.e., marginal, distributions unconditionally describing the state of each CI, plus an outer function capturing the interdependence. This outer function is a *copula* and essentially is the mathematical function describing the dependencies visualized in the graph $G = (V, E)$. Formally, we have $F_{S_1,\ldots,S_n}(1, \ldots, 1) = C(F_{S_1}(1), \ldots, F_{S_n}(1))$. In case of stochastic independence, the last term simplifies to $\prod_{i=1}^{n} \Pr(S_i = 1)$, i.e., we have $C(x_1, \ldots, x_n) = x_1 \cdot x_2 \cdots x_n$.

Since we assume that at the beginning every asset is working properly (i.e., in state 1) we know that the likelihood of returning to that state after $t$ time steps is $u_1 P^t$. Further, the Markov Chain returns to the starting distribution if it is a limiting distribution, i.e., if $u_1 = u_1 P$ holds. So, equation (1) is fulfilled if the vector $u_1$ is a stationary distribution of each of the involved Markov chains.

In case $k$ is an absorbing state (i.e., if there is no recovery from a failure, say, if the CI is irreparably destroyed) for all involved Markov chains, we find that $\Pr(S_i = k) = 1$ and thus $\Pr(S = k) = 1$. Otherwise, the probabilities $\Pr(S_i = k)$ can be determined by the law of total probability

$$\Pr(S_i = k) = \sum_{j=1}^{k} \Pr(S_i = k | v_i = j) \cdot \Pr(v_i = j)$$

where $v_i$ is the $i$-th provider that yields to state $S_i$ according to the transition probability $p_{jk} = \Pr(S_i = k | v_i = j)$. The probabilities $\Pr(v_i = j)$ depend in turn on the providers of the provider $v_i$, which makes an explicit analysis challenging.

## V. Dependencies Between Providers

The basic model introduced in Section III assumes that providers are independent and the effect of a limited availability is in not influenced by the state of other providers. However, the effect of a problem in one provider might be limited as long as there is another one of the same type that is fully working. The basic model can be extended to capture a certain degree of dependency between providers, e.g., if they are of the same or very similar kind and can be used as substitutes. This particularly applies to the situation where

a CI has contracts with several providers to reduce the damage in case the provider is not available.

Let $S_1, \ldots, S_n$ denote the states of a CI due to the state of the corresponding providers according to the transition matrices. Until now we have assumed that the state $S$ communicated to other CIs is determined by the maximum principle, yielding

$$S = \max\{S_1, \ldots, S_n\}. \quad (2)$$

However, this is a very conservative view, as it ignores the fact that due to dependencies between providers the overall state may be better than caused by a single provider. For example, some critical infrastructures use several providers of the same type to avoid this strong dependency, e.g., they have contracts with more than one telecommunication provider. In that case, we may replace the relation given in (2) by the more general form

$$S = f(S_1, \ldots, S_n)$$

where $f$ is any function that aggregates the $n$ values into one "overall" state. If we know about dependencies between providers of the same type (e.g., one may be a substitute) we can reformulate this as

$$S = f(C(S_{t_1}, \ldots, S_{t_y}), S_{r_1}, \ldots, S_{r_m}) \quad (3)$$

where providers $t_1, \ldots, t_y$ are of the same type and thus assumed to be dependent to some extent (while providers $r_1, \ldots, r_m$ are not of this type). We model their joint distribution with a suitable copula or a more general function. Possibilities include the following here at least:

- $\min$-operator: this is a copula, and in setting the overall state of a CI to the minimum state of all its providers, our convention that "healthy" has a state representation number less than that for "failure", we end up with the following semantic:
  A CI will not change into failure state unless *all* its providers have failed. This is an "OR-aggregation" since the CI remains intact if any of its providers is intact.
- $\max$-operator: this provides the reverse semantic as above since a CI will go into failure state if at least one of its providers fails. Logically, they are in that sense "AND-connected".
- Combinations of the two, where a complex Boolean term can have each its connectives represented by an artificial intermediate node with min or max aggregations to model AND or OR operations therein. We leave this as a simple extension and not go into formal details at this point[1].

The decomposition in equation (3) can be further refined by using several copulas for several types of providers, i.e.,

$$S = f(C_t(S_{t_1}, \ldots, S_{t_y}), \ldots, C_s(S_{s_1}, \ldots, S_{s_z}))$$

for copulas $C_t, \ldots, C_s$.

