TOWARDS RISK BASED PREVENTION OF GROOMING ATTACKS

Dimitrios Michalopoulos, Ioannis Mavridis
University of Macedonia, Thessaloniki, Greece
dimich@uom.gr, mavridis@uom.gr

Keywords: Risk modeling methods, grooming detection.

Abstract: The increasing incidents of children sexual exploitation through cyberspace demand for proper protection with technological defense mechanisms. This paper aims to present and evaluate methods and tools that are appropriate towards the prevention of child sexual abuse through Internet based communications. Attacking categories and strategies that predators follow are analyzed and modeled. Moreover, a comparative review of existing risk modeling methods, which is based on a set of proposed criteria, is presented. This comparison results in the conclusion that only two of the reviewed risk modeling methods can be adapted on the intended grooming attack detection system: Bayesian and Markovian. The proposed approach is concluded with a discussion on particular methods and tools for accurate attack probability calculation.

1. INTRODUCTION

During recent years Internet has been growing rapidly. Along with the World Wide Web online communication forms has grown as well. Chat rooms, instant messaging IM, social networks like facebook and MySpace are becoming very popular among children and teenagers. The spend lot of time on these online communities talking with friends, classmates or strangers. At the same time many incidents of children sexual exploitation (grooming attacks) are reported (Subrahmanyam *et al.* 2006). Parents are very concerned about how safe are their children while spending hours on the internet talking on these modern communication forms. In parallel, as they are older they do not have the proper knowledge or experience to protect properly their children from online hazards.

These reported grooming incidents bring up the syllogism about what can be done to protect minors. Many governments around the world start training police officers to identify and arrest online predators. Moreover, ISPs and social networking websites are enforced to keep log files with to help police with investigations. Besides, teachers at schools inform minors about the dangers from talking with strangers. Children are also informed on how they can avoid revealing personal information through their internet profiles and conversations.

However, the above mentioned defense mechanisms are not enough for proper protection. Because grooming attacks are based on the new technological forms of communications, additional defense mechanisms are expected to come from technological perspective and prevent grooming attacks. This paper analyzes and models the hazards that minors are exposed to while talking online. In addition, existing risk modeling methods are analyzed with a comparative review on how they can be adopted on the indented grooming attack detection system.

In section 2 the issues of Internet related hazards for youth are analyzed and modeled. A comparative review of existing risk modeling methods is presented and discussed in section 3. And the paper concludes with a discussion on methods and techniques for accurate grooming attack probability calculation.

2. PROBLEM ANALYSIS

The hazards that children are exposed to while talking online vary through age and sex and can be divided into three main categories: (a) cyberbullying, (b) sexual exploitation or grooming and (c) exposing to illegal material.

Cyberbullying refers to all kind of attacks that terrify a young user with threats for his/her life, parents, friends etc. Many times teenagers are exposed to anonymous threats through the internet (Bauman 2007). Despite the fact that it is often among online communications, very few victims talk about that with their parents. Even fewer actions are reported to police and other authorities. Indeed, only 10% of the minors who have experienced such an activity online have talked to their parents or police (Finkelhor and Ormrod 2000).The most usual types of cyberbullying are (Bauman 2007):

- Sycophantic defamation
- Assaulting and abusive messages
- Menace against life
- Social exclusion from online communication networks

