Information Sciences 342 (2016) 111–131

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Probabilistic reasoning with graphical security models☆

Barbara Kordy[a,b,*], Marc Pouly[c], Patrick Schweitzer[b]

[a] INSA Rennes, IRISA, 35708 Rennes, France
[b] University of Luxembourg, SnT, 1359 Luxembourg, Luxembourg
[c] Lucerne University of Applied Sciences and Arts, 6048 Horw, Switzerland

ARTICLE INFO

Article history:
Received 30 July 2014
Revised 10 November 2015
Accepted 1 January 2016
Available online 12 January 2016

Keywords:
Attack–defense trees
Attack trees
Bayesian networks
Dependent actions
Probabilistic analysis of security
Semiring theory

ABSTRACT

This work provides a computational framework for meaningful probabilistic evaluation of attack–defense scenarios involving dependent actions. We combine the graphical security modeling technique of attack–defense trees with probabilistic information expressed in terms of Bayesian networks. In order to improve the efficiency of probability computations on attack–defense trees, we make use of inference algorithms and encoding techniques from constraint reasoning. The proposed approach is illustrated on a running example and the computations are automated with the help of suitable software tools. We show that the computational routines developed in this paper form a conservative generalization of the attack–defense tree formalism defined previously. We discuss the algebraic theory underlying our framework and point out several generalizations which are possible thanks to the use of semiring theory. Finally, our results apply directly to the analysis of the industrially recognized model of attack trees.

1. Introduction

Attack–defense trees [15] extend the well-known attack tree methodology [24,36], by considering not only actions of an attacker, but also possible countermeasures of a defender. Since the augmented formalism models interactions between an attacker and a defender explicitly and is able to capture evolutionary aspects of attack–defense scenarios, it allows for a more thorough and accurate security assessment process compared to classical attack trees. The necessity for including defensive nodes into the attack tree formalism and the advantages of attack–defense trees over attack trees were discussed in [19]. A large industrial case study presented in [2] validated the usefulness of the attack–defense tree methodology for the analysis of real-world security problems. Furthermore, in [16], we have proven that the analysis using attack–defense trees interpreted in the propositional semantics, i.e., when the trees are formalized with Boolean functions, is computationally not more expensive than the analysis using attack trees. These results show that attack–defense trees have the potential to become an efficient and practical security modeling and risk assessment tool.

Quantifying probabilistic aspects of attacks is one of the most important issues in security evaluation. Decisions concerning which defensive mechanisms or countermeasures should be implemented are based on the success probability of attacks. Furthermore, estimation of probability is necessary in order to evaluate risk related measures, because they all combine frequency or probability of an attack with its impact or costs [41,42]. Hence, a fully fledged methodology for security

☆ A preliminary version of this work has been published in [21].

* Corresponding author at: INSA Rennes, 20 Avenue des Buttes de Coësmes, F-35708 Rennes, France. Tel.: +33 02 99 84 7527.
E-mail addresses: barbara.kordy@irisa.fr (B. Kordy), marc.pouly@hslu.ch (M. Pouly), patrick.schweitzer@gmail.com (P. Schweitzer).

http://dx.doi.org/10.1016/j.ins.2016.01.010
0020-0255/

analysis needs to contain a mature framework for probabilistic computations. Unfortunately, the standard bottom-up approach for quantitative analysis of attack tree-based formalisms [19,24] can *only* be used for computing probabilities under the assumption that *all considered actions are independent*. This is a very strong assumption which is unrealistic for real-life situations.