[1]Extending this Boolean approach further, we can even model a logical negation by setting the conditional probabilities for a state change accordingly: suppose that $S$ has only a single provider $P$, then we can have S to "invert" the state of $P$ by specifying that $\Pr(S = 5|P = 1) = 1$ and $\Pr(S = 1|P = 5) = 1$, meaning that $S$ is in a good/bad condition if and only if $P$ is in a bad/good condition. Intermediate states $k = 2, \ldots, 4$ would herein be triggered with the likewise defined conditional probabilities.

## VI. An Illustrative Example

To illustrate the application of the presented model, we evaluate a small example of high-level dependencies between a few critical infrastructures. When applying the model to real use cases, one needs to decide on the granularity of the network representation that corresponds to the degree of detail in the description. If the aim of the analysis is getting an overview of dependencies between critical infrastructures (as we do in this small example), then each node represents a single CI. Dependencies are assessed on a general level, and the results are accordingly general. If more data on a CI are available, then this CI can, in turn, be represented as an entire network where components represent significant components of the network (this approach has been illustrated in [45]).

Let us consider a subnetwork of dependent CIs, which consists of a hospital that depends on a water provider, an electricity provider as well as transportation infrastructures (roads). The dependencies between the different components in the network are classified as either "minor", "normal" or "critical" depending on how important the service provisioning is for the CI. In this small example, we classified input from the electricity provider as "normal" (as we assume existence of an emergency power system), input from a water provider as "critical" (substitution by bottled water is usually just possible for a limited period of time, but substituting water for waste management or fire extinguishing systems is even more problematic to replace; we do not cover these details hereafter), and the transport connection as "minor", since even if roads are temporarily congested or blocked, aerial transportation remains possible for critical patients.

The effects of an outage of each provider may be different, and to ease matters as before with the edge classes, we propose specifying transition matrices *per dependency criticality level*, i.e., transitions from working into a failure state become more likely the more critical the dependency on the respective provider is. Consequently, we will below specify three transition matrices for dependencies of levels "minor", "normal" and "critical".

In [1], we considered transmission matrices that are estimated with certainty. Dropping this assumption now, we show a case where experts are not certain about their assessments, i.e., they provide information as described in Section III-B. Further, we assume that two experts do the assessment for the critical dependency on water with potentially different backgrounds. For this example, we consider 5 possible states for each node, where 1 represents the situation where everything works smoothly, while 5 stands for serious problems including total failure.

Suppose for example that data, as shown in Table II, has been collected, where we represent the confidence levels by numbers ranging from 1 ("totally unsure") to 3 ("totally sure").

We assign the same weight (importance) to both experts doing assessments of the dependency on water, that is we chose $\alpha_1 = \alpha_2 = 0.5$. This yields individual transmission matrices $T_{\text{dependency-criticality-level}} = (t_{ij})_{i,j=1}^{5}$, in which the $ij$-th

(a) EXPERT ASSESSMENT FOR ELECTRICITY

| State of Provider | Predicted Value | Confidence Level |
|---|---|---|
| 1 | 1 | 2 |
| 2 | 1 | 1 |
| 3 | 1 | 1 |
| 4 | 4 | 1 |
| 5 | 4 | 1 |

(b) EXPERT ASSESSMENT FOR TRANSPORT

| State of Provider | Predicted Value | Confidence Level |
|---|---|---|
| 1 | 1 | 2 |
| 2 | 1 | 2 |
| 3 | 1 | 2 |
| 4 | 1 | 1 |
| 5 | 1 | 1 |

(c) FIRST EXPERT ASSESSMENT FOR WATER

| State of Provider | Predicted Value | Confidence Level |
|---|---|---|
| 1 | 1 | 1 |
| 2 | 2 | 1 |
| 3 | 4 | 1 |
| 4 | 4 | 2 |
| 5 | 5 | 2 |

(d) SECOND EXPERT ASSESSMENT FOR WATER

| State of Provider | Predicted Value | Confidence Level |
|---|---|---|
| 1 | 1 | 2 |
| 2 | 2 | 1 |
| 3 | 4 | 1 |
| 4 | 5 | 2 |
| 5 | 5 | 3 |