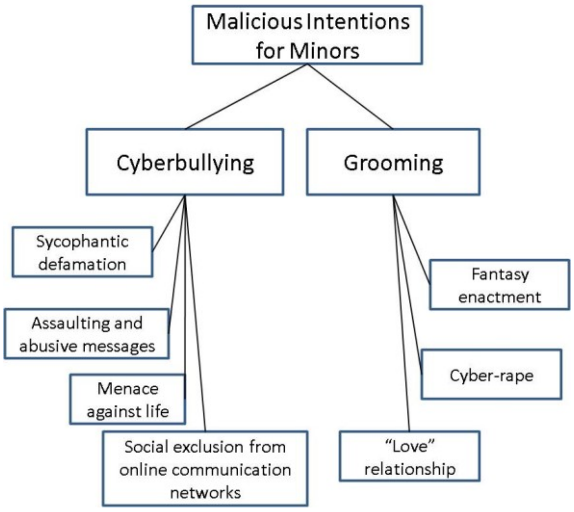
Sexual exploitation or grooming attacks are performed by people, who feel sexual attracted to children and under age people, use the modern communication methods through internet to find their victims. Predators take advantage of the anonymity provided through internet communications and build a profile suitable for their purpose starting their malicious work (Dean 2007). The research that has been published on this area has shown the there is a similar strategy followed in most of the cases by predators (Subrahmanyam *et al.* 2006); (O'Connell 2003); (Stanley 2001); (Krone 2005). Of course, each case is unique and the way it takes place is according to various characteristics such as age, location and the character of people in conversation. The types of grooming are (O'Connell 2003):

- Forming a "love" relationship
- Cyber-rape
- Fantasy enactment

Indeed, there is little difference between cyberbullying and grooming. In many cases, a cyberbullying attack follows a grooming one, as predators terrify and threat their victims in order not to talk about what has been done or intended to be done. Moreover, some times cyberbullying comes first from grooming as attackers intend to have their victims in fear for providing less resistance. Furthermore, Cyberbullying phenomena occur more often in older teenagers (13+) where grooming techniques are followed more often by pedophiles whose target range victims are children bellow 13 years old. However, this threshold of 13 years old is not limited. It is used as an attempt to categorize underage users in order to achieve better results.

Exposing to illegal material includes many types of images, video, music. Indeed, this category cannot be modeled: The World Wide Web is a huge source of information so children can search for inappropriate material many times motivated by curiosity and not by a third person as in previous two categories. Actually, exposing to illegal material is not an attack type, children many times are exposed motivated by curiosity and not by a predator.

Figure 1 bellow presents the IM and Chat attack tree that categorizes attacks on children through internet communications:



Malicious Intentions for Minors
Cyberbullying
Grooming
Sycophantic defamation
Assaulting and abusive messages
Menace against life
Social exclusion from online communication networks
Fantasy enactment
Cyber-rape
"Love" relationship

Figure 1: IM and Chat attack tree.

The analysis on cyber-threats for children brings up a major challenge: What can be done from technological perspective to reduce the hazards for children while they talk online. In this paper most of the effort is focused on grooming attacks for two reasons: At first grooming affects on children are more important and secondly cyberbullying incidents are more difficult to be detected and analyzed.

The first step for preventing grooming incidents is the analysis of how predators act and which their aims are. Similarly, O'Connell (2003) investigated grooming incidents and indicated specific stages that predators follow to perform an attack: The friendship stage, the relationship stage, the risk management stage and the sexual stage. The final one, sexual stage, includes three categories of attack as they are analyzed previously and presented at the attack tree of figure 1. For simplicity reasons the three initial stages before the sexual stage, friendship, relationship and risk management stage, are merged in one: the risk management stage including the preliminary actions of a predator before an attack. The possible transitions between the aforementioned stages that predators follow. based on the published research work (Subrahmanyam *et al.* 2006); (O'Connell 2003); (Stanley 2001); (Krone 2005), are depicted in the state-transition diagram of figure 2.

