Dynamic graphical models for security and safety joint modeling

July 12th 2015 GraMSec Workshop, Verona



Marc Bouissou^{1,2} Siwar Kriaa^{1,2} Ludovic Piètre-Cambacédès¹ ¹EDF R&D, ²École Centrale Paris



Context: pervasive computing





Outline

Introduction

- Safety/security convergence
- Why Petri nets, SAN and BDMP
- Petri nets and SAN
 - Formalism description
 - Use case: security of a metro station

BDMP

- Formalism description
- Use case: a pipeline

Conclusion

Introduction



Industrial systems are more and more complex and interconnected





- Industrial systems targeted by cyber-attacks
- Large consequences on the system's environment
- Their requirements converge for complex systems



Terminology

Safety and security (SEMA referential) [1]



[1] L. Pietre-Cambacedes and C. Chaudet, "The SEMA referential framework: Avoiding ambiguities in the terms "security" and "safety"," International Journal of Critical Infrastructure Protection, Vol. 3 Issue 2, pp. 55-66, July 2010.

Safety and security



Differences

- Random vs intelligent
- Stability vs evolution
- Access to information
- Vocabulary

Similarities

- Protection aim
- Risk = fundamental notion
- Not "additive"
- Importance of human factors

Synergy between the two communities: possible & desirable



Interdependences Safety Security

Interdependences

- Antagonism
- Conditional dependence
- Mutual reinforcement
- Independence

Stakes

- Correct risk evaluation
- Cost optimization





Dynamic graphical models to study such interdependencies

- We need a holistic approach
- Single model describing both safety and security aspects
- State of the art [2] identified the following dynamic graphical formalisms:
 - Stochastic Petri nets and SANs
 - BDMP
 - Dynamic Bayesian nets
- All of them can be simulated and have a probabilistic basis
- Formalisms too specific of one domain have been discarded (e.g. Mobius/ADVISE)

[2] A Survey of Approaches Combining Safety and Security for Industrial Control Systems Siwar Kriaa, Ludovic Pietre-Cambacedes, Marc Bouissou, and Yoran Halgand

SPN & SAN

Stochastic Petri Nets and Stochastic Activity networks



Stochastic Petri nets

- Standard SPN must be used in a bottom-up manner
- Patterns can ease the model construction
- The resulting model is flat and lacks structure
- Assessing methods:
 - Markovian Petri net => all Markov analysis methods
 - Non Markovian => Monte Carlo simulation





Example taken from [3]



[3] Flammini et al. A Petri Net Pattern-Oriented Approach for the Design of Physical Protection Systems. Safecomp 2014



Assembling patterns

Security of a Metro station [3]



edf



Stochastic Petri nets pros and cons



Theoretically, unlimited modeling power (Turing machine)



Not suited for representing structure functions (nor instantaneous far reaching interactions)



Spaghetti plate syndrome => validation is very hard







Stochastic Activity Networks [4]



[4] W. H. Sanders and J. F. Meyer, "Stochastic Activity Networks: Formal Definitions and Concepts" Lecture Notes in Computer Science no. 2090, pp. 315-343. Berlin: Springer, 2001.



SAN atomic model = Stochastic Petri net + following extensions

- Activities (= transitions) can have several outputs (probabilistically chosen)
- Input gates: contain the definition of a Boolean function of the input places marking that defines the enabling of the activity, and the modification of the input places marking when the transition fires
- Output gates: contain a set of actions to perform on output places when the transition fires
- Input and output gates are defined using C++ syntax => the graph can "hide" a lot of information



Communication between submodels

Shared places Shared variable

(not apparent on the GUI)

Shared variables





SAN pros and cons



Can solve the problem of structure function representation (but not graphically)



Instantaneous far reaching interactions? Maybe, with very complicated input and output gate functions



- In a "normal" use
 - Lots of small spaghetti plates with sauce validation is still very hard
 - Sauce can be hot chili!
 - (input and output functions, shared variables are hidden)



BDMP

Boolean logic Driven Markov Processes



BDMP CV

Since 2002, Interest proven in reliability and safety engineering





- Invented and used at EDF (NPP safety, substations, data centers reliability,...)
- Complete theory and software framework

•edf

\Rightarrow Adapted to attack and defense modeling [5]

[5] L. Pietre-Cambacedes, M. Bouissou, Attack and defense dynamic modeling with BDMP. MMM-ACNS 2010, St Petersbourg, September 2010.