TABLE II. ASSESSMENTS FOR ELECTRICITY AND TRANSPORT

entry corresponds to the conditional likelihood $t_{ij} := \Pr(\text{CI gets into state } j \mid \text{provider is in state } i$ and given the respective criticality of the dependency). We choose

$$T_{minor} = \begin{pmatrix} 2/3 & 1/3 & 0 & 0 & 0 \\ 2/3 & 1/3 & 0 & 0 & 0 \\ 2/3 & 1/3 & 0 & 0 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \end{pmatrix},$$

$$T_{normal} = \begin{pmatrix} 2/3 & 1/3 & 0 & 0 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \end{pmatrix}$$

and

$$T_{critical} = \begin{pmatrix} 13/30 & 8/30 & 3/30 & 3/30 & 3/30 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 8/80 & 8/80 & 18/80 & 28/80 & 18/80 \\ 0 & 0 & 3/24 & 10/24 & 11/24 \\ 0 & 0 & 0 & 1/6 & 5/6, \end{pmatrix},$$

where $T_{critical}$ is the linear combinations of the two matrices induced by the two expert assessments.

Figure 3 (left side) displays the dependencies graphically, with arrows annotated according to the criticality of the dependency. The right part of Figure 3 shows how the inner model of Figure 1 corresponds to a dependency and is instantiated according to the matrices above. For example, if the dependency's criticality is "minor" and the respective provider is in state 4 (i.e., it has rather serious problems), this will yield to a state 5 of the critical infrastructure that depends on it with a likelihood of $1/5 = 0.2$.

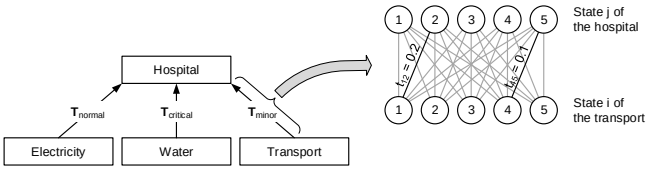


Fig. 3. Example instance

The results of our analysis can be interpreted in (at least) two ways. On one hand, it provides information on a specific node in the network (such as it frequency of failure), including information about which node caused the failure. On the other hand, it provides an overview on the average number of nodes in a specific state, so in particular on the expected number of failing components.

Initially, we assumed that all components are in state 1 (i.e., operate smoothly) except for the water provider that is in state 2 facing some (maybe temporary) problems. This scenario yielded to a critical state for the hospital in 201 out of 1000 cases. Note that in this example, this critical state can only be caused by the state of the water provider since a CI of normal or even minor importance will never cause a critical level while being in state 1 (i.e., both entries in the transition matrices are zero). In a more elaborated example with numerous dependencies and other components facing problems state changes may be caused by other components as well. The simulation may then be used to track which provider caused the worse state (if this information is stored as well). This information may help providers to identify critical dependencies and indicates where future investments may be useful (e.g., it might make sense to have a substitute for a provider that often causes problems).

Table III shows the average number of nodes (CIs) that are in each of the 5 possible states. This gives a general overview on the situation of the entire network, e.g., it can be seen that on average 1.921 nodes are in a state worst than 1 (indicating that they have a problem). Further, it gives an estimate of the number of components in the worst case, which might help planing resources needed to fix problems. Additionally, such information may be used to measure resilience of a component [41].

TABLE III. AVERAGE NUMBER OF AFFECTED NODES DUE TO INCREASED LEVEL OF CRITICALITY

| Criticality | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Nodes | 2.079 | 1.313 | 0.201 | 0.206 | 0.201 |

## VII. CONCLUSION

A basic stochastic model of dependencies between critical infrastructures can capture issues such as the impossibility of exact predictions in a network of interdependent critical infrastructures. It describes the degree of availability of a provider by different states and let this potentially cause a state change in the dependent CI. In the simplest case, the

actual state of a CI depending on many providers is assumed to be the worst of all states caused by any supporting CI, corresponding to a worst-case view.

In this work, we addressed practical issues when applying the model by describing how experts assessment can be incorporated without the need of an agreement between several experts. Instead, opinions of different experts might be combined, and each expert is asked to rate the confidence in his prediction. Further, we illustrated how to implement the simulation either in the statistical software `R` or in the event simulation tool `OMNeT++` and exemplified the approach with a small example.

