Representing Domains and Scenarios by Means of Model Replication and Composition

Daniele Codetta-Raiteri[1], **Roberto Nai**[1]
Dipartimento di Informatica, Università del Piemonte Orientale
Alessandria, Italy

Abstract

We consider a domain as a particular system or a portion of the system, while a scenario is a sequence of effects on the domain, originated by a particular event or condition. We show how it is possible to build first the model of the domain by replication and composition of atomic models, each representing a particular aspect of the domain. Then, the models of the scenarios are obtained from the domain's model, by composing further atomic models representing the events originating the scenarios. In particular, we take into account the domain consisting of one control centre and a set of substations inside an electrical distribution grid, communicating by means of a network. We consider scenarios originated by threats such as the denial of service attack to the communication network, and the temporary unavailability of substations due to the failure and the repair of the internal components. Stochastic Activity Networks (SAN) are the modelling formalism. The simulation of the models representing the scenarios, estimates the impact of the threats on the communication reliability.

1. Introduction

The CRUTIAL project (*CRitical UTility InfrastructurAL resilience*) [1] has investigated the ways to obtain the resilience of the *Electrical Power System* (EPS); this means the capacity of the EPS to provide its service despite of the occurrence of failures or attacks. For instance, an attack to a communication network may affect the data exchange among the EPS sites connected by that network, compromising an automation function depending on such data, as in the case of the voltage regulation [2]. One of the activities in CRUTIAL is the evaluation of the critical scenarios [2] consisting of particular event sequences occurring in a specific portion of the EPS (scenario domain [3]), as a consequence of an attack or a failure. One of the ways to estimate the effects of such threats on the domain is the simulation of the stochastic models of the scenarios.

In [4], we considered the domain consisting of a control centre and a set of substations, connected by a communication network, inside a distribution grid of the EPS, and we evaluated in reliability terms, scenarios characterized by the communication network unavailability, and by intrusions generating fake commands directed to the substations.

In this paper, we focus more on the modelling approach and we apply it in the same domain (Sec. 2), to the scenarios where the communication is affected by *denial of service* (DoS) attacks and failures of the substation components (Sec. 5). According to the modelling approach, first we represent the domain

(Sec. 4) by replicating and composing *Stochastic Activity Network* (SAN) [5] (Sec. 3) models, each representing a particular aspect of the domain. Then, the scenarios are represented (Sec. 6) by composing the domain model with further SAN models, each concerning a certain threat. The resulting models are simulated (Sec. 7) in order to estimate the communication reliability in terms of probability and quantity of failed communication sessions. The design, the replication, the composition and the simulation of the SAN models are supported by the tool *Möbius* [6].

2. Domain specification

The scenarios under exam take place in a domain composed by one control centre, a set of 10 substations, and 2 redundant communication networks (Fig. 1), located in a distribution grid of the EPS. Substations execute the commands coming from the control centre and concerning the electrical lines connected to the substations. An example of command is the arming or disarming order [2]. The generation of a command may occur as a consequence of the state of the electrical lines, described by the signals transmitted from the substations to the control centre. Such information allows the control centre to monitor and rule the distribution grid under its control. So, the communication of commands and signals has to be reliable in order to avoid malfunctioning.

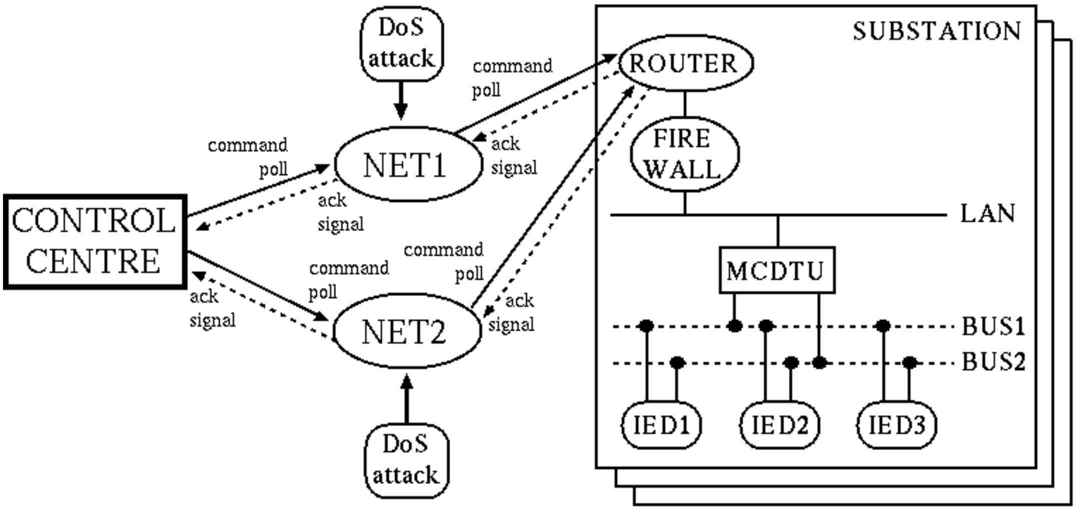

Figure 1. The scheme of the domain (and the threats (Sec. 5))

