Computer Networks 58 (2014) 70–86

Contents lists available at ScienceDirect

Computer Networks

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

Trustworthy placements: Improving quality and resilience in collaborative attack detection

Manuel Gil Pérez [a,*], Juan E. Tapiador [b], John A. Clark [c], Gregorio Martínez Pérez [a], Antonio F. Skarmeta Gómez [a]

[a] Departamento de Ingeniería de la Información y las Comunicaciones, University of Murcia, 30071 Murcia, Spain
[b] Computer Science Department, Carlos III University of Madrid, 28911 Leganés, Madrid, Spain
[c] Department of Computer Science, University of York, York YO10 5GH, United Kingdom

ARTICLE INFO

Article history:
Received 18 February 2013
Received in revised form 24 June 2013
Accepted 15 August 2013
Available online 7 September 2013

Keywords:
Trust diversity
Reputation
Adaptive self-configuration
Detection quality
Collaborative attack detection

ABSTRACT

In distributed and collaborative attack detection systems decisions are made on the basis of the events reported by many sensors, e.g., Intrusion Detection Systems placed across various network locations. In some cases such events originate at locations over which we have little control, for example because they belong to an organisation that shares information with us. Blindly accepting such reports as real encompasses several risks, as sensors might be dishonest, unreliable or simply having been compromised. In these situations trust plays an important role in deciding whether alerts should be believed or not. In this work we present an approach to maximise the quality of the information gathered in such systems and the resilience against dishonest behaviours. We introduce the notion of *trust diversity* amongst sensors and argue that detection configurations with such a property perform much better in many respects. Using reputation as a proxy for trust, we introduce an adaptive scheme to dynamically reconfigure the network of detection sensors. Experiments confirm an overall increase both in detection quality and resilience against compromise and misbehaviour.

1. Introduction

Many security threats in current computing environments can only be detected by gathering and correlating evidence obtained at different locations [1,2]. In some cases, evidence may come from sources over which we have little control. This is, for example, the case when organisations choose to share information about detected security events. In other cases, the integrity of the source may be questioned, for example if there is evidence that it may be malfunctioning, exhibiting a dishonest behaviour, or simply compromised.

Whatever the case, gaining such evidence in a prompt and timely manner is essential for minimising risk exposure. A major challenge in such Collaborative Intrusion Detection Networks (CIDNs) [3,4] is the ability to properly assess how much trust we can place in each piece of information and what risks we incur by believing or not believing it. Blindly accepting each report as truth is certainly dangerous. Consider, for example, an intelligent adversary who, after compromising a few detection sensors, forces them to stop sending alerts related to his intended malicious activities. Conversely, sensors may report too many false alerts, either to undermine our confidence in them or divert attention from a more serious event taking place. Ultimately, deciding what to believe depends on available information about the source and its operational context. This topic has received too little attention so far, despite being crucial for a proper functioning of collaborative

* Corresponding author. Tel.: +34 868 887645; fax: +34 868 884151.
E-mail addresses: mgilperez@um.es (M. Gil Pérez), jestevez@inf.uc3m.es (J.E. Tapiador), jac@cs.york.ac.uk (J.A. Clark), gregorio@um.es (G. Martínez Pérez), skarmeta@um.es (A.F. Skarmeta Gómez).

1389-1286/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.comnet.2013.08.026

detection efforts where there is some degree of distrust amongst parties or regarding the resilience of the sensors against attack.

In this paper, we propose a scheme aimed at increasing the *quality* of the decisions made about pieces of evidence in which we place different degrees of trust. In our proposal, we first quantify the confidence we place in sensors behaviour according to how they have behaved in the past, normally referred to as reputation [5]. However, the use of reputation alone has shortcomings, and systems based solely on assessments of past behaviour have severe limitations. We address this issue by introducing the notion of *trust diversity*. In essence, trust diversity measures the dispersion among the trust values of a population of detection sensors placed in a given domain, with low diversity values indicating that sensors have similar trustworthiness, and vice versa.

