ARTICLE IN PRESS

COMPUTERS & SECURITY ■■ (2015) ■■–■■

Available online at www.sciencedirect.com

ScienceDirect

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

Computers & Security

Near-optimal blacklisting

Christos Dimitrakakis, Aikaterini Mitrokotsa *

Chalmers University of Technology, Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 21 August 2014
Received in revised form 25 April 2015
Accepted 26 June 2015
Available online

Keywords:
Decision theory
Blacklisting
Markov decision process
Optimal stopping
Expected loss
Network management

ABSTRACT

Many applications involve agents sharing a resource, such as networks or services. When agents are honest, the system functions well and there is a net profit. Unfortunately, some agents may be malicious, but it may be hard to detect them. We consider the *decision making* problem of how to permanently *blacklist* agents, in order to maximise expected profit. The problem of efficiently deciding which nodes to permanently blacklist has various applications ranging from efficient intrusion response, network management, shutting down malware infected hosts in an internal network and efficient distribution of services in a network. In this paper, we propose an approach to efficiently perform this blacklisting while minimising the cost of the service provider. Although our approach is quite general and could be applied to all the previously mentioned applications, to ease understanding we consider the problem in which an Internet service provider (ISP) needs to decide whether or not to blacklist a possibly misbehaving node. This is not trivial, as blacklisting may erroneously expel honest nodes (agents). Conversely, while we gain information by allowing a node to remain, we may incur a cost due to malicious behaviour. We present an efficient algorithm (HIPER) for making near-optimal decisions for this problem. Additionally, we derive three algorithms by reducing the problem to a Markov decision process (MDP). Theoretically, we show that HIPER is near-optimal. Experimentally, its performance is close to that of the full MDP solution, when the (stronger) requirements of the latter are met.

1. Introduction

We consider the *decision making* problem of blacklisting potentially malicious nodes or agents that share a resource or network based on partial information. As motivation, consider a communication network which is monitored by a network management system employed for example by an Internet service provider (ISP). All nodes pay to participate in the network (i.e. fees to the ISP). However, nodes can be of one of two types: malicious (e.g. sending spam emails, creating undue congestion as part of a denial of service attack, participating in phishing) or honest. For instance, the malicious nodes might be compromised and part of a botnet; thus, their actions disrupt other clients and might pose a threat to the ISP and its reputation.

The malicious nodes thus incur a cost due to their participation, which is not possible to measure directly. However, at each time-step (e.g. reporting period), the administrator/ISP gets a set of measurements, giving some information about the behaviour of each node during that period. These measurements could be alarms from an intrusion detection system monitoring for instance clients connected to malicious adresses. The decision problem at each time step is whether to blacklist a node, or maintain it in the system (i.e. clients of the ISP) for one more time-step.

We consider that for every honest node in the network, we have some fixed tangible gain at each time period. This would

* *Corresponding author.* Tel.: +46 31 772 10 40
E-mail addresses: chrdimi@chalmers.se (C. Dimitrakakis), aikmitr@chalmers.se (A. Mitrokotsa).
http://dx.doi.org/10.1016/j.cose.2015.06.010
0167-4048/

be the case if all participation was done through a subscription model as in ISPs. On the other hand, we incur a (hidden) cost for each malicious node that participates (i.e. maintenance cost for each node in the network, impact on the reputation of an ISP, disruption of the network). Thus, it is in our interests to kick out malicious nodes as soon as possible, but never to expel honest ones. More precisely, in the ISP setting we want the ISP to keep providing fair services to nodes (clients) that respect the service policies of the ISP and we want to blacklist and expel nodes that misbehave consistently.

We should emphasise that this is *not* an intrusion detection problem. In fact, the readings that we obtain for each node could be seen as the output of some intrusion detection system (IDS) that we consider that the service provider/administrator has at his disposal. Rather, we are more concerned about the decision making aspect as a response to a misbehaving, "malicious" node: *what is the* optimal response *to the IDS outputs, given assumptions about the cost of malicious behaviour?*

This setting of keeping suspicious nodes in the network until we become more certain about their type appears in many applications such as: (i) blacklisting clients of an ISP, (ii) shutting down malware-infected hosts in an internal network, and (iii) expelling selfish nodes from a peer-to-peer network. In all of the above cases, any single piece of information is not enough to condemn a node to blacklisting. Rather, a sufficient amount of statistics has to be collected before we are sure that removing a node is more beneficial than keeping it. In this paper, we propose and consider a number of algorithms for tackling this problem in a general setting.