2.1 Communication sessions

In our domain, we suppose that each command generated by the control centre has to be executed by all the substations; therefore, a copy of the command is sent to each substation. Moreover, we assume that the execution of a command by a substation is notified to the control centre by the transmission of an acknowledgment coming from the substation. The following sequence of operations is called *"command session"*:

1. the control centre opens the command session: it generates the command and starts collecting the acknowledgments coming from the substations and concerning the command execution, until a certain time out expires;
2. a copy of the command is transmitted on the available communication network to each substation;
3. each substation executes the command and generates an acknowledgment proving the execution of the command;
4. each acknowledgment is transmitted on the available communication network to the control centre;
5. the time out for the acknowledgments collection expires and the command session is closed.

We suppose that signals are not sent by a substation in an autonomous way, but we assume that they are generated as a reply to a poll request: periodically the control centre polls all the substations by sending a poll request to each of them, and they reply by sending a signal to the control centre. A *"signal session"* consists of the following sequence of operations:
1. the control centre opens the signals session: it generates a poll and starts collecting signals coming from the substations, until a certain time out expires;
2. a poll request is transmitted on the available communication network to each substation;
3. each substation generates the signal;
4. each signal is transmitted on the available communication network to the control centre;
5. the time out for the signals collection expires and the signal session is closed.

We assume that at most one command (signal) session is running at any time. In the domain under study (and in the scenarios (Sec. 5)), the time for an event to occur can be deterministic or random; in the second case, such time is ruled by the negative exponential distribution. The occurrence (mean) times for the events in a command or signal session are reported in Tab. 1.

Event	Type of event	(mean) time to occur	occurring rate
command generation	stochastic	6.00000E+0 h	0.16667 1/h
command execution	stochastic	2.77778E-4 h	3600 1/h
time out for ack.	deterministic	5.55555E-3 h	-
poll generation	deterministic	8.33333E-2 h	-
signal generation	stochastic	2.77778E-4 h	3600 1/h
time out for signals	deterministic	5.55555E-3 h	-
packet transmission	stochastic	2.77778E-4 h	3600 1/h

Table 1. The (mean) occurrence time (and rates) for the events in a session

2.2 Transmission of packets

In our domain, the transmission of the several kinds of packets (command copies, acknowledgments, poll requests and signals) is performed by means of the redundant communication networks NET1 and NET2. NET1 is usually used for the communication between the control centre and the substations. We suppose that the bandwidth of each communication network is equal to 16 Kbit/sec and that the transmission of each packet consumes 1 Kbit/sec of the bandwidth. This means that no more than 16 packets can be transmitted on the same communication network at the same time. It may happen that the current available bandwidth of NET1 is not enough to transmit all the packets. For instance, if a command session and a signal session are running in parallel way, it may happen that 10 acknowledgments and 10 signals have to be transmitted to the control centre at the same time. In this case, 16 of such packets will be transmitted by NET1, while the remaining 4 packets will be directed to NET2 for the transmission.

Actually, we could have specified that the transmission of a packet requires less than 1 Kbit/sec of the bandwidth, or that a communication network has a bandwidth higher than 16 Kbit/sec; in this way, the communication network would be able to transmit more than 16 packets at the same time. Our choice depends on the fact that one of the goals of the scenarios is evaluating the effect of the bandwidth consumption, on the communication reliability. To this aim, if the communication networks had a higher transmission capacity, then we would need to consider more than 10 substations in the case study, possibly making the simulation computing cost worse.

3. SAN formalism

SAN can be considered as a particular form of Stochastic Petri Net (SPN) [7]; so, a SAN model contains places (appearing as circles), activities (appearing as bars) and arcs. A place contains a certain number of tokens (marking). An instantaneous activity (transition) completes (fires) as soon as it is enabled; a timed activity instead, completes after a certain amount of time which can be random or deterministic. A particular condition on the marking of a certain set of places enables the completion (firing) of activities whose effect is modifying in some way the marking of the places. Such condition and effect can be expressed by connecting the activity to the places by means of oriented arcs, as it is possible in SPN. Another way consists of using input gates. An input gate is connected to an activity and to a set of places, and is characterized by two expressions:

- a *predicate* consists of a Boolean condition expressed in terms of the marking of the places connected to the gate; if such condition holds, then the activity connected to the gate is enabled to complete.
- a *function* expresses the effect of the activity completion on the marking of the places connected to the gate.

A SAN model can contain output gates as well. An output gate has to be connected to a certain activity and to a set of places. Its role is specifying only the effect of the activity completion, so it is characterized only by a function. Gates graphically appear as triangles (input gate: ◄ - output gate: ►).