The main contribution of this paper is to provide a *complete framework for probability computations on attack–defense trees*. Our approach combines the security methodology of attack–defense trees with the probabilistic framework of Bayesian networks. This allows us to overcome the mentioned limitation of the bottom-up approach and to *perform probabilistic computations in the presence of dependent actions*. The main strengths of our design are

- In our framework, the security model and the probability model are *kept separate*. Hence, they can be created and maintained by different experts. The overlay network is only implicitly constructed when probabilistic results are calculated.
- We show that the proposed computational procedure is compatible with the propositional semantics for attack–defense trees, i.e., that propositionally equivalent attack–defense trees yield the same probability values. Furthermore, in the case of attack–defense scenarios without dependent actions, our framework coincides with the standard bottom-up approach for quantitative evaluation of attack–defense trees. These two results show that the framework developed in this paper is a *sound extension of the attack–defense tree methodology* from [19].
- Since attack–defense trees allow for modeling of interleaved attacks and defenses, the proposed approach *improves upon existing frameworks* that combine AND-OR graphs and Bayesian networks.
- Finally, as attack trees are formally a subclass of attack–defense trees, our framework *applies directly for the analysis of the former model*. Thus, the paper also provides a full-fledged analysis technique for attack trees which are widely accepted and commonly used by industry [20].

We start by describing related work in Section 2. Then, we give a brief overview of the attack–defense tree methodology in Section 3. After recalling basic concepts for Bayesian networks, we present our framework for dependent probability computations on attack–defense trees, in Section 4. Sections 5 and 6 are concerned with methods for improving the efficiency of the proposed framework. In Section 7, we elaborate on practical aspects of the introduced methodology. We conclude the article and give an outline of future work in Section 8.

2. Related work

Attack–defense trees (ADTrees) have been first proposed in [15], where their theoretical foundations, the syntax, and numerous formal semantics have been introduced. A bottom-up procedure for quantitative analysis of ADTrees has been formalized in [19]. In [18], the most popular quantitative measures for ADTrees have been classified and guidelines for their specification have been presented. The authors of [14] have established a relation between ADTrees and game theory, by proving that ADTrees under the propositional semantics are equivalent to two-player, binary, zero-sum games. Finally, to ease the use of the ADTree formalism, a free software tool, called ADTool [17], has been developed. The ADTool supports creation, sharing and management of ADTree models and automates their quantitative bottom-up analysis.

Since classical attack trees cannot model dependencies between involved actions, several improved and alternative models have been proposed to lift this limitation [5,25,28,40]. A recent survey, by Kordy et al. [20] presents a complete state of the art in the field of DAG-based approaches for modeling of attacks and defenses. The paper summarizes existing formalisms, compares their features, and proposes their taxonomy. In the remainder of this section, we concentrate on the most prominent, existing approaches that combine AND-OR graphs with Bayesian networks and make the probabilistic evaluation of security scenarios involving dependent actions possible.

Qin and Lee are one of the pioneers in applying Bayesian networks for security analysis [33]. They propose a conversion of regular attack trees into Bayesian networks, in order to make use of probabilistic inference techniques to evaluate the likelihood of attack goals and predict potential upcoming attacks. Edges representing disjunctive refinements in the tree are also present in the corresponding Bayesian network, because they represent cause-consequence relations between components. Contrary to our interpretation, a conjunction in attack trees is assumed to have an explicit or implicit order in which the actions have to be executed. This allows to convert conjunctions into a directed path in the Bayesian network, starting from the first child, according to the given order, and ending with the parent node. The construction from [33] implies that the Bayesian network and the attack tree contain the same set of nodes. Furthermore the Bayesian network models cause-consequence relationships that correspond to the child-parent connections in the underlying attack tree. In our case, the Bayesian network depicts *additional* dependencies that represent how different basic actions are influenced by each other.

In 2008, Frigault and Wang advance a different model, called *Bayesian attack graphs*, which is used to calculate general security metrics regarding information system networks [8]. They construct a Bayesian network starting from an attack graph model which depicts how multiple vulnerabilities may be combined in an attack. The resulting directed acyclic graph contains all nodes of the original attack graph. The nodes are then populated with probability values based on the Common Vulnerability Scoring System (CVSS) [26]. These values and the conditional relationships represented by the edges of the network are encoded in the conditional probability tables assigned to each node. In a follow-up paper the approach is extended with a dynamic dimension by interpreting multiple copies of the Bayesian network as time slices [9]. Attack graphs may contain cycles, which need to be eliminated in order to construct the corresponding Bayesian network. An