The basic model can be extended to take into account dependencies between providers. Further, we explained how the model fits into a risk analysis as a tool to estimate the impact of an incident affecting parts of a network of CIs and how it can be used to compare the current situation with a scenario in which countermeasures have been implemented before.

## ACKNOWLEDGMENT

This work was done in the context of the project "Cross Sectoral Risk Management for Object Protection of Critical Infrastructures (CERBERUS)", supported by the Austrian Research Promotion Agency under grant no. 854766.

## REFERENCES

[1] S. König and S. Rass, "Stochastic dependencies between critical infrastructures," in *SECURWARE 2017: The Eleventh International Conference on Emerging Security Information, Systems and Technologies*, IARIA, Ed., 2017, pp. 106–110.

[2] S. Fletcher, "Electric power interruptions curtail California oil and gas production," *Oil Gas Journal*, 2001.

[3] M. Schmidthaler and J. Reichl, "Economic Valuation of Electricity Supply Security: Ad-hoc Cost Assessment Tool for Power Outages," *ELECTRA*, no. 276, pp. 10–15, 2014.

[4] J. Condliffe, "Ukraine's Power Grid Gets Hacked Again, a Worrying Sign for Infrastructure Attacks," 2016, URL: https://www.technologyreview.com/s/603262/ukraines-power-grid-gets-hacked-again-a-worrying-sign-for-infrastructure-attacks/ [accessed: 2017-07-26].

[5] *Managing the Complexity of Critical Infrastructures: A Modelling and Simulation Approach*, ser. Studies in Systems, Decision and Control. Cham: Springer, 2016, vol. 90.

[6] R. Klein, E. Rome, C. Beyel, R. Linnemann, W. Reinhardt, and A. Usov, "Information modelling and simulation in large interdependent critical infrastructures in irriis," in *Critical Information Infrastructure Security: Third International Workshop, CRITIS 2008, Rome, Italy, October13-15, 2008. Revised Papers*. Berlin, Heidelberg: Springer, 2009, pp. 36–47.

[7] E. Luiijf, M. Ali, and A. Zielstra, "Assessing and improving SCADA security in the Dutch drinking water sector," *International Journal of Critical Infrastructure Protection*, vol. 4, no. 3-4, pp. 124–134, 2011.

[8] A. Abou El Kalam and Y. Deswarte, "Critical infrastructures security modeling, enforcement and runtime checking," in *Critical Information Infrastructure Security: Third International Workshop, CRITIS 2008, Rome, Italy, October13-15, 2008. Revised Papers*. Berlin, Heidelberg: Springer, 2009, pp. 95–108.

[9] S. Rinaldi, J. Peerenboom, and T. Kelly, "Identifying, Understanding, and Analyzing Critical Infrastructure Interdependencies," *IEEE Control Systems Magazine*, pp. 11–25, 2001.

[10] N. K. Svendsen and S. D. Wolthusen, "Analysis and Statistical Properties of Critical Infrastructure Interdependency Multiflow Models," in *2007 IEEE SMC Information Assurance and Security Workshop*, June 2007, pp. 247–254.

[11] M. Theoharidou, P. Kotzanikolaou, and D. Gritzalis, "Risk assessment methodology for interdependent critical infrastructures," *International Journal of Risk Assessment and Management*, vol. 15, no. 2-3, pp. 128–148, 2011, URL: http://www.inderscienceonline.com/doi/abs/10.1504/IJRAM.2011.042113 [accessed: 2018-11-24].

[12] P. Kotzanikolaou, M. Theoharidou, and D. Gritzalis, "Assessing $n$-order dependencies between critical infrastructures," *International Journal of Critical Infrastructures*, vol. 9, no. 1-2, pp. 93–110, 2013, URL: http://www.inderscienceonline.com/doi/abs/10.1504/IJCIS.2013.051606 [accessed: 2018-11-24].

[13] R. Setola, S. De Porcellinis, and M. Sforna, "Critical Infrastructure Dependency Assessment Using the Input-Output Inoperability Model," *International Journal of Critical Infrastructure Protection (IJCIP)*, vol. 2, pp. 170–178, 2009.

[14] Y. Y. Haimes, "Hierarchical Holographic Modeling," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 11, no. 9, pp. 606–617, 1981.

