Available online at www.sciencedirect.com

ScienceDirect

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

Computers & Security

GARS: Real-time system for identification, assessment and control of cyber grooming attacks

Dimitrios Michalopoulos [a,*], *Ioannis Mavridis* [a], *Marija Jankovic* [b]

[a] *Dept. of Applied Informatics, University of Macedonia, 54006 Thessaloniki, Greece*

[b] *Dept. of Information Systems, Faculty of Organizational Sciences, University of Belgrade, 11000 Belgrade, Serbia*

ARTICLE INFO

Article history:
Received 13 September 2013
Received in revised form
30 November 2013
Accepted 25 December 2013

Keywords:
Child abuse
Fraud
Computer forensics
Intrusion detection
Security management
Risk estimation

ABSTRACT

In this paper, the Grooming Attack Recognition System (GARS) is presented. The main objectives of GARS are the real-time identification, assessment and control of cyber grooming attacks in favor of child protection. The system utilizes the processes of document classification, personality recognition, user history and exposure time recording to calculate specific risks children are exposed to during chat conversations. The above processes are repeated after each new message and three of them feed corresponding fuzzy logic controllers that provide particular but homogenized risk values as outputs. The weighted sum of the particular risk values results in a total value that indicates the current cyber grooming risk the child is exposed to, as the conversation evolves. Depending on predefined thresholds, the total risk value can be used to trigger alarms for various scopes (children, parents, etc). The practical use of GARS is demonstrated with a case study based on real grooming dialogs. Furthermore, an evaluation of the proposed approach through the discussion of applicability and performance results is discussed.

1. Introduction

The growing problem of cyber predators is a major concern for parents in their effort to protect children against malicious acts on the Internet. Moreover, the rapid development of Internet communications has brought new means for predators to approach potential victims, such as social networks, instant messaging applications and other sources (Armagh and Battaglia, 2006). The results of a successful attack on children can be numerous and catastrophic, and include psychological and physical effects such as abnormal psychology, anxiety, depression, aversion from school and social activities, behavioral and learning difficulties, tendency towards drugs and alcohol abuse and deliberate self-harm incidents (Noll et al., 2003).

In fact, there is a variety of hazards that children are exposed to. First of all, sexual exploitation attacks, known as child grooming, are primarily hazardous due to their catastrophic consequences. The term ‘child grooming’ refers to actions performed by predators with an aim to establish a sexual or emotional relationship with children (Olson et al., 2007). Moreover, cyberbullying and cyberstalking attacks refer to threatening acts through the use of online communications. On one hand, cyberbullying can be defined as an aggressive action, performed via electronic media, such as mobile phone text messages and Internet (Bauman, 2007). Recent trends involve the use of social networks, like Facebook or MySpace, where predators take advantage of the anonymity and increasing penetration to trap and attack the victim. On the other hand, cyberstalking involves activities such as harassing e-mail,

[*] *Corresponding author*. Tel.: +30 2310891868.
E-mail addresses: dimich@uom.gr, dmixalo@gmail.com (D. Michalopoulos), mavridis@uom.gr (I. Mavridis), marija.jankovic@fon.bg.ac.rs (M. Jankovic).

0167-4048/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cose.2013.12.004

flaming (online verbal abuse), mass unsolicited e-mail and forging profiles on social networks. All activities tend to threaten victims and affect personalities (Schmalleger and Pittaro, 2009).

1.1. *Background*

In terms of parental control, the ambivalent knowledge about computing and related activities, as well as physical time constraints, limit both the observation of Internet use and therefore child protection. This limitation created the need for parental control software tools which however, have their own restrictions. First, security filters can be bypassed easily by sophisticated users and a simple matching of keywords in a html text or improper images exchange is not enough to conclude whether the communication link has malicious intentions in order to restrict Internet access (Think Digit, 2009). Second, most tools lack the intelligence to understand how risk is progressing throughout conversation. This further limits their applicability with false warnings (Winder, 2013). Besides, each grooming incident is unique in nature due to time-related and physical factors. This relation becomes more complex when different personalities interact online using different styles of language. For example, in typing common expressions Internet users may use idioms and abbreviations instead of formal language. Therefore, dialog analysis should overcome these challenges and recognize potential grooming threats even if the captured text is not in a formal language format.