Tools associated to BDMP formalism



* And Petri nets!

Download: http://sourceforge.net/projects/visualfigaro/



An example of BDMP in security: attack of a remote access server





RAS attack **BDMP** – Step 0 (attack just started)





RAS attack BDMP – Step 1





RAS attack BDMP – Step 2





RAS attack **BDMP** – Attacker's objective reached





An important mechanism of BDMP: filtering of relevant events









Principles of sequences exploration in a locally defined Markov chain (Figseq)



Quantification (1/2) – Time-domain analysis

Taking advantage of the BDMP framework

- Efficient sequence exploration with trimming
 - Probability to reach the objective in a given time
 - Overall mean time to the attack success
 - Probability of each explored sequence
 - Ordered list of sequences

	0.55
	1.07 x 10 ⁵ s
	Cf. hereunder

Sequences Attack steps	Probability in mission time	Average duration after init.	Contribution
[Wardialing, Bruteforce]	0.2717	4.878x10 ³	0.4877
[Wardialing, Find_vuln, Bruteforce]	0.1272	9.7561x10 ³	0.2329
[Wardialing, Find_vuln, Exploit_vuln]	0.1272	9.7561 x10 ³	0.2329
[Wardialing, Social_eng.]	0.0136	4.8780 x10 ³	0.0249
[Wardialing, Find_vuln, Social_eng.]	0.0064	9.7561 x10 ³	0.0116



Quantification (2/2) – Time-independent

Classical values attributed to attack tree leaves

- Fixed probabilities \rightarrow (dynamically) covered by stochastic processes
- Monetary cost \rightarrow scenario cost, average attack cost
- Boolean indicators (specific requirements, properties)
 - Need of internal knowledge, internal support
 - Need of specific tool, piece of information
 - \rightarrow Characterization of selected scenarios
- Minimum attacker skills
- (Generalization) Continuous, Boolean, Discrete attributes
 - All computable thanks to the Attack tree structure



An example in safety: system to be modeled





The BDMP in KB3





Simulation of a sequence of events





Simulation of a sequence of events



On demand failures are not modeled (here)



Simulation of a sequence of events




Simulation of a sequence of events





Simulation of a sequence of events





BDMP main ideas

The total independence of leaves of a fault-tree is replaced by simple dependencies. Each leaf has two modes:

required/active (1) and not required/idle (0). Transitions between those two modes define instantaneous states in which probabilistic choices can be triggered.

Any Markov process can be associated to each mode of a leaf

Formalism "Boolean logic Driven Markov Process" (BDMP)



Graphical representation of a BDMP



triggered Markov processes Pi

A gate/basic event is TRUE when:

- a failure is present (for safety related parts)
- an attack is successful (for security related parts)



Examples of leaves behaviors (safety)



BDMP for attack modeling – Types of leaves

- Attack scenarios \Rightarrow 3 kinds of security leaves
 - Modeling of attacker's actions
 - AA (*Attacker Action*) leaves, timed leaves ($1/\lambda = MTTS$)
 - Modeling of security events
 - TSE (*Timed Security Event*) leaves, timed as well
 - ISE (Instantaneous Security Event) leaves, instantaneous (γ)













edf

Definition of required/active mode in a BDMP (1)

Very powerful concept, because it is hierarchical
 Requirement signal transmitted by the branches of the fault-tree



a gate or leaf is in mode 1 except if it receives a signal of mode 0 from : all its fathers or directly via a trigger Makes it easy to model cascade

standby redundancies/hierarchy of attack steps



Definition of required/active mode in a BDMP (2)

- The origin of a trigger can be any Boolean function of the states (true or false) of the leaves
- This origin is often a gate corresponding to a sub-tree of the fault-tree defining the structure function of the system, but it is not compulsory





What if a non standard model is needed for a leaf?

Use a «Petri leaf», associated to an arbitrary Petri net, the transitions of which are enabled/disabled according to the mode (required or not required) of the leaf





Definition of irrelevant events



- After a failure of f2, all others fi become irrelevant
- An event is said to be irrelevant if the propagation of the effects of its fulfillment in the fault-tree only concerns gates which are already in the «true» state

Number of sequences leading to top event r

= n if irrelevant events are trimmed: (f1,h; f2,h...)

