*14th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP14*
*Dublin, Ireland, July 9-13, 2023*

# Comparison of Cascading Effects in Critical Infrastructure Networks based on Simulation

Sandra König
*Scientist, Center for Digital Safety & Security, Austrian Institute of Technology, Vienna, Austria*

Stefan Schauer
*Senior Scientist, Center for Digital Safety & Security, Austrian Institute of Technology, Vienna, Austria*

ABSTRACT: Assessing impacts of threats to critical infrastructures is challenging due to the high complexity. Using probabilistic simulations seems adequate to capture the uncertainty but makes it at the same time more challenging to compare different situations. Such comparison is particularly important when actions for improved protection are considered. We here propose a simulation that gives a probability distribution over a (finite) set of states and to compare the probability distributions in such a way that actions with lower chance of worst-case damage are preferred. The approach is illustrated with data from the H2020 project PRECINCT where the impact of a flooding on a city is investigated.

## 1. INTRODUCTION

Comparison of two situations with uncertain consequences is the first step to understand which protective actions should or should not be taken in case of a dangerous event. Assessing the effects of natural hazards, such as floods, on critical infrastructures (CIs) is challenging since a precise prediction of the consequences is impossible due to the high complexity of CIs and their environment. In particular, the interdependencies among CIs yield many indirect consequences and cascading effects, of which we are often unaware. Numerous factors influence these consequences, such as the magnitude, the duration, or the time of the initial event (day or night), or whether other events are taking place simultaneously, such as an epidemic as we have experienced it during the last three years with COVID19. In the light of all this, a probabilistic model seems most suitable to describe the effects of an incident.

Once it is possible to estimate the impact of an incident for a given situation, it is interesting to see how this impact changes for different settings. The situation can change due to

- different threats, incl. different types of threats, but also different magnitudes of floodings
- different reactions of the system to the same threat, e.g., due to protective actions taken by the CI operator

In this paper, we mainly focus on the second point. The motivation behind this is to evaluate different potential actions to reduce the impact of a specific threat and evaluate which actions reduce the impact the most. Therefore, we describe a probabilistic approach for the simulation of cascading effects in Section 2.1 by highlighting how different scenarios can be evaluated and compared as well as discussing the advantage and main challenges of this approach. Further, we illustrate our approach with an example from the research project PRECINCT[1] in Section 2.2. Therein, we provide an overview of how a flooding in a metropolitan area can be described and how its cascading effects on the CIs in that area can be simulated. By analyzing the effects of different setups (based on the implementation of different countermeasures), we show how comparing these effects can be used to

[1] https://www.precinct.info/

support decision makers in finding optimal strategies to prepare against such a flooding.

## 2. PROBABILISTIC SIMULATION OF CASCADING EFFECTS

The intrinsic uncertainty about the effects of natural hazards on CIs calls for a probabilistic model. In this paragraph, we demonstrate how a simulation based on such a probabilistic model helps compare different situations (due to different threats or various protective measures).

### *2.1. Cascading Effects Simulation*

#### *2.1.1. Build-up*

Probabilistic models for the impact of incidents such as natural hazards are frequently applied (Schaberreiter et al. 2013; Sun et al. 2016; Zhang et al. 2020). Many of these models are inspired by spreading models of infectious diseases (Cohen et al. 2001; König et al. 2016), and therefore, in general, only consider a binary impact ("affected" or "not affected"). When dealing with CIs, such a binary representation is not sufficient. A model that describes the condition of a CI in more detail is (König et al. 2019). It uses multiple states (typically 3 or 5) to represent different levels of functionality or availability of the individual CIs (or their relevant components). This state may change due to an incident, e.g., the respective infrastructure may not function as before the incident happened. This change is described probabilistically because, in general, it is influenced by too many factors to be described deterministically. A CI or one of its components is a highly complex ecosystem, which cannot (and should not) be described in every detail; the model is designed to rather capture the behavior of the CI or the component on a higher level of abstraction.

Formally, this model is a directed graph, called the interdependency graph, describing CIs (or relevant components of CIs) through nodes and the dependencies between them through directed edges. In order to describe the state changes in a CI, each node contains a probabilistic Mealy automaton. This automaton receives an input (such as a notification of a dangerous event) and describes the change of state depending on the input. If the state changes, i.e., there is a problem in the node due to the input, a notification is sent to all neighbors. The current implementation provides the following output.

(1) Statistics including the relative frequencies of all states for each node over all simulation runs
(2) A visualization of the network for every time unit for each simulation run with nodes colored according to their state
(3) A timeline of events for each simulation run.