We then propose to quantify the quality of a particular sensor placement through its trust diversity, and to dynamically search for high quality placements. Informally speaking, a high quality placement is one where trusted sensors are deployed in the vicinity of others we have doubts about. Similarly, low quality placements are those where all sensors have roughly the same trustworthiness. We pursue two main goals with the use of trust diversity as a measure of placement quality:

1. Firstly, by simultaneously maximising trust diversity across all network domains we guarantee that no domain is left poorly protected (i.e., monitored exclusively by untrusted sensors).
2. Secondly, since each domain will have sensors with varying trust values, the most reliable among them can contribute to the assessment of the reputation of others, for example by identifying those that behave differently when presented with the same events.

The overall result is that sensor placements with a sufficient amount of trust diversity exhibit two interesting properties: (i) they facilitate the early identification of misbehaving sensors, a fact that implicitly contributes to a better assessment on the truth of the events they generate; and (ii) they are more resilient to compromise by external attackers.

In our scheme, trust diversity is used to dynamically place sensors when and where they are most needed, both to lessen uncertainty about what is actually happening in the network infrastructure and also to make defences more resilient to threats. Assume, for example, that contradictory events are reported by sensors in the same network area. If our confidence in all these sources is similar, we face a problem when deciding what is actually going on. Even if at present time there is little we can do about this, a careful deployment of alternative or additional sensors will help to resolve such conflicts in the future and, indirectly, to re-assess our confidence on the sensors currently deployed there.

As we discuss later, such re-deployments (or re-configurations of the monitoring infrastructure) can be triggered under many circumstances, possibly in a fully automated reconfiguration. This allows defences to react to the presence of uncertain information by seeking ways of improving their function. We believe this is a very interesting property for adaptive security systems and a prerequisite for self-healing networks.

The remainder of this paper is organised as follows. Section 2 presents some definitions and outlines our system model. Section 3 describes the adaptive model designed to reduce the uncertainty arising from discrepancies. Section 4 introduces both the reputation system used to assess sensors' behaviour and the quantification of trust diversity. Section 5 reports some experimental results to illustrate how the system can obtain better evidence by maximising trust diversity. Section 6 discusses related work in this area and, finally, Section 7 summarises our contributions and identifies future research directions.

2. Definitions and system model

Any *information system* (IS), regardless of the services it offers, can be modelled on the basis of all the elements that make up its underlying network. In this sense, the information system administrator can internally structure services and resources within a well-defined set of *domains* (D). Such a grouping may be based on the site's security policy and/or some modelling of the services, e.g., with respect to their type, so as to allow a seamless scalability as the number of services increases [6]. Governance functions could also be expected to monitor the proper operation of all services deployed in the information system, thereby defining a *monitoring system* as a surveillance centre.

Each domain imposes a set of *requirements* (R), or liabilities, for the proper operation of its services and, consequently, for the proper operation of the entire information system. As these requirements are given by the needs of each service, each domain can automatically extract its requirements from those of the services it contains. Requirements may differ in importance and so, consequently, may their monitoring. The administrator should provide a weight for each requirement $R_k \in R$, $Imp(R_k) \in [0,1]$, indicating the impact or importance on the information system if R_k is compromised.

Using the notation defined earlier, we formally define an information system with a total number of u requirements that can be demanded by m domains as follows:

$$IS = \{D_1, D_2, \ldots, D_m\}$$
$$R = \{R_1, R_2, \ldots, R_u\}$$
$$R(D_i) = \{R_{i_1}, R_{i_2}, \ldots, R_{i_x}\}$$

where each domain $D_i \in IS$ $(1 \leqslant i \leqslant m)$ is defined in accordance with its needs with x requirements $(x \leqslant u)$ for the proper operation of its services.

An example of a generic information system is depicted in Fig. 1. Note that this figure includes some other concepts not yet defined, belonging to the monitoring system, that will be formally introduced later.

This information system illustrates a distribution of the services in three domains, $IS = \{D_1, D_2, D_3\}$. The information