[15] T. J. Gordon and H. Hayward, "Initial experiments with the cross impact matrix method of forecasting," *Futures*, vol. 1, no. 2, pp. 100–116, 1968.

[16] M. Turoff, "An alternative approach to cross impact analysis," *Technological Forecasting and Social Change*, vol. 3, pp. 309–339, 1971.

[17] V. A. Bañuls and M. Turoff, "Scenario construction via Delphi and cross-impact analysis," *Technological Forecasting and Social Change*, vol. 78, no. 9, pp. 1579–1602, 2011.

[18] Y. Y. Haimes and J. Pu, "Leontief-Based Model of Risk in Complex Interconnected Infrastructures," *Journal of Infrastructure Systems*, vol. 7, no. 1, pp. 1–12, 2001.

[19] J. R. Santos and Y. Y. Haimes, "Modeling the demand reduction input-output (I-O) inoperability due to terrorism of interconnected infrastructures," *Risk Analysis: An Official Publication of the Society for Risk Analysis*, vol. 24, no. 6, pp. 1437–1451, 2004.

[20] R. Setola, S. De Porcellinis, and M. Sforna, "Critical Infrastructure Dependency Assessment Using the Input-Output Inoperability Model," *International Journal of Critical Infrastructure Protection*, vol. 2, pp. 170–178, dec 2009.

[21] M. I. Jordan, Ed., *Learning in graphical models*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1999.

[22] Y. Y. Haimes, J. Santos, K. Crowther, M. Henry, C. Lian, and Z. Yan, "Risk Analysis in Interdependent Infrastructures," in *Critical Infrastructure Protection*, ser. IFIP International Federation for Information Processing. Springer, Boston, MA, 2007, pp. 297–310.

[23] Z. Yan, Y. Y. Haimes, and M. G. Wallner, "Hierarchical coordinated Bayesian model for risk analysis with sparse data," Baltimore, USA, 2006.

[24] Y. Y. Haimes, J. Santos, K. Crowther, M. Henry, C. Lian, and Z. Yan, "Risk Analysis in Interdependent Infrastructures," in *Critical Infrastructure Protection*. Boston, MA: Springer US, 2007, vol. 253, pp. 297–310.

[25] T. Schaberreiter, S. Varrette, P. Bouvry, J. Röning, and D. Khadraoui, *Dependency Analysis for Critical Infrastructure Security Modelling: A Case Study within the Grid'5000 Project*. Berlin, Heidelberg: Springer, 2013, pp. 269–287.

[26] T. Schaberreiter, S. Varrette, P. Bouvry, J. Röning, and D. Khadraoui, "Dependency Analysis for Critical Infrastructure Security Modelling: A Case Study within the Grid´5000 Project," in *Security Engineering and Intelligence Informatics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, vol. 8128, pp. 269–287.

[27] R. Caire, J. Sanchez, and N. Hadjsaid, "Vulnerability analysis of coupled heterogeneous critical infrastructures: A Co-simulation approach with a testbed validation," in *IEEE PES ISGT Europe 2013*. IEEE, 2013, pp. 1–5.

[28] R. Jaromin, B. Mullins, J. Butts, and J. Lopez, "Design and Implementation of Industrial Control System Emulators," in *Critical Infrastructure Protection VII*, ser. IFIP Advances in Information and Communication Technology. Berlin, Heidelberg: Springer, 2013, vol. 417, pp. 35–46.

[29] H. Lin, S. Sambamoorthy, S. Shukla, J. Thorp, and L. Mili, "Power system and communication network co-simulation for smart grid applications," in *ISGT 2011*. IEEE, 2011, pp. 1–6.

[30] M. Faschang, F. Kupzog, R. Mosshammer, and A. Einfalt, "Rapid control prototyping platform for networked smart grid systems," in *Proceedings IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*. Vienna, Austria: IEEE, 2013, pp. 8172–8176.

[31] M. Findrik, P. Smith, J. H. Kazmi, M. Faschang, and F. Kupzog, "Towards secure and resilient networked power distribution grids: Process and tool adoption," in *Smart Grid Communications (SmartGridComm), 2016 IEEE International Conference on*. Sidney, Australia: IEEE Publishing, 2016, pp. 435 – 440.