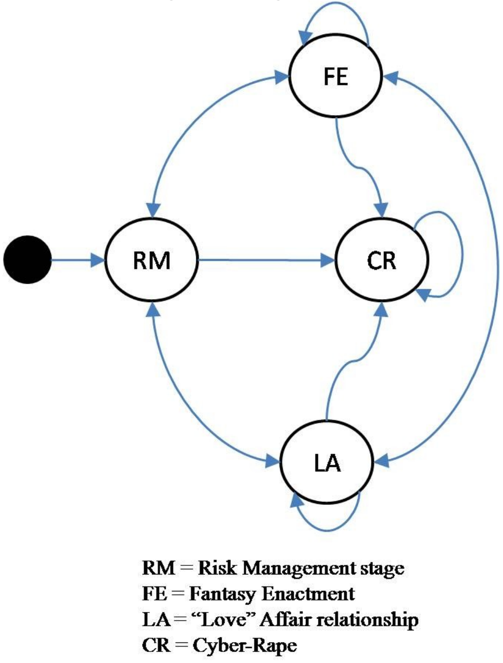


Figure 2: Grooming Attack state-transition diagram.

A predator, for example, may stay a lot of time at risk management stage (RM), perform a Fantasy Enactment attack (FE), go back at risk management stage and then perform an attack other than FE (O'Connell 2003).

3. COMPARATIVE REVIEW OF RISK MODELING METHODS

The potential system detects grooming attacks and sends a warning signal in case an attack is detected. Indeed, the decision of sending a warning signal or not is crucial. In case of a false positive, of false grooming attack detection with warning signal transmission, the system becomes irritating. Similarly, in case of a false negative, of grooming attack incident that did not be identified the consequences can be catastrophic for the minor user. Therefore, the decision making algorithm of the potential grooming attack detection system is going to decide if a warning signal will be send or not through a risk modeling process.

Indeed, which one of the existing risk modeling methods is proper for grooming attack detection? In this chapter, existing risk modeling methods are analyzed and compared based on specific criteria focused on grooming attack detection. Risk modeling methods can be very accurate on engineering, estimating the number of failures that come up each time period in many detailed form. Manufacturers provide a probability density function (PDF) of failure of each material, so implementing the proper model the risk factor can be calculated. However, how risk modeling methods can be implemented on grooming detection? Which are the criteria for such an effort? Based on the published research on this area (Subrahmanyam *et al.* 2006); (O'Connell 2003); (Stanley 2001); (Krone 2005), the following criteria are specified and proposed for grooming attack risk modeling:

- C1.Memory of the previous stages is required. The performing of a grooming attack is not based only in present stage but is related to previous ones.
- C2.There are component dependencies – items are not physically independent as the presence of one stage is depended on the previous one.
- C3.The approach is probabilistic-quantitative. The decision making algorithm about sending or not a warning signal demands for probabilistic approach
- C4.The present state should be clear. The clearance of the present state is crucial for accuracy in attack probability calculations.
- C5.The attack flow is both towards and backwards. The attack flow is not precise, the predator may return to the previous stage, stay more type and then perform a different

Table 1. Comparison of Risk Modeling Methods

	C1	C2	C3	C4	C5
Block Diagram	No	Yes	No	Yes	Towards
Attack Tree	No	Yes	No	Yes	Towards
Master Logic Diagram	No	Yes	No	Yes	Towards
Event Tree	No	Yes	No	Yes	Towards
FMEA - FMECA	Yes	Yes	Yes	Yes	Towards
Bayesian Network	Yes	No	Yes	Yes	Both
Markov Diagram	Just for the previous	No	Yes	Yes	Both
Hidden Markov Model	Just for the previous	No	Yes	No	Both
Kalman Filter	Just for the previous	No	Yes	No	Both

type of attack.

What follows is a brief review of the existing risk modeling methods with pros and cons for each one. A sort analysis is also provided about where the criteria, which are mentioned before, are matched and where they are not.

Block Diagram Method. This method usually approaches the physical arrangement of the items. For each item, the reliability-risk is calculated. Then the whole system reliability-risk can be found depending on how systems items are connected to each other: in Parallel, in Series, in Standby or a combination of the above (Modarres *et al.* 1999). For this purpose all possible path-sets and cut-sets should be found. This method is not suitable for systems with a small number of units and without a complex arrangement of them. The higher the number of items or the complexity of the system, the higher is the number of path sets and cut sets as well.