It is possible to include artificial nodes, too, to describe the impact of an incident, e.g., a "people" or "society" node that measures how strongly humans are affected by the incident, both directly and indirectly.

#### *2.1.2. Evaluation of different scenarios*

The existing simulation model supports risk management in at least two ways. On the one hand, it allows comparison of different scenarios, e.g., the consequences of different levels of flooding. For this, the initial event, i.e., the "trigger", is different for each scenario because it causes a specific damage to the first node than for other scenarios. In this case, the network model remains unchanged, including the local dynamic described by the Mealy automaton inside the nodes. The results of the different simulations show how the impact on the entire network changes for different threats.

On the other hand, it is possible to evaluate the effect of different proactive and protective measures that can be taken by CI operators. Such protective measures can change the reaction of one or more nodes, which are described in the model through the probabilistic Mealy automaton. Such protective actions focus on a few nodes since each operator can only influence a part of the network and since budget is limited. To evaluate the benefit of a preventive action, the simulation is run with the corresponding changes in the model and the simulation results are compared with the results before the changes.

As described in the previous subsection, the simulation provides as a result a set of statistics on how frequently each node is in each state, i.e., a probability distribution $(p_1, \dots, p_5)$ over all possible states, e.g., from $1, \dots, 5$, where 1 is the best and 5 the worst state. To compare two scenarios, we now compare the vectors of probabilities using a lexicographic order. In this case, the highest priority is put on the worst state 5 in the sense that the scenario with the lowest probability for ending up in the worst case is preferred. In case the probabilities for the worst case are equal, the probabilities of the second worst case are compared and so on. This corresponds to the idea that the probability of the worst case should be minimized.

With this lexicographic ordering of the probabilities of the states, it is possible to rank different scenarios or different actions and to identify choices that minimize the probability for the worst state of a node. The decision on which node is used for this analysis is up to the user since it depends on the value of each node for a CI. If the focus is on the entire network, it is possible to average the probabilities with weights that represent the importance of each node.

A comparison of two different scenarios in the way described above paves the way for optimization. In a situation where the goal is to protect a network against multiple threats, game theory identifies optimal actions. When the payoffs are vector-valued (as in the situation considered here), a generalized framework can be applied to identify optimal actions (Rass et al. 2020) .

#### *2.1.3. Advantages*

The main advantage of such an evaluation and comparison of different scenarios using simulation is the low cost in terms of resources. Once the model is set up, it can be used by the security or risk officers within the CI operators after a short training and does not require special knowledge or equipment on the very technical details of the CI. Compared to a digital twin, it is not necessary to touch the real system or understand it to it very technical detail, which is more expensive and might even be dangerous as defining or training a digital twin could interfere with the operation of the CI's systems.

#### *2.1.4. Contrast*

The approach described above differs from existing methods in two main points. First, it compares different situations not just heuristically but in a formal way. The focus on the worst case agrees with a conservative view that is often used in the context of CIs. An optimal protection strategy identified with a game-theoretic model is provably optimal since deviations do not yield to a better result.

Second, the approach considers intrinsic uncertainty but provides more information on the components than most other models. As already mentioned above, in many models, only two states are considered for a node (i.e., "healthy" or "infected"), but our approach provides a broader distinction and describes the functionality or availability in more detail. Using probability distributions, we decide based on the full knowledge available rather than aggregating, e.g., using an expected value.

#### *2.1.5. Disadvantages*

The main disadvantage of the proposed method is that – depending on the complexity of the network that needs to be described – some or a large amount of effort may be required to build up a useful model. On the one hand, the individual nodes, i.e., CIs and their components, need to be identified together with the interdependencies among the CIs, which might be explicitly visible but could also be implicit. In recent projects, it showed that discussions with experts from different domains are one of the best ways to obtain this information, but it can also be rather time consuming. However, such discussions are essential for a common understanding of the dependencies among CIs.

Further, the parametrization of the local dynamics (i.e., the transition probabilities of the Mealy automaton) also requires some time since experts need to be interviewed about the behavior of the individual systems (CIs or relevant components of them). However, this effort may be reduced by grouping similar nodes (König and Shaaban 2022) or by using already existing simulation models such as digital twins of the CIs' components. The parametrization of the local dynamics could then be learned from these existing models.

For both aspects, one of the main disadvantages is that the quality of the simulation model and its outputs heavily relies on the information available and provided by the experts. Usually, CI operators are reluctant to share much information about their systems since this information is highly sensitive and could be used by adversaries to attack the respective CIs.