Exponential function K(n) if they are not trimmed: (f1,h; f1,f2,h; f1,f3,h...)

K(n) = n + n K(n-1).For example, K(10) = 9.864.100, and $K(15) > 3.5 \ 10^{12}$



Effect of irrelevant events trimming on Markov chain size



Supposing all leaves represent repairable components



Exploitation of irrelevant events

Trimming of irrelevant events:

- Non repairable system -> dramatic reduction of the Markov chain size, with exact calculation of reliability
- Repairable system -> dramatic reduction of the Markov chain size , with approximate calculation of reliability and availability
- Note that in many cases the model with trimming is more realistic than without
 - (e.g.: electrical components, mutually exclusive failure modes, competition between attack techniques...)

Attack detection Modeling

Main points

- The IOFA distinction:
 <u>Initial / On-going / Final / A posteriori</u>
- Changes in the parameters and/or the leaves behavior
- Introduction of a "Detection status indicator" D_i
 - New Boolean function of the time, associated to each element of the BDMP



Theoretical framework extension - overview

- Introduction of a "Detection status indicator" D_i
- Some change in the modes, related to this new dimension
 - "Active" is divided in "Active Undetected" and "Active Detected"
 - Allows for parameter change, and even leaf cancellation
 - The mode is selected based on $X_i D_i$

$X_i D_i$	00	01	10	11
Mode	Idle	e (I)	Active Undetected (AU)	Active Detected (AD)

Extension of the leaves' Markov models

- New states and transitions, modeling detections & reactions effects
- New probability transfer functions



Detections/reactions for AA leaves





From Idle to Active Undetected (AU) mode





From AU mode to Active Detected mode



And so on...

Five probability transfer functions...

$$= \left\{ f_{0 \to 10}^{i}, f_{0 \to 11}^{i}, f_{10 \to 11}^{i}, f_{10 \to 0}^{i}, f_{11 \to 0}^{i} \right\}$$

• $f_{11\rightarrow 10}^{i}$ is not defined: a detected attack never comes back undetected

- ...for each type of leaf
 - Attacker Action (AA)
 - Timed Security Event (TSE)
 - Instantaneous Security Event (ISE)
- With their own Markov chains per mode
- In fact, extension of the triggered Markov process definition $\left\{Z_0^i(t), Z_{10}^i(t), Z_{11}^i(t), f_{0 \to 10}^i, f_{0 \to 11}^i, f_{10 \to 11}^i, f_{10 \to 0}^i, f_{11 \to 0}^i\right\}$



Detection and reaction integration in BDMP

Three approaches for reaction "propagation" modeling

- Strictly local incidence: straightforward but not satisfactory
- Global incidence: meaningful and direct implementation
- Extended and selective reactions: reaction triggers (not formalized)

Formal foundations – snapshot 1/3

A BDMP (A, r, T, P) is made of

- A fault/attack tree $\mathcal{A} = \{E, L, g\}$
 - a set E = G U B, where G is a set of gates and B a set of basic events
 - (*E*, *L*) a directed acyclic graph, with *L* a set of oriented edges (i, j)
 - a function g, defining the gates $(g:G \rightarrow N^*, with g(i))$ the gate parameter k)



- A main top event r
- Set of triggers T





Formal foundations – snapshot 2/3

- $P = \{P_i\}_{i \in B}, \text{ triggered Markov Processes}$ $P_i = \{Z_0^i(t), Z_{10}^i(t), Z_{11}^i(t), f_{0 \to 10}^i, f_{0 \to 11}^i, f_{10 \to 11}^i, f_{10 \to 0}^i, f_{11 \to 0}^i\}$
 - $Z_0^i(t), Z_{10}^i(t)$ and $Z_{11}^i(t)$ three homogeneous Markov processes
 - For k in {0, 1} (modes), A_k^i state-space of $Z_k^i(t)$
 - $S_k^i \subset A_k^i$, subset of states for which the leaf is true
 - $\circ \ D_k^i \subset A_k^i$, subset of detected states
 - $f_{0\to 10}^{i}(x)$ [...] $f_{11\to 0}^{i}(x)$ "probability transfer functions" with
 - $\forall x \in A_0^i, f_{0 \to 10}^i(x)$ is a probability distribution on A_{10}^i such that $x \in S_{10}^i \Rightarrow \sum_{j \in S_1^i} (f_{0 \to 10}^i(x))(j) = 1$ and $x \in D_{10}^i \Rightarrow \sum_{j \in D_1^i} (f_{0 \to 10}^i(x))(j) = 1$ • [....] x 5 $\{f_{0 \to 11}^i, f_{10 \to 11}^i, f_{10 \to 0}^i, f_{11 \to 0}^i\}$