The *Replicate/Join* formalism [5] was conceived for SAN models and expresses by means of a tree structure, the way to compose together several SAN models in a unique large composed model: leaf nodes are atomic SAN models, each non leaf node is a *Join* or *Replicate* operator, and the root node is the resulting model. In particular, the *Join* operator composes two or more SAN models sharing places. The *Replicate* operator constructs a model consisting of a number of identical copies of a certain SAN model (copies can share places).

3.1 Motivating the use of SAN

As Petri Net based formalisms in general, SAN expresses the system states and behaviour in terms of places containing tokens, and transitions modifying their quantity. So, the system dynamics is represented by the token game, avoiding the modeller to enumerate the complete state space of the system. This is useful in particular when the system behaviour is characterized by the occurrence of concurrent events. Deterministic and random completion times are available in SAN. This is a reason why SAN is suitable to model the domain and the scenarios in this paper, where both stochastic and deterministic events occur.

Another advantage of SAN is the presence of gates which allow to set complex firing conditions or effects that would be very complicated (or impossible) to express in a SPN, only by means of oriented arcs. This simplifies the graph structure of the model when we represent complex systems. Moreover, the modeller can concentrate its attention on each particular aspect of the system behaviour and represent it in form of SAN; then, the SAN models can be easily composed in order to obtain the model of the whole system.

4. Building the model of the domain

In our modelling approach, we first model in form of SAN, each aspect of the domain, in isolation. Then, the SAN models are replicated and joined in order to obtain the model of the whole domain. Actually several places are shared by the SAN models and they act as points of connection when the models are composed. For the sake of brevity, in this section we briefly describe the SAN models of the domain aspects, while all their details can be found in [8].

The functions of the control centre (Sec. 2.1) are represented by the SAN model appearing in Fig. 2 where the upper part concerns the command generation and the collection of acknowledgments, while the lower part is about the poll generation and the collection of signals. The functions performed by a substation are modelled in the SAN model in Fig. 3: the upper part of the model is about the execution of commands and the generation of acknowledgments, while the lower part concerns the generation of signals.

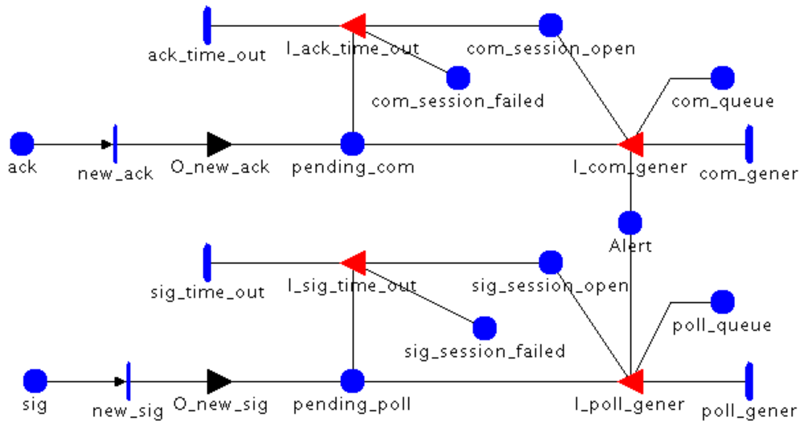

Figure 2. The atomic model of the control centre functions

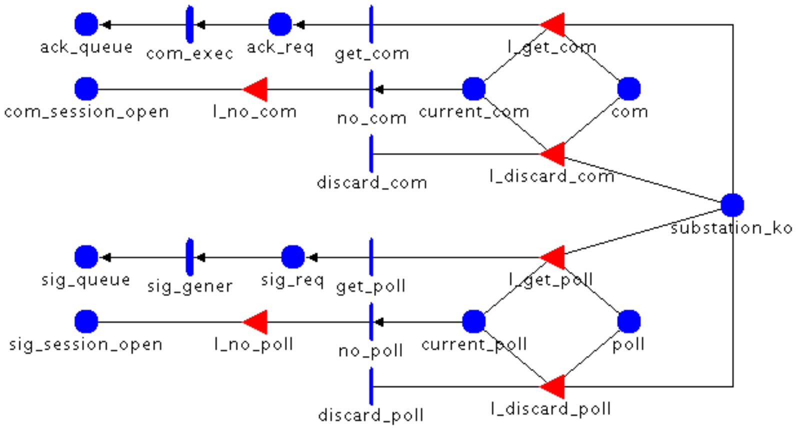

Figure 3. The atomic model of the substation functions

The transmission of packets can be performed by the communication network NET1 or by NET2 (Sec. 2.2); packets can be command copies, acknowledgments, poll requests or signals. The SAN model in Fig. 4 represents this situation. The markings of several places in this model represent packets waiting to be transmitted on the available communication network: the tokens inside the places *com_queue* and *poll_queue* represent command copies and poll requests respectively, and they appear also in the SAN model of the control centre (Fig. 2); the tokens inside the places *ack_queue* and *sig_queue* represent acknowledgments and signals respectively, and they appear also in the SAN model of the substation functions in Fig. 3. Other places in the SAN model in Fig. 4 represent instead packets that have been delivered: the markings of the places *ack* and *sig* represent the acknowledgments and the signals respectively, delivered to the control centre; such places appear in the SAN model of the control centre (Fig. 2) as well. The tokens inside the places *com* and *poll* represent the command copies and the poll requests respectively, delivered to the substations; therefore these places belong also to the SAN model of the substation functions (Fig. 3).