## 2.2. Application

### 2.2.1. Case selection

To illustrate the idea, we give an overview of a use case from the H2020 EU project PRECINCT. A description of this use case in full detail is not possible due to the sensitivity of the data. Still, a high-level analysis already provides valuable insights into (potentially unknown) interdependencies. The threat scenario considered is a flooding hitting a city, and the focus lies on how impacted people are by the incident, both directly and indirectly.

### 2.2.2. Scenario description

The considered scenario of a flooding affecting a city is described through an interdependency graph. It has been developed based on discussions with end users of the PRECINCT project and is shown in Figure 1. The star in the upper left corner represents the threat (flood) and triggers the cascading effects.

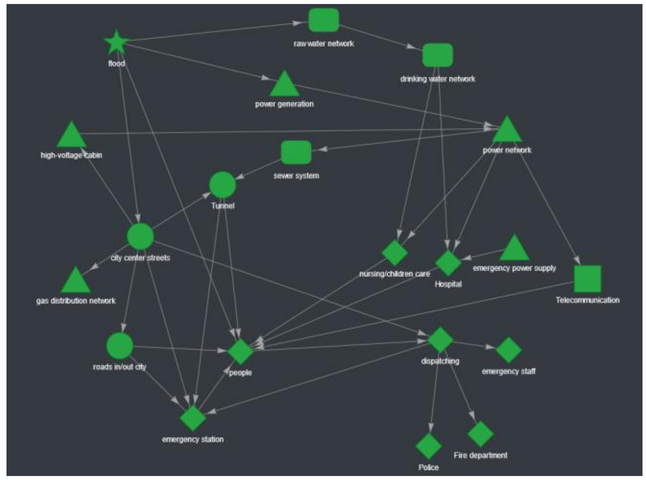


*Figure 1: Interdependency graph for flooding scenario.*

It represents the threat (star node) and the main components of the involved CIs:

- Transportation: tunnel, city center streets, peripheral roads (in/out)
- Water: raw water network, drinking water network, sewer system
- Energy: power generation, power network, high-voltage cabin, gas distribution network
- Health and emergency services: hospital (with emergency power supply), nursing/children care, dispatching of emergency services, emergency station, emergency staff, police, fire department
- Telecommunication

An artificial node representing people has been introduced to measure how strongly and how fast the people living in the city are affected.

The local dynamics of each node are described through the probabilistic Mealy automaton it contains. All nodes have the same number of states, but the transition probabilities are specific to each node (or group of nodes to reduce the parametrization effort) as they describe its characteristics. Since this reaction to threats is considered sensitive data of the CI, it is not possible to provide a full description of the transition probabilities here.

### *2.2.3. Simulation of a flooding*

Running 100 simulations yields statistics on the state of each of the nodes, as shown in Figure 2.


*Figure 2: Simulation results for each node in the graph.*

The simulation results for the people node are shown in Table 1. State 5 is by far the most likely case, so people are expected to be impacted very strongly. In only 4% of the simulation runs the impact on people has remained low.

*Table 1: Distribution over states of people node.*

| *State* | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| *Probability* | *0.0* | *0.04* | *0.11* | *0.22* | *0.63* |

For each simulation run, it is possible to have a closer look on when people become affected and why. Figure 3 shows an extract of the timeline of events.

| | Entity Name | Event | New State | Because of |
|---|---|---|---|---|
| 1 | flood | flood | 4 | |
| 2 | city center streets | flood | 3 | flood |
| 2 | raw water network | flood | 3 | flood |
| 2 | power generation | flood | 5 | flood |
| 2 | people | flood | 2 | flood |
| 3 | roads in/out city | city streets blocked | 3 | city center streets |
| 3 | gas distribution network | city streets blocked | 2 | city center streets |
| 3 | high-voltage cabin | city streets blocked | 2 | city center streets |
| 3 | dispatching | city streets blocked | 3 | city center streets |
| 3 | Tunnel | city streets blocked | 4 | city center streets |
| 3 | emergency station | city streets blocked | 4 | city center streets |
| 3 | drinking water network | threat water | 4 | raw water network |
| 3 | power network | threat power | 2 | power generation |
| 4 | people | peripheral roads blocked | 3 | roads in/out city |
| 4 | power network | hvc limited | 3 | high-voltage cabin |
| 4 | Fire department | dispatching affected | 2 | dispatching |
| 4 | emergency station | Tunnel blocked | 5 | Tunnel |
| 4 | people | emergency station limited | 4 | emergency station |
| 4 | Hospital | limited drinking water | 3 | drinking water network |

*Figure 3: Timeline of events for simulation of original scenario.*

At timestep 2, people are impacted directly (e.g., it is impossible to enter or leave houses). At timestep 4, people are affected indirectly since the peripheral roads are blocked and because the emergency station is only partially available.