Attack tree method. According to the attack tree method, the undesired - disasterful event or "top event" is mentioned and then all the possible ways than an attacker can reach the top event are identified and presented in a diagram (Scheiner 2009). The diagram is designed is such a way that the top event is at the top of "tree", leaves represent the events and the brunches represent all possible ways that the top event can be reached. Events can be physical or software vulnerabilities, human actions and everything else that can lead to the top event. The way that events are connected to each other is with logical gates (AND, OR, Exclusive OR etc.) according to who the system is designed and how the combination of these events is needed for the top event. This method is widely used in information systems and software engineering.

Master Logic Diagram. It is mostly used in large and complex systems with several autonomous subsystems. It represents the logical representation of the system showing the relationships between independent subsystems (Modarres *et al.* 1999). A dependency matrix is extracted from the diagram and then with logical representations we can calculate the probability (and then the risk) of the "top event" occur.

Event Tree method. This method underlines the discrete states of a system. It is suitable in cases where the attack depends on the chronological order of events (Modarres *et al.* 1999). A diagram is structured and then the final event - or top event - result is calculated. It is based on binary format (occur or not occur) and as a result the final event is based on the occurrence or not of the previous events.

The above three methods, called Logic trees (Block Diagram, Attack tree, Master Logic Diagram, Event Tree), have something in common. All of them are based on Logical or Qualitative evaluation (Boolean) evaluation. However, this approach is not suitable for grooming detection project. Methods that are analyzed above are more focused on probabilistic – quantitative approach:

Failure Mode and Effect Analysis - FMEA. Failure Mode and Criticality Analysis - FMECA (Bouti and Kandy 1994).This is a very useful method in risk analysis and includes 2 steps. The first one focuses on failures and the second one with the effects the might have. This method analyzes the function of each system and all possible combinations of events which could lead to failure are discovered. Then these modes are classified according to their criticality and the consequences that they might have.

Bayesian Network. This is a very powerful mathematical model for probability calculation (Refsdal and Stolen 2009); (Krause and Clark1993); (Moore 2009). The Bayesian Network is both a graphical and a probabilistic model that is used to predict events based on known conditions.

Probabilities can be calculated both forwards and backwards. This means that the probability of the final event can be calculated knowing that some previous events have taken place or the probability of specific events is calculated knowing that the final event has taken place in combination with more observations.

Markov Diagrams. This model is widely used in economics, computer science, assurance etc. In many cases it is used in computer science as well. It is a stochastic method for prediction sequences of events and analyses the probability of each event to occur (Weisstein 2009); (Kemeny and Snell 1976); (Ayyub 2003). Markov analysis is suited on discrete systems states and performs calculations with aim to figure the probability of each state transfer. A major advantage that this model has is the fact that it does not assume that all components are completely independent. As an example, Markov model would be suitable for prediction of future states during a Monopoly play whereas Bayesian Network would be suitable for calculating probabilities during a blackjack game where the cards that have already be revealed affect the probability of the next cards (Murphy 2009). What is more, this model requires an extract knowledge of the present state just like Bayesian Network as well.

Hidden Markov model. (HMM) This is a statistical model, similar to Markov one with the difference that the present state is unobserved (Rabiner and Juang 1986); (Kemeny and Snell 1976). This method is used in computer science in terms of speech recognition, keystroke analysis, biometrics etc. As it is mentioned before, in a Markov model the present state is directly visible. For example, in a forecast prediction model, the present state is the weather today which is clearly obvious. In a hidden Makrov model the present state is not clearly obvious but each state has a probability distribution model for each possible output (Huang *et al.* 1990). Similar with the Bayesian Network, probability calculation is required in advance for each state transmission. What is more, it requires disaggregation between states, mostly on time.