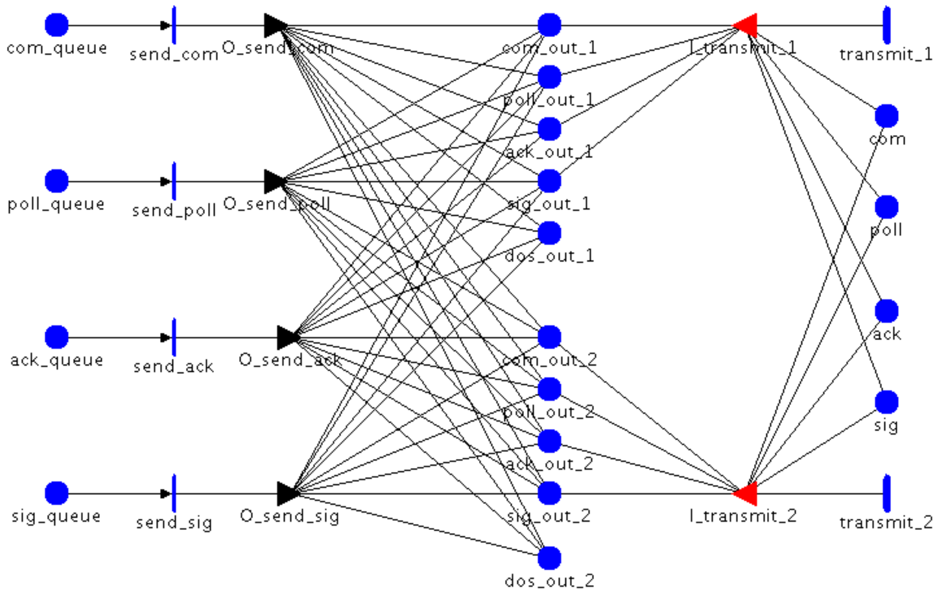

Figure 4. The atomic model of the packets transmission

Besides representing the packets transmission, the SAN model in Fig. 4 acts as a "bridge" to join the previous SAN models in order to build the model of the whole domain. This is done in Fig. 5 where the SAN model of the substation is replicated 10 times by means of the *Rep* operator (Sec. 3), in order to represent the presence of 10 substations in the domain (Sec. 2.1). The result of the replication is joined with the SAN model of the control centre (Fig. 2) and with the SAN model of the packets transmission (Fig. 4) according to the common places mentioned above. This is done by means of the *Join* operator (Sec. 3) and generates the model of the domain.

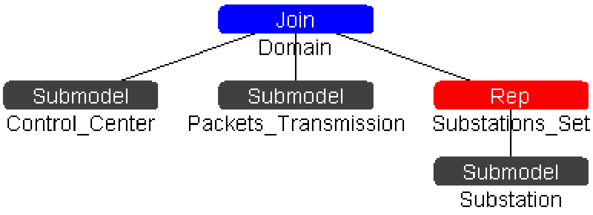

Figure 5. The composed model of the domain

5. Scenarios specification

In absence of attacks or failures, the communication between the control centre and the substations cannot fail. In case of threats instead, some packets (command copies, acknowledgments, poll requests, signals) may be lost. If the number of substations is N, we consider a command (signal) session as successful if at least N −1 acknowledgments (signals) are received by the control centre before that the time out expires (N = 10 in the domain under study). In other words, if more than one acknowledgment (signal) is missing when the time out expires, then the command (signal) session is considered to be failed.

As mentioned in Sec. 1, each scenario is characterized by the occurrence of a particular kind of attack or failure, and in this paper we are interested in evaluating three scenarios:

- Scenario 1: the DoS attacks may occur;
- Scenario 2: the substations failures may occur;
- Scenario 3: both the substations failures and the DoS attacks may occur.

5.1 DoS attack

During a DoS attack, the attacker sends a huge amount of packets on the affected communication network: the effect is the gradual reduction of the bandwidth available for normal communication, leading to the complete unavailability of the bandwidth. We assume that a DoS attack may affect NET1 or NET2 (Fig. 1). Both communication networks may be attacked several times, but a communication network cannot be the object of more than one attack at the same time. It may happen that both networks are under attack at the same time, but in this case, two distinct attacks are running and each affects one communication network.