The statistics in Figure 2 shows that the node 'emergency station' end up most frequently in the worst state (in 67 out of 100 cases). Since this node also impacts people (in the example from Figure 5 in timestep 4), it might be beneficial to consider options to increase the protection of this node. In the first setup, blocked roads in the city center influenced the emergency station strongly (if everything works fine before the flooding, the most likely new state is 3, but also 2 or 4 are possible). If this could be reduced such that the most likely state is 2, e.g., through an additional station at another location, it would be less likely to go to the worst state and therefore affect people less frequently or less strongly. To investigate if this change indeed improves the situation for people, we rerun the simulation with modified transition probabilities.

The results with adapted transitions probabilities of the node 'emergency station' are shown in Table 2.

*Table 2: Distribution over states of people node for modified simulation model.*

| *State* | *1* | *2* | *3* | *4* | *5* |
|---|---|---|---|---|---|
| *Probability* | *0.00* | *0.01* | *0.12* | *0.14* | *0.73* |

In the 100 simulation runs under the modified model, the people node was still very frequently in the worst state 5. Comparing the two simulation results in a lexicographic way in Eq. (1).

$$(0, 0.04, 0.11, 0.22, 0.63\,) \leq_{lex} (0, 0.01, 0.12, 0.14, 0.73) \quad (1)$$

shows that the new situation is not preferred to the original one. This shows that in complex situations, it is not always clear for the user how a situation can be improved and supports the simulation approach that we follow.

The timeline of a simulation of the modified scenario in Figure 4 shows that people are still strongly affected by limited hospital capacity or limited nursing capacity, so protecting the 'emergency station' node is not sufficient to improve the situation.

| | Entity Name | Event | New State | Because of |
|---|---|---|---|---|
| 1 | flood | flood | 4 | |
| 2 | city center streets | flood | 2 | flood |
| 2 | raw water network | flood | 3 | flood |
| 2 | power generation | flood | 3 | flood |
| 3 | roads in/out city | city streets blocked | 2 | city center streets |
| 3 | gas distribution network | city streets blocked | 3 | city center streets |
| 3 | high-voltage cabin | city streets blocked | 2 | city center streets |
| 3 | dispatching | city streets blocked | 3 | city center streets |
| 3 | Tunnel | city streets blocked | 3 | city center streets |
| 3 | emergency station | city streets blocked | 3 | city center streets |
| 3 | drinking water network | threat water | 4 | raw water network |
| 3 | power network | threat power | 2 | power generation |
| 4 | Police | dispatching affected | 2 | dispatching |
| 4 | Hospital | limited drinking water | 2 | drinking water network |
| 4 | nursing/children care | limited power | 2 | power network |
| 4 | Hospital | limited power | 4 | power network |
| 4 | Telecommunication | limited power | 2 | power network |
| 5 | people | limited hospital capacity | 2 | Hospital |
| 5 | people | limited nursing capacity | 3 | nursing/children care |
| 5 | people | limited hospital capacity | 4 | Hospital |

*Figure 4: Timeline of events for simulation of modified scenario.*

### 2.2.4. *Visualization*

An overview on the state of the entire network after the simulation of a flooding event is shown in Figure 5.

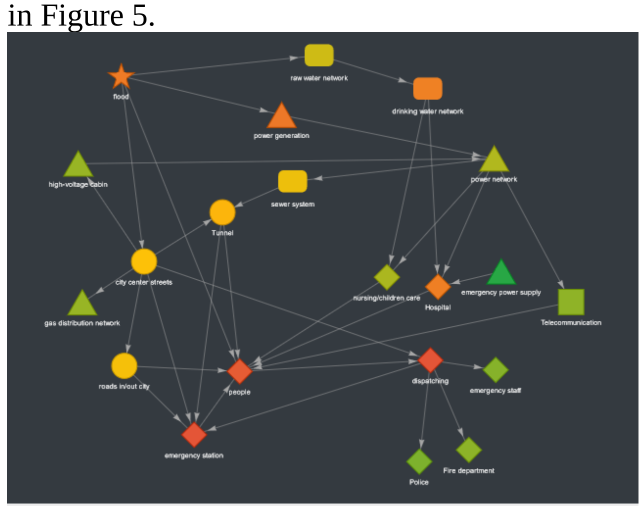


*Figure 5: Simulation results for flooding scenario (current setup).*

### 2.2.5. *Discussion*

The main insights from the considered example is that an action cannot always easily be judged as 'good' or at least better than another action if it has both direct and indirect consequences. In the considered situation of a network of CIs, uncertainty about the impact and especially cascading effects make it very hard to foresee the impact of an action. Therefore, the choice of protection actions should be guided by mathematical analysis.

