Chapter 24
The World is Polluted With Leaked Cyber Data

Ivan D. Burke
University of Pretoria, South Africa & Rhodes University, South Africa

Renier P. van Heerden
University of Pretoria, South Africa & Nelson Mandela University, South Africa

ABSTRACT

Data breaches are becoming more common and numerous every day, where huge amount of data (corporate and personal) are leaked more frequently than ever. Corporate responses to data breaches are insufficient, when commonly remediation is minimal. This research proposes that a similar approach to physical pollution (environmental pollution) can be used to map and identify data leaks as Cyber pollution. Thus, IT institutions should be made aware of their contribution to Cyber pollution in a more measurable method. This article defines the concept of cyber pollution as: security vulnerable (such as unmaintained or obsolete) devices that are visible through the Internet and corporate networks. This paper analyses the recent state of data breach disclosures Worldwide by providing statistics on significant scale data breach disclosures from 2014/01 to 2016/12. Ivan Burke and Renier van Heerden model security threat levels similar to that of pollution breaches within the physical environment. Insignificant security openings or vulnerabilities can lead to massive exploitation of entire systems. By modelling these breaches as pollution, the aim is to introduce the concept of cyber pollution. Cyber pollution is a more tangible concept for IT managers to relay to staff and senior management. Using anonymised corporate network traffic with Open Source penetration testing software, the model is validated.

INTRODUCTION

In the digital age data is everywhere and is constantly being generated by everyday tasks. The safeguarding of these the data items is becoming increasingly more difficult. In the case of Internet of Things (IoT) and Machine to Machine (M2M) communications, humans are often completely absent from the data creation process.

DOI: 10.4018/978-1-5225-7912-0.ch024

Bruce Schneier warns in his book, Data and Goliath, that: "Data is the pollution problem of the information age, and protecting privacy is the environmental challenge" (Schneier, 2015). This echoes the words of Corry Doctorow, "Every gram - sorry, byte - of personal information these feckless data-packrats collect onus should be as carefully accounted for as our weapons-grade radioisotopes, because once the seals have cracked, there is no going back" (Doctorow, 2011). This brings to mind the question, if the loss of data is truly this dangerous, why is the regulation to prevent data loss not as strict as that of pollution prevention.

The term data breach commonly refers to a security incident by which sensitive data becomes exposed to individuals which are not authorized to access the data. Data breaches can usually be categorised into one of three main types of data being leaked: Personal Health Information (PHI), Personally Identifiable Information (PII), or Intellectual Property (IP). A data breach is exactly the type of loss of control that Schneier and Doctorow were concerned about. Unlike physical pollution, digital pollution often leaves no trace and can go undetected for many years.

In this paper, the current state of data breaches will be reviewed in Section 2, the review will focus on data breaches within Europe. In Section 3, current legislation regarding data breach disclosure and prevention will be discussed. In section 4, a basic experimental scenario will be defined on how to potentially model data breach detection within an organisation. In Section 5, the results of the experimental model will be discussed. In Section 6, a conclusion will be provided with recommendations for future expansion of the work.

CURRENT STATE OF DATA BREACHES

In the past few years large volumes of records have been leaked due to data breaches. Detailed data is available at www.breachlevelindex.com. Since the web service started tracking data breaches in 2013, approximately nine billion records have been leaked worldwide. The data has been filtered to exclude the United Stated and excluded data breaches where the data breach only consisted of publicly available data. Data breaches from the Unites states dwarf that of other countries, as is shown in Figure 2. According to the Breach Level Index report (2017) approximately 52 percent of breaches that were reported have no data on the amount of records leaked. The web service only reports on publicised data breaches, any non-disclosed or unknown/unconfirmed breaches have been omitted. Of the reported breaches, only 4 percent of the stolen records had been encrypted and unusable by the perpetrator of the breach.

In the following figures, we demonstrate the current state of data breaches. In Figure 1, the Technology and Entertainment industries are significant more impacted by data breaches than the other sections. This represent the strong presence of technology and entertainment in the United States compared to the rest of the world. In Figure 2 we show that the Unites States alone has the same of even more data breaches that the rest of the world combined. Thus, we concentrate on the rest of the world data breaches in this paper since the Unites States significantly skew results into their unique environment.

In Figure 3, surprisingly three Asian countries have the most data breaches: China, South Korea and Turkey. The sectors of the breaches do not form a pattern, except retail that dwarfs all other breaches in China. Although the Government related breaches do figure highly in Turkey, Mexico and India. This could present the challenges that emerging counties have in their government sectors in securing data due to skill or other shortages.