Event	mean time to occur	occurring rate
DoS occurrence	720 h	0.00139 1/h
DoS duration	12 h	0.08333 1/h
Bandwidth reduction by 1 Kbit/sec	0.1875 h	5.33333 1/h

Table 2. The occurrence mean times and rates of the events in the DoS attack

NET1 and NET2 are redundant; so, in case of NET1 under attack, its bandwidth is gradually consumed by the packets transmitted by the attacker; therefore also NET2 has to be exploited to transmit. If the global available bandwidth of both NET1 and NET2 is not enough to transmit all the packets, some of them will not be transmitted becoming lost.

We suppose that the mean time to completely consume the bandwidth of NET1 is 3 h: since the bandwidth of NET1 and NET2 is 16 Kbit/sec respectively, then the bandwidth occupancy by the DoS attack is increased by 1 Kbit/sec every 675 sec. (Tab. 2). When the DoS attack ends, the bandwidth consumed by the attack becomes available again for the normal communication.

5.2 Substation failure

We assume that a substation is composed by three subsystems (Fig. 1):

- the MCDTU (*Monitoring Control and Defence Terminal Unit*) [2] is the core of the substation and consists of a particular device in charge of managing the requests for command execution or for signal generation. The MCDTU is connected to both the substation LAN and to the substation bay.
- The LAN (*Local Area Network*) acts as a bridge between the MCDTU and the external communication networks NET1 and NET2: all the packets

transferred from NET1 or NET2 to the LAN, then to the MCDTU (commands and polls), or in the opposite sense (acknowledgments and signals), are directed by a router and filtered by a firewall.

• The bay contains all the electrical devices necessary to physically perform the commands received by the MCDTU, and to generate the signals. We assume that the bay contains three redundant IED (Intelligent Electronic Device) [2] components connected to the MCDTU by means of two redundant electrical buses: the MCDTU controls the IEDs ordering them the execution of the commands or the retrieval of signals.

Component	MTTF	Failure Rate	MTTR	Repair Rate
bus	4380 h	2.28311E-4 1/h	24 h	4.16667E-2 1/h
IED	4380 h	2.28311E-4 1/h	48 h	2.08333E-2 1/h
MCDTU	8760 h	1.14155E-4 1/h	12 h	8.33333E-2 1/h
router	17520 h	5.70776E-5 1/h	6 h	1.66667E-2 1/h
firewall	17520 h	5.70776E-5 1/h	6 h	1.66667E-2 1/h

Table 3. The mean time to failure, the failure rate, the mean time to repair, and the repair rate for each of the substation components

The failure mode of the substation is expressed by the *Fault Tree* (FT) [7] in Fig. 6. A substation cannot execute commands or generate signals while it is unavailable because of its internal failure. We assume that all the substation components are repairable, so the substation can be available again. We suppose that each repair process acts on a single component, and such processes can be executed in parallel way. The mean time to failure, the mean time to repair and the corresponding rates are reported in Tab. 3. Actually we do not resort to the *Fault Tree Analysis* (FTA) [7] in the scenarios evaluation. The FT model is exploited only as a graphical representation of the failure mode of the substation, and in Sec. 6, it will be converted in SAN form.

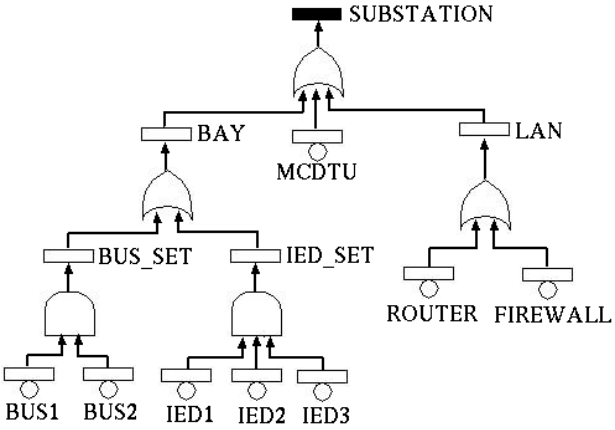

Figure 6. The Fault Tree model of the substation failure mode

6. Building the models of the scenarios

The model of a scenario is obtained by representing the threat characterizing the scenario in form of SAN, and joining it with the model of the domain, still exploiting the common places.

The Scenario 1 is characterized by the occurrence of DoS attacks (Sec. 5.1). A single DoS attack is modelled by the SAN in Fig. 7; it contains the place *dos_out* modelling the occupancy of the bandwidth by the packets transmitted by the DoS attack. Since this may affect NET1 or NET2, two instances of the DoS attack model are composed with the model of the domain in order to obtain the model of the Scenario 1 (Fig. 8): one instance represents the DoS attack to NET1, so its place *dos_out* corresponds to the place *dos_out_1* in the SAN model of the packets transmission (Fig. 4). The other instance concerns the attack to NET2; therefore its place *dos_out* corresponds to the place *dos_out_2* of the packets transmission model. In this way, the model in Fig. 4 takes into account the bandwidth consumption also by means of the DoS packets, and acts as a bridge also to include the DoS attack in the scenario model.