Parental software tools demand system resources in order to process and save data which in turn diminish the performance of computerized systems (Think Digit, 2009). Moreover, the lack of customization according to a user profile limits the applicability of such tools. This is because default settings usually prohibit access to web or conversation content recognized malicious at first sight. Although the use of these settings may result in preventing a grooming attack from realization, it rarely explains why a distinct site or conversation has been banned. This implies that neither the parent nor the child is informed about their Internet habits (Amanda et al., 2012).

Olson et al. have analyzed the communicative process predators follow to entrap their victims into a sexual relationship (Olson et al., 2007). This process refers to the time and effort a predator uses to deceive the victim into accepting sexual proposals known as the ‘cycle of entrapment’. Similar research on the nature of grooming attacks reveals that predators follow specific strategies to catch and maintain the attention of victims. For example, a predator may build a high social profile to get acquainted with a victim and turn into a friendly profile during the conversation. Interestingly, it is quite common that after the attack, predators delete cookies so as to erase their history and thus reduce the risk of being caught by the authorities (O’Connell, 2003).

1.2. *Motivation*

Despite the development of parental control software, grooming attacks are still on the rise. Therefore, there is a need, from a technological perspective, to develop effective defenses against grooming attacks. The respective effective defenses should be transparent without restricting access to any content, in order not to be disturbing for the minor user. Besides, the defenses should not store any communication data in order not to violate any legislation. Moreover, the parent should be warned in real time about the potential grooming threat. It is worth noting that controlling grooming attacks prior to their realization is of fundamental importance. The early recognition is vital for the parent in order to follow all necessary actions to eliminate the threat before the grooming attacks becomes irreversible.

1.3. *Contribution*

This paper presents the Grooming Attack Recognition System (GARS) as a tool to identify, assess and control in real-time, grooming attacks in favor of child protection. GARS operation is dedicated to the calculation of the running total risk value which reflects the grooming threat that the child is exposed to as the conversation evolves. This total risk value is the synthesis of four (4) distinct risk values combined with a weight balancing method. Three processes provide inputs in corresponding fuzzy logic controllers which calculate respective risk values as outputs. At the same time, the exposure time process calculates the corresponding risk value as a function of the time the child is exposed to possible threats. Whenever the total risk value overlaps predefined thresholds the alarm mechanism is triggered. Simultaneously, the alarm mechanism sends an instant warning signal to the parent and in addition, the child is also warned about the criticality of risk the conversation possesses with a colorful signal.

Related work as an intelligent system for grooming attack identification is the ChatCoder 2.0 (Kontostathis et al., 2009). This system identifies and analyzes predatory posts using machine learning algorithms and a set of 15 attributes related to grooming attacks. However, limitations such as the amount of false positive/negative alarms require further tuning of the system (Kontostathis et al., 2012). In parallel, the recognition of abnormal behavior and the effective risk assessment during an online communication has been considered as a critical process in most information systems (Polemi et al., 2013). In terms of grooming recognition, most major challenges refer to the security issues associated with predator’s gain of trust. Related research has revealed serious security threats that arise from the insider, in cases where the predator gains the victim’s trust (Kandias et al., 2010). Similarly, recently published research has acknowledged that forensic data from smartphones can be used towards minor protection (Mylonas et al., 2013).

The paper is organized as follows: Section 2 explains the proposed framework. Section 3 describes the reasoning underlying fuzzy logic and weight balancing. Section 4 demonstrates the applicability of GARS with an implementation example and further discusses the results of performance tests. Section 5 concludes the paper and outlines ideas for future work.

2. The proposed framework

The GARS system operates as a linear one (Fig. 1) having as input four (4) particular processes, namely: Document