### *2.2.6. Comparison with earlier work*

This paper extends earlier work in at least two ways. First, it provides a more detailed simulation model to assess the impact of an incident than the one used in (König et al. 2018). Second, it illustrates the multi-categorial risk assessment framework proposed in (König et al. 2022) with data from a research project with several CI operators rather than using dummy data.

## 3. CONCLUSIONS

Cascading effects of incidents in interdependent CIs make it challenging to assess the impact for each CI. Due to the high complexity, the impact is not fully described by a single number. Instead, we propose to use probability distributions over all possible outcomes. Comparison of different situations, e.g., due to protective measures, is required to judge which action is better or to measure how much the current situation improves. Once it is possible to compare the consequences of two actions, it is possible compare any finite number of actions and thus identify the best action(s) among multiple options. If protective actions are supposed to protect a CI against more than one threat, game theoretic optimization is recommended (with an adapted notation of equilibrium such as (Rass et al. 2015)). This will be our major focus of future work, based on the information collected in the PRECINCT project.

## 4. ACKNOWLEDGEMENT

The PRECINCT project has received funding from the European Union's HORIZON 2020 research and innovation program under Grant Agreement No 101021668.

## 5. REFERENCES

Cohen, R., K. Erez, D. ben-Avraham, and S. Havlin. 2001. "Breakdown of the Internet under Intentional Attack." *Physical Review Letters*, 86 (16): 3682–3685.

König, S., A. Gouglidis, B. Green, and A. Solar. 2018. "Assessing the Impact of Malware Attacks in Utility Networks." *Game Theory for Security and Risk Management*, S. Rass and S. Schauer, eds., 335–351. Cham: Springer International Publishing.

König, S., S. Rass, B. Rainer, and S. Schauer. 2019. "Hybrid Dependencies Between Cyber and Physical Systems." *Intelligent Computing*, 550–565. Cham: Springer.

König, S., S. Schauer, and S. Rass. 2016. "A Stochastic Framework for Prediction of Malware Spreading in Heterogeneous Networks." *Secure IT Systems: 21st Nordic Conference, NordSec 2016, Oulu, Finland, November 2-4, 2016. Proceedings*, B. B. Brumley and J. Röning, eds., 67–81. Cham: Springer International Publishing.

König, S., S. Schauer, M. Soroudi, I. Ko, P. Carroll, and D. McCrum. 2022. "Risk Management with Multi-categorical Risk Assessment." Grenoble, France.

König, S., and A. M. Shaaban. 2022. "Parametrization of Probabilistic Risk Models." *Proceedings of the 17th International Conference on Availability, Reliability and Security*, 1–6. Vienna Austria: ACM.

Rass, S., S. König, and S. Schauer. 2015. "Uncertainty in Games: Using Probability-Distributions as Payoffs." *Decision and Game Theory for Security: 6th International Conference, GameSec 2015, London, UK, November 4-5, 2015, Proceedings*, M. Khouzani, E. Panaousis, and G. Theodorakopoulos, eds., 346–357. Cham: Springer International Publishing.

Rass, S., König Sandra, and A. Alshawish. 2020. *HyRiM: Multicriteria Risk Management using Zero-Sum Games with vector-valued payoffs that are probability distributions*.

Schaberreiter, T., P. Bouvry, J. Röning, and D. Khadraoui. 2013. "A Bayesian Network Based Critical Infrastructure Risk Model." *EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation II*, O. Schütze, C. A. Coello

Coello, A.-A. Tantar, E. Tantar, P. Bouvry, P. Del Moral, and P. Legrand, eds., 207–218. Berlin, Heidelberg: Springer Berlin Heidelberg.

Sun, X., J. Dai, P. Liu, A. Singhal, and J. Yen. 2016. "Towards probabilistic identification of zero-day attack paths." *2016 IEEE Conference on Communications and Network Security (CNS)*, 64–72. Piscataway, NJ: IEEE.

Zhang, L., Y. Wang, Y. Chen, Y. Bai, and Q. Zhang. 2020. "Drought Risk Assessment in Central Asia Using a Probabilistic Copula Function Approach." *Water*, 12 (2): 421. https://doi.org/10.3390/w12020421.