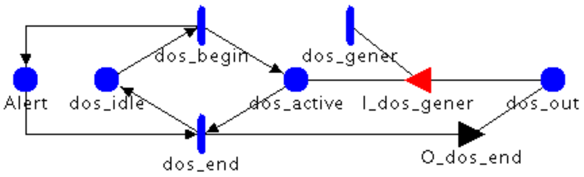

Figure 7. The atomic SAN model of the DoS attack

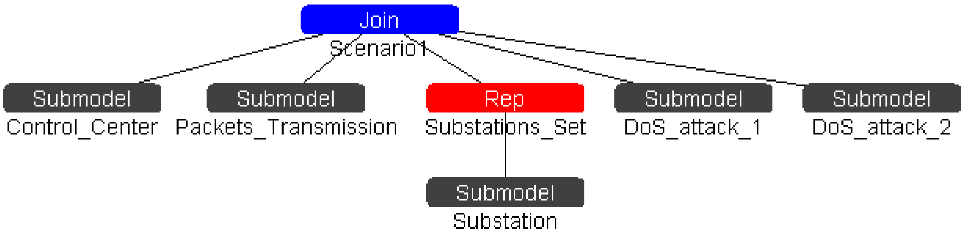

Figure 8. The composed model of the Scenario 1

In the Scenario 2, the communication may be compromised by the unavailability of the substations (Sec. 5.2). The SAN model in Fig. 9 consists of the conversion into SAN, of the FT model in Fig. 6, with the addition of the repair processes, each involving a single component. In particular, this SAN model contains the place *substation_ko* indicating if the substation is currently unavailable or not. The composed model of the Scenario 2 in Fig. 10 is derived from the domain model (Fig. 5) in this way: before the replication, the SAN model of the substation functions (Fig. 3) is joined with the SAN model of the substation failure and repair (Fig. 9), by means of the common place *substation_ko*. In this way, in the resulting model of the substation, its functions are disabled if such place is marked (the substation is unavailable). Then, such model is replicated in order to represent the set of 10 substations in the domain.

Finally, the Scenario 3 takes into account both the DoS attacks and the substations failures. So, its composed model (Fig. 11) is obtained from the model of the domain by including two instances of the DoS attack SAN model, and the SAN of the substation failure and repair. The details about the atomic SAN model in Fig. 7 and in Fig. 9 are available in [8].