Formal foundations – snapshot 3/3

Four families of Boolean functions of the time

Structure functions $(S_i)_{i \in E}$ $\forall i \in G, S_i = \sum S_j \ge g(i)$ $\forall j \in B, \quad S_j \equiv Z_{X_i}^{j \in sons(i)} \in S_{X_i}^{j}, \text{ with } X_j = 0 \text{ or } 1, \text{ indicating the mode in which } P_j \text{ is at time } t$ Process selectors $(X_i)_{i \in E}$ If *i* is a root of \mathcal{A} , then $X_i = 1$ else $X_{i} \equiv \neg \left[\left(\forall x \in E, (x, i) \in L \Longrightarrow X_{x} = 0 \right) \lor \left(\exists x \in E / (x, i) \in T \land S_{x} = 0 \right) \right]$ **Relevance indicators** $(Y_i)_{i \in E}$ If i = r (final objective), then $X_i = 1$ else $Y_i \equiv (\exists x \in E / (x, i) \in L \land Y_x \land S_x = 0) \lor (\exists y \in E / (i, y) \in T \land S_y = 0)$ Detection status indicators $(D_i)_{i \in F}$ $\forall i \in B, \ D_i \equiv \left(Z_{x_i}^i \in D_{x_i}^i \right) \lor \left(\exists j \in B / j \neq i \land D_i = 1 \right) \ \forall j \in G, \ D_i \equiv \left(\exists i \in B / D_i = 1 \right)$

Mathematical properties

Robustness

- Theorem 1: $(S_i)(X_i)(Y_i)(D_i)_{i \in F}$ are computable whatever the BDMP structure
- Theorem 2 : Any BDMP, defined at time t by the modes and the P_i states, is a valid homogeneous Markov process
- Combinatory reduction by "relevant event filtering"



- After attack step P_2 , all the others P_i are not relevant anymore: nothing is changed for "r" if we inhibit them
 - The number of sequences leading to the top objective is
 - n, if we filter the relevant events $(\{P_1, Q\}, \{P_2, Q\}, \dots)$
 - exponential otherwise $(\{P_1, Q\}, \{P_1, P_2, Q\}, \{P_1, P_3, Q\}, ...)$

Theorem 3: if the P_i are such that $\forall i \in B, \forall t, \forall t' \ge t, S_i(t) = 1 \Rightarrow S_i(t') = 1^*$ $Pr(S_r(t)=1)$ is unchanged whether irrelevant events $(Y_i=0)$ are trimmed or not

* This is always the case in security (~ non-repairable in reliability studies)



BDMP pros and cons



Concise, hierarchical and powerful formalism
 All dynamic behavior can be inferred from graphical

representation => relatively easy validation



BDMP (just like fault trees, Petri nets etc.) are difficult to reuse. True re-usability can only be achieved with a tool like KB3 that generates automatically calculation models



Combinatorial explosion, of course, still exists. The largest BDMP ever processed with sequence exploration had around 300 leaves. With MC simulation, problem of rare events.



BDMP are not good at all at modeling systems in which objects are created, destroyed, or even simply change places



CASE STUDY SYSTEM ARCHITECTURE



Case study of a pipeline and its control system

Example taken from: S. Kriaa, M. Bouissou et al. Safety and security modeling using the BDMP formalism: case study of a pipeline. SafeComp'2014







□ Model leaves <-> parameters estimated based on assumptions

MTTS -> security events
 MTTF -> safety events
 Probability -> instantaneous events

□ Pollution probability ~ 2e-2 for a mission time of one year

□ Attack scenarios are the most likely to happen



Transitions	МТ	Contrib.	
Name	Rate	proba	
attack_occurrence	2.28e-5	1.31e-2	0.67
access_to_RTU	0.0208		
understand_syst_operation	0.0208		
falsify_data_sent_to_CC	0.6		
falsify_data_sent_to_other_RTUs	0.6		
falsify_instructions_sent_to_equipments	0.7		
high_pumping_pressure_activation	0.7		
closing_valve	0.7		

