Chapter 6
Anomaly Detection in Cloud Environments

Angelos K. Marnerides
Liverpool John Moores University, UK

ABSTRACT

Cloud environments compose unique operational characteristics and intrinsic capabilities such as service transparency and elasticity. By virtue of their exclusive properties as being outcomes of their virtualized nature, these environments are prone to a number of security threats either from malicious or legitimate intent. By virtue of the minimal proactive properties attained by off-the-shelf signature-based commercial detection solutions employed in various infrastructures, cloud-specific Intrusion Detection System (IDS) Anomaly Detection (AD)-based methodologies have been proposed in order to enable accurate identification, detection, and clustering of anomalous events that could manifest. Therefore, in this chapter the authors firstly aim to provide an overview in the state of the art related with cloud-based AD mechanisms and pinpoint their basic functionalities. They subsequently provide an insight and report some results derived by a particular methodology that jointly considers cloud-specific properties and relies on the Empirical Mode Decomposition (EMD) algorithm.

INTRODUCTION

Undoubtedly, cloud computing has evolved as a critical asset regarding the adequate deployment of large-scale, always-on services that are nowadays considered as a necessity within a range of important socio-economical ICT environments (e.g. online banking, high frequency trading systems, e-health databases/services). In practice, cloud computing is a paradigm that enables the deployment of dynamic and scalable virtualized resources to the end user through the Internet. Throughout recent years, a plethora of companies such as Google, Microsoft, Amazon and eBay have placed enormous efforts and investments towards the development, maintenance and upgrade of data-centers in order to improve their cloud-based services and further provide the best Quality of Service (QoS) as well as Quality of Experience (QoE) to the end user as indicated by Chengwei et al. (2010). Hence, their thorough analysis and proposition of

DOI: 10.4018/978-1-5225-8176-5.ch006

sufficient frameworks that support the various dimensions (e.g. security, availability, resilience) of the aforementioned domains of QoS and QoE has been prioritized in the agenda of the research community. An extremely core design element towards the healthy operation of virtualized cloud environments is regarded as the provision of mechanisms that may sufficiently confront security challenges that are likely to emerge due to the highly complex and inter-connected persona that persists in such environments.

By virtue of the intra-cloud hardware and software multi-layered nature of several components as well as the direct dependency with the Internet, the cloud composes a number of unique security concerns that need to be efficiently addressed. Apart from the networking aspect in regards to functionality and security, the cloud encompasses many technologies ranging from databases, resource scheduling, transaction management, load balancing up to operating systems and concurrency control. Thus, cloud networks trigger diverse security concerns such as storage security, data security, network security and secure virtualization. Moreover, in contrast with distributed systems as deployed over the Internet in the past where data owners had a full control over their data, their successors which are formulated by cloud environments hold intrinsic beneficial properties such as service transparency and elasticity which at the same time hold a complete control of the original owners' data. Hence, despite the end-user benefits gained by the virtual components that constitute the basis of such systems do also come with a range of threats that exploit the security holes on virtual machines (e.g. rootkit attacks on virtual machines investigated by Christodorescu et al. 2009) as well as with mutated cloud-specific Internet-based attacks that aim to compromise cloud networks (e.g. malware as studied by Gruschka et al. 2010; Marnerides et al. 2013), DDoS attacks on cloud services by Gruschka et al. 2010). According to Chen et al. (2010), blackhat hackers have already identified the potentials of the cloud since the manifestation, maintenance and operation of botnets seems to be much more efficient under a cloud paradigm in comparison with how it was in the past.

Furthermore, due to the aforementioned transparency and shared resource environment offered by the virtualized nature of the cloud, the work in Ristenpart et al. (2009), has demonstrated that hacker techniques have also transformed and evolved. In particular, it was noticed that attackers could easily construct side channels that could allow passive eavesdropping in the intra-cloud network as well as they could create covert channels that in practice send malicious data through the network. These vulnerabilities were achieved by exploiting the Virtual Machine (VM) placing method conducted by the cloud management software by allocating the attacking VM on a physical machine of the underlying datacenter and further by initiating an SSH keystroke timing attack (Song et al. 2001). Hence, the operational architecture and design of the cloud has indirectly aided the construction of new types of attacks that need to be adequately faced.

Hence, there has been a rapid development of cloud-specific security solutions that target to proactively and reactively detect cloud-specific threats either by adjusting the attack signatures of Intrusion Detection/Prevention Systems (i.e. IDS and IPS) or with statistical methods that encompass the notion of anomaly detection. IDS and IPS systems have been the main commercial solution for a number of years in the traditional Internet security domain as well as in current cloud environments and their efficiency has been questioned in several cases (Chengwei et al. 2011; Marnerides et al. 2013). Due to their signature-based concept and their full dependency on monitoring already known threats, such solutions tend to not be in a position at efficiently detecting new types of attacks that may manifest. However, the research community achieved to address this issue by suggesting a number of techniques that go beyond traditional rule and signature-based systems by implementing sophisticated statistical