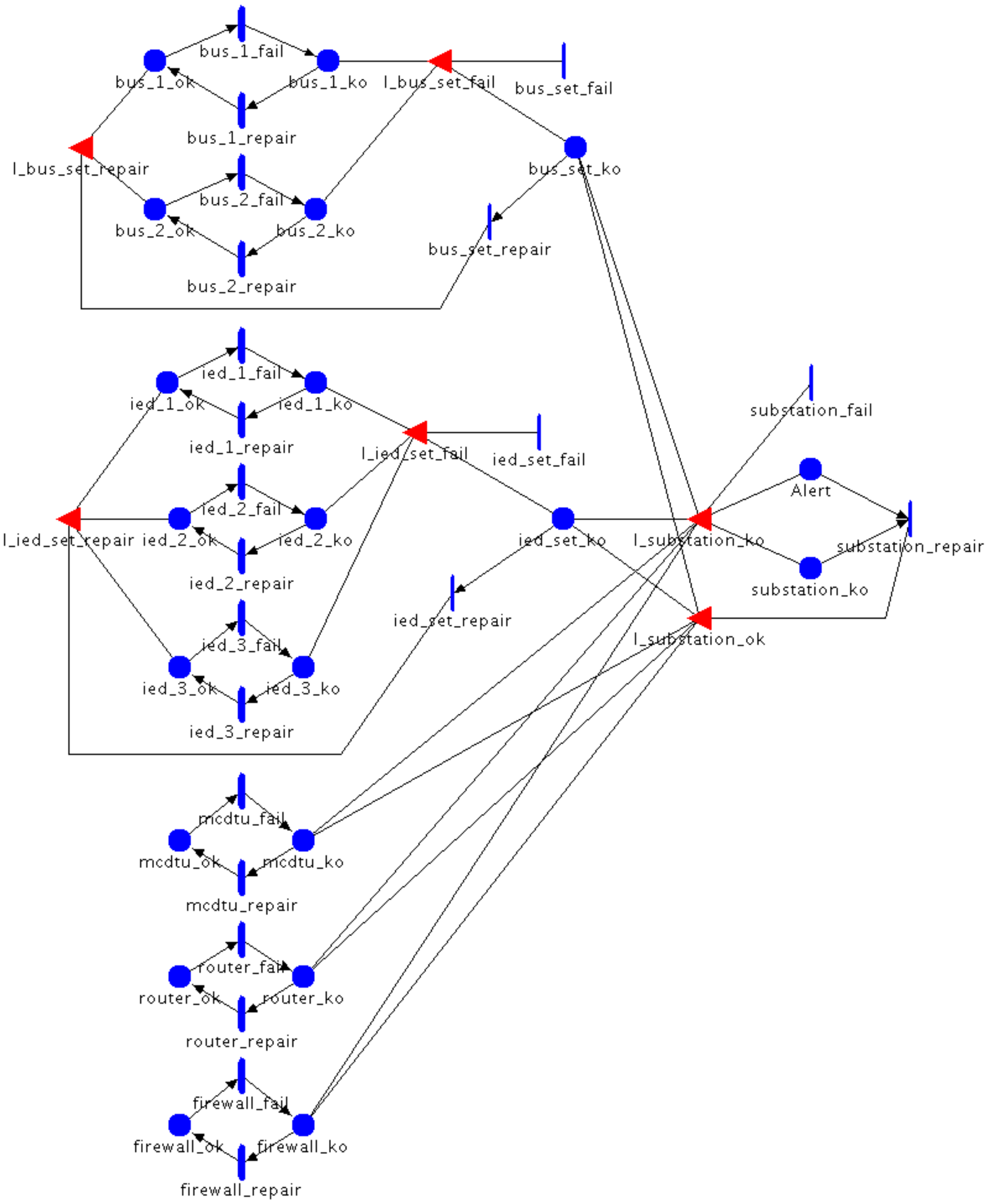

Figure 9. The atomic model of the failure and the repair of the substation

7. Simulating the scenarios

For each scenario model described in the previous section, 10000 simulation batches have been performed by means of the tool *Möbius*, setting a confidence level of 0.95, and a relative confidence interval of 0.1. The measures computed by the simulation are:

- *Pr_com(t)*: the probability that at least one command session has failed at a certain time;

- *Pr_sig(t)*: the probability that at least one signal session has failed at a certain time;
- *Num_com(t)*: mean number of failed command sessions at a certain time;
- *Num_sig(t)*: mean number of failed signal sessions at a certain time.

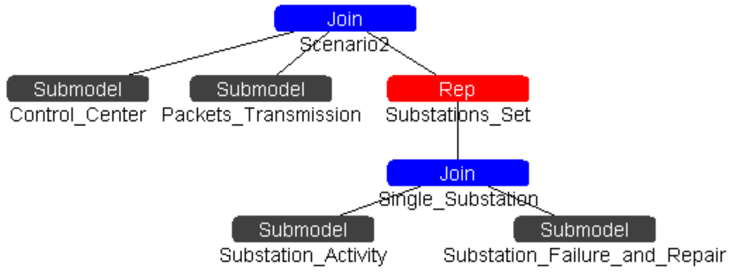

Figure 10. The composed model of the Scenario 2

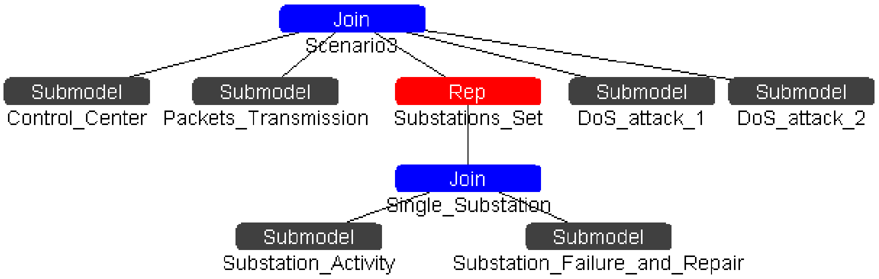

Figure 11. The composed model of the Scenario 3

The functions expressing such measures in terms of place markings are reported in [8]. All measures are computed for a mission time *t* varying between 0 and 10000 h. The values of *Pr_com(t)* and *Pr_Sig(t)* returned by the simulation in each scenario are depicted in Fig. 12 where we notice that the DoS attacks (Scenario 1) determine an higher probability of command or signal session failure, with respect to the substation failures (Scenario 2). This is confirmed in terms of number of failed sessions, by the results obtained for the measures *Num_com(t)* and *Num_sig(t)*, as shown in Fig. 13. The simulation results in numerical form are available in [8] where they are commented with more detail.

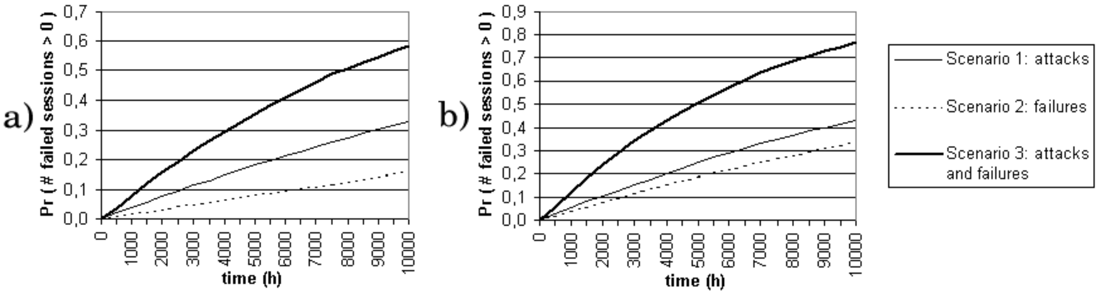

Figure 12. **a)** *Pr_com(t)* **b)** *Pr_sig(t)*

In the CRUTIAL project, the results of the models simulation have been compared with the results of test-beds execution on prototypal power system management architectures. The evaluation of scenarios both in form of models and in form of test-beds, has supported and inspired the definition of architectures resilient to attacks and failures. The achievements of the project are reported in [9].