Most probable attack scenario



Transitions	МТ	Contrib	
Name	Rate	proba	Contrib.
attack_occurrence	2.28e-5		0.87
access_to_RTU	0.0208	4.03e-4	
understand_syst_operation	0.0208		
falsify_data_sent_to_CC	0.6		
falsify_data_sent_to_other_RTUs	0.6		
falsify_instructions_sent_to_equipments	0.7		
no_realization(high_pumping_pressure_activation)	0.3		
pipe_breaks_accidentally	1.14e-5		

Most probable hybrid scenario



Transitions	MT	Contrib.	
Name	Rate	proba	
pipe_breaks_accidentally	1.14e-5	1.98e-5	1e-3
good(CC_RTU_communication_lost) good(Control_Center) good(RTU) good(faulty_operator) faill(faulty_sensor_measure) good(inter_RTU_communication_lost)	0.99954 0.999886 0.999862 0.99977 0.00023 0.9993		

Most probable accidental scenario



CASE STUDY SAFETY AND SECURITY INTERDEPENDENCIES

Mutual reinforcement



Pollution probability with and without reflex action

→ The reflex action decreased the pollution probability by 13%

➔ To succeed into causing pollution, the attacker has to deactivate the reflex action.

NB: Reflex action = shutdown decided by the set of RTUs without intervention of the centralized control system



CASE STUDY SENSITIVITY ANALYSIS



Pollution probability without attacks and with attacks without detection

→Security-related scenarios increase considerably the pollution probability
 →Conditional dependency between safety and security



CASE STUDY SENSITIVITY ANALYSIS



Effect of two detection strategies on pollution probability

bad detection: detection and reaction measures chosen arbitrarily

good detection: detection and reaction measures placed on the elements appearing in the most probable scenarios

 γ : detection probability

Conclusion and perspectives

- Importance of considering safety and security together in the risk evaluation and management process
- Petri nets and SAN: unlimited modeling power in theory, but in practice, limits due to validation problems
- BDMP still have a good modeling power, while being easier to use
 - Readability all essential information is graphically represented
 - Top-down approach, each "refinement" is manageable
 - Qualitative and quantitative analysis
 - Can easily be extended to take different probability distributions into account (requires Monte Carlo simulation). Cf. McQueen *et al.*
- Qualitative and quantitative analysis => identification of:
 - the most probable scenarios
 - the most vulnerable points in the system
 - the best detection and reaction strategies



Conclusion and perspectives

Common limitation of all these dynamic models

Estimation of security metrics (MTTS...)

Perspectives

- Robustness of the quantitative results
- Deal with uncertainties related to security parameters (uncertainty propagation)
Some references

• On BDMP & KB3

- M. Bouissou, J.L. Bon, "A new formalism that combines advantages of fault-trees and Markov models: Boolean logic Driven Markov Processes," *Reliability Engineering and System Safety*, Vol. 82, Issue 2, nov. 2003, pp. 149-163
- M. Bouissou, "Automated Dependability Analysis of Complex Systems with the KB3 Workbench: the Experience of EDF R&D," *Proceedings of CIEM 2005*, Bucharest, Romania, oct. 2005

Marc Bouissou's homepage: http://marc.bouissou.free.fr/

On BDMP & Security

- L. Piètre-Cambacédès et M. Bouissou, "Attack and defense dynamic modeling with BDMP," in *Proc. 5th International Conference on Mathematical Methods, Models, and Architectures for Computer Networks Security (MMM-ACNS-2010)*, St Petersburg, Russia, sept. 2010
- L. Pietre-Cambacedes, Y. Deflesselle and M. Bouissou, "Security modeling with BDMP: from theory to implementation," in *Proc. 6th IEEE International Conference on Network and Information Systems Security (SAR-SSI 2011)*, La Rochelle, France, may 2011
- L. Piètre-Cambacédès and M. Bouissou, "Modeling safety and security interdepedencies with BDMP (Boolean logic Driven Markov Processes)," *Proc. IEEE International Conference on Systems, Man, and Cybernetics (SMC 2010*), Istanbul, Turkey, oct. 2010.

Ludovic Pietre-Cambacedes' homepage: <u>http://perso.telecom-paristech.fr/~pietreca/</u>



QUESTIONS?

http://marc.bouissou.free.fr/

74