More precisely, in our setting, the nodes can be one of two types: honest or malicious. However, we initially start out without knowing what type each node is. Consequently, we must gather data (observations) to reduce our uncertainty about their types. Unfortunately, we can only do so while a node remains within the network (e.g. receive normally the services of an ISP). However, the longer we maintain a malicious node in the network, the more loss we incur. Conversely, once we remove an honest node, we will obtain no more profit from it. So, the problem can be reduced to deciding at what time, or under which conditions, to remove a given node from the network, if at all. Thus, our scenario can be seen as a type of *optimal stopping* problem.

Optimal stopping problems can be modelled as a Markov decision process (DeGroot, 1970) (MDP) in a Bayesian setting, where the states equal the set of information states of the problem. While good approximate solutions to general MDPs are computationally demanding, the algorithm we propose is simple, provably efficient, requires fewer assumptions, and has similar or better performance than MDP approximations.

In this paper, we demonstrate how difficult it is to perform efficient decision making in realistic conditions and problems such as efficient intrusion response and effective network management. Although the use of MDP for solving the intrusion response has been identified before, the existing methods approach the problem in a simplistic way. More precisely, in most existing works it is assumed that the administrator has access to the normally hidden awards and thus that he/she is able to see if he/she gets a cost or gain for each decision that he/she takes. Our paper is the first one that addresses this in a more realistic manner. More precisely, we describe how the problem of blacklisting could be modeled as a *hidden reward stopping* problem.

We treat a sub-class of this problem in which some prior knowledge is given to administrators (ISP or in general a decision-maker). This prior knowledge includes that there are two types of behavior (normal/malicious) and that for each node some information is given by an existing monitoring mechanism (e.g. an intrusion detection system). We consider that these assumptions are not very far from reality since this is usually the setting under which an administrator/ISP needs to take a decision. He/she usually has access to a monitoring system and an intrusion detection system that has a detection rate and a false alarm rate. The actual intrusion detection/monitoring system is beyond the scope of this paper (e.g. a system as this proposed in Liu et al., 2014) and we consider that such a system is available to the administrator/ISP.

However, we consider that the ISP does not know the actual type of each node neither how much he gains or losses at every time step. Our theoretical and experimental analyses show that even this case (sub-class) of the problem is not that simple. We provide a novel algorithm – called High Probability Efficient Response algorithm (HiPER) – that could be employed for this efficient decision making and compare its performance with some approximations of MDP. We show that the performance of our algorithm is near optimal, even though the prior information required by HiPER is much less than actual MDP approximations.

The paper is organised as follows. In the remainder of this section we give some background and present related work. Section 3 introduces notation while Section 4 specifies the loss model. Section 5 presents the proposed HiPER algorithm as well as the bounds on the *worst-case expected loss*. Section 6 describes the decision-theoretic approaches which model the problem as an MDP and are used in the performance comparisons with the HiPER algorithm while Section 7 describes the evaluation experiments. Finally, Section 8 concludes the paper. The appendix provides proofs of technical lemmas and some useful auxiliary results.

2. Related work

As previously mentioned, our setting corresponds to a type of stopping problem. This has been extensively studied in general (DeGroot, 1970), while partial monitoring games in general have also received a lot of attention recently (Cesa-Bianchi and Lugosi, 2006). However, to the best of our knowledge, the *general hidden reward stopping problem* has not been previously studied in the literature. On the other hand, the specific application we consider can be seen as a type of *optimal intrusion response*.

The problem of intrusion response has received a lot of attention in the literature (Mitrokotsa et al., 2007a, 2007b). Most of the previous research on intrusion response has concentrated on the partially observable Markov decision processes (POMDP) formalism. Indicative publications are those by Zonouz et al. (2009), Zan et al. (2010), and Zhang et al. (2009), which have all proposed an intrusion response through modelling the process as a partially observable Markov decision process (POMDP) (Smallwood and Sondik, 1973). More precisely, Zonouz et al. (2009) proposed a Response and Recovery Engine (RRE)

Please cite this article in press as: Christos Dimitrakakis, Aikaterini Mitrokotsa, Near-optimal blacklisting, computers & security (2015), doi: 10.1016/j.cose.2015.06.010