8. Conclusions

In this paper, we focused on the modelling approach to represent domains and their critical scenarios. In particular, we dealt with the domain consisting of the control centre and the substations in a distribution grid of the EPS. According to the modelling approach, the composed SAN model of the domain has been built by applying the *Join* and *Replicator* operators to the atomic SAN models dedicated to the control centre, the substation, and the communication network, respectively. Then, the models of the scenarios have been obtained by extending the domain composed model, with the atomic models representing the DoS attack and the substation failure respectively. The reliability of the communication in each scenario has been evaluated in terms of probability and quantity of command or signal session failures, by simulating the corresponding SAN model, by means of *Möbius*.

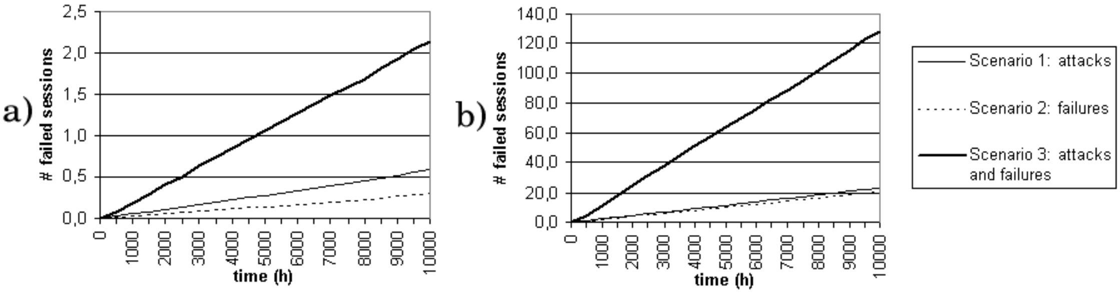

Figure 13. **a)** *Num_com(t)* **b)** *Num_sig(t)*

Modelling only the failure mode of the system, as we do in FTA, is not enough to evaluate the communication reliability in the scenarios: actually we had to concentrate our attention first on the model of the normal functioning of the system (domain model), and then to the occurrence of threats with their effect on the system functioning (scenario model). The possibility of replicating and composing atomic models simplified this task. The domain model can be useful for performance evaluation, while the scenario model is effectively oriented to estimate the reliability. If we had to model other scenarios in the same domain, we would have to create only the models of the new threats, and combine them with the existing model of the domain, as we did in [4] (Sec. 1).

References

1. CRUTIAL project's web page. http://crutial.rse-web.it.
2. Garrone F. (editor), Brasca C., Cerotti D., Codetta-Raiteri D., Daidone A., Deconinck G., Donatelli S., Dondossola G., Grandoni F., Kaâniche M. and Rigole T., Deliverable D2: Analysis of new control applications, CRUTIAL project, http://crutial.rse-web.it (2007).

3. Cerotti D., Codetta-Raiteri D., Donatelli S., Brasca C., Dondossola G. and Garrone F., UML diagrams supporting domain specification inside the CRUTIAL project, *Lecture Notes in Computer Science*, **5141** 106-123 (2008).
4. Codetta-Raiteri D. and Nai R., Evaluation of communication scenarios inside the Electrical Power System, *International Journal of Modelling and Simulation*, **30** [3] 345-352, ACTA Press (2010).
5. Sanders W.H. and Meyer J.F., Stochastic activity networks: Formal definitions and concepts, *Lecture Notes in Computer Science* **2090** 315–343 (2001).
6. Deavours D., Clark G., Courtney T., Daly D., Derisavi S., Doyle J., Sanders W.H. and Webster P.G., The Möbius Framework and its Implementation, *IEEE Transactions on Software Engineering*, **28** [10] 956–969 (2002).
7. Sahner R.A., Trivedi K.S. and Puliafito A., "Performance and Reliability Analysis of Computer Systems; An Example-based Approach Using the SHARPE Software Package.", Kluwer Academic Publishers (1996).
8. Codetta-Raiteri D. and Nai R., “Simulating the exchange of command and signals in a distribution grid”, Technical Report TR-INF-2009-12-08-UNIPMN, Dipartimento di Informatica, Università del Piemonte Orientale, *http://www.di.unipmn.it* (2009).
9. CRUTIAL project’s deliverables, http://crutial.rse-web.it/Dissemination/DELIVERABLES-OF-THE-PROJECT.asp