[32] J. Butts, *Critical Infrastructure Protection VII: 7th IFIP WG 11. 10 International Conference, ICCIP 2013, Washington, DC, USA, March 18-20, 2013, Revised Selected Papers*, ser. IFIP Advances in Information and Communication Technology. Berlin/Heidelberg: Springer, 2013, vol. v.417.

[33] S. M. Rinaldi, "Modeling and simulating critical infrastructures and their interdependencies," in *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*. IEEE, 2004, pp. 1–8.

[34] E. Wiseman, "Critical Infrastructure Protection and Resilience Literature Survey: Modeling and Simulation," URL: https://apps.dtic.mil/dtic/tr/fulltext/u2/1003598.pdf [accessed: 2018-11-24].

[35] S. Rass and S. Schauer, Eds., *Game Theory for Security and Risk Management: From Theory to Practice*, ser. Static & dynamic game theory : foundations & applications. Cham, Switzerland: Birkhäuser, 2018.

[36] B. A. Carreras, D. E. Newman, P. Gradney, V. E. Lynch, and I. Dobson, "Interdependent Risk in Interacting Infrastructure Systems," in *40th Annual Hawaii International Conference on System Sciences, 2007. HICSS 2007*, Hawaii, USA, 2007, pp. 112–112.

[37] N. Svendsen and S. Wolthusen, "Analysis and Statistical Properties of Critical Infrastructure Interdependency Multiflow Models." IEEE, 2007, pp. 247–254.

[38] M. Rahnamay-Naeini and M. M. Hayat, "Cascading Failures in Interdependent Infrastructures: An Interdependent Markov-Chain Approach," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 1997–2006, 2016.

[39] Z. Wang, A. Scaglione, and R. J. Thomas, "A Markov-Transition Model for Cascading Failures in Power Grids," in *2012 45th Hawaii International Conference on System Sciences*, 2012, pp. 2115–2124.

[40] S. Wu and M. T. Chu, "Markov chains with memory, tensor formulation, and the dynamics of power iteration," *Applied Mathematics and Computation*, vol. 303, pp. 226–239, 2017.

[41] S. König, T. Schaberreiter, S. Rass, and S. Schauer, *A Measure for Resilience of Critical Infrastructures*, 2018, (in press).

[42] T. Grafenauer, S. König, S. Rass, and S. Schauer, "A simulation tool for cascading effects in interdependent critical infrastructures," in *International Workshop on Security Engineering for Cloud Computing (IWSECC 2018)*, 2018, (in press).

[43] R. B. Nelsen, *An Introduction To Copulas*, ser. Lecture Notes in Statistics 139. Springer, 1999.

[44] T. Ross, J. M. Booker, and W. J. Parkinson, *Fuzzy Logic and Probability Applications: Bridging the gap*. ASA SIAM, 2002.

[45] S. König, T. Grafenauer, S. Rass, and S. Schauer, "Practical risk analysis in interdependent critical infrastructures - a How-To," in *SECURWARE 2018: The Twelfth International Conference on Emerging Security Information, Systems and Technologies*. IARIA, 2018, pp. 150–157.

[46] S. Schauer, "A Risk Management Approach for Highly Interconnected Networks," in *Game Theory for Security and Risk Management*, ser. Static & dynamic game theory : foundations & applications. Cham, Switzerland: Birkhäuser, 2018, pp. 285–311.

[47] M. Findrik, P. Smith, J. H. Kazmi, M. Faschang, and F. Kupzog, "Towards secure and resilient networked power distribution grids: Process and tool adoption," in *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Nov 2016, pp. 435–440.

[48] B. S. Onggo, J. Busby, and Y. Liu, "Using agent-based simulation to analyse the effect of broadcast and narrowcast on public perception: A case in social risk amplification," in *Proceedings of the Winter Simulation Conference 2014*, Dec 2014, pp. 322–333.

[49] S. König, A. Gouglidis, B. Green, and A. Solar, *Assessing the Impact of Malware Attacks in Utility Networks*. Cham: Springer International Publishing, 2018, pp. 335–351.

[50] Bundesamt für Sicherheit in der Informationstechnik, "BSI-Standard 100-2: IT-Grundschutz Methodology," https://www.bsi.bund.de/cln_156/ContentBSI/grundschutz/intl/intl.html, May 2008, version 2.0, english.