Kalman Filter. Similar with the Hidden Markov model, is the Kalman filter, developed by Kalman (1960). This is a very powerful tool for predicting, through mathematical equations, the state of a linear process. The estimation is based on observations and estimates the existing noise. The main difference from the HMM is that the hidden state can take values from a non predefined space, where in HMM the hidden state is among discrete values.

Table 1 presents a synopsis of all above methods and how they are matching the predefined criteria. The comparison of the risk modeling methods denotes that two of the methods match the defined criteria: Bayesian and Markovian. Indeed, the implementation of the Bayesian demands for the calculation of conditional probabilities for the transmission in each stage. Similarly, the implementation of the Markovian demands for the calculation of the transmission matrixes for each transmission.

Indeed, the basic challenge is how these transmission-conditional probabilities can be calculated with accuracy. The proposed method for these calculations is the stochastic simulation (Modarres *et al.* 1999). The analysis of a large number of grooming incidents will lead to accurate estimations about the transmission probabilities. These grooming incidents can be found for example in the web site www.perverted-justice.com or from a live process where the researcher can pretend a minor user through chat room or IM conversations. The categorization among the attack categories will be achieved through keyword identification. Dialog analysis will indicate basic keywords that indicate the presence in specific attack stage.

4. CONCLUSIONS

The implementation of a grooming attack detection system demands for deep analysis of the methodologies that predators follow. Besides, the decision making algorithm about sending or not a warning signal, leads to a probabilistic approach for risk modeling. In this paper, most of the existing risk modeling methods are analyzed and compared according to a set of proposed criteria and in order to be implemented on the intended grooming attack detection system. Bayesian and Markovian methods seem to match the criteria. However, the implementation of each one method demands for proper calculation of the conditional-transmission probabilities. For this purpose, stochastic simulation through a dialog analysis of a large number of known grooming incidents is selected for use. This dialog analysis should also include the categorization of captured dialogs into various attack categories.

The basic advantage of the intended grooming attack detection system is instant warning. The system analyzes the captured dialogs, calculates the probability of grooming attack and then decides whether to send a signal or not. Thus, parents of minor users can be warned about a possible danger on time and can make all the necessary actions to prevent any catastrophe for their child.

REFERENCES

Ayyub, B., 2003 *Risk Analysis in Engineering and Economics* Taylor & Francis Ltd 2003 ISBN 1584883952

Bauman, S., 2007 C*yberBullying: a Virtual Menace*, National Coalition Against Bullying, Melbourne, Australia,www.ncab.org.au/Assets/Files/Bauman,%20 S.%20Cyberbullying.pdf accessed October 2009

Bouti A, Kadi A.D., 1994 *A State-of-the-art review of FMEA/FMECA* International Journal of Reliability, Quality and Safety Engineering, vol1 pp.515-543

Dean, S., (2007) Sexual Predators: How to recognize them on the internet and on the street - how to keep your kids away, Silver Lake Publishing.

Finkelhor, D. and Ormrod, R., 2000 *Characteristics of Crimes against Juveniles*, Juvenile Justice Bulletin, pp. 1-11.

Huang X., Jack, M. and Ariki Y.,1990 *Hidden Markov Models for Speech Recognition*, Edinburgh University Press. ISBN 0748601627

Kalman, R.E., 1960 *The Seminal Kalman Filter Paper* http://www.cs.unc.edu/~welch/kalman/kalmanPaper.html accessed October 2009

Kemeny, J. G. and Snell, J. L., 1976 *Finite Markov chains*, Springer-Verlag, ISBN: 978-0-387-90192-3

Krause, P., Clark, D.,1993 *Representing Uncertain Knowledge: An Artificial Intelligenc Approach*, Springer, 1 edition, ISBN-10: 0792324331

Krone, T., 2005 Queensland Police Stings in Online Chat rooms no. 301, Australian Institute of Criminology www.aic.gov.au/documents/B/C/E/%7BBCEE2309-71E3-4EFA-A533-A39661BD1D29%7Dtandi301.pdf accessed October 2009

Madan, B.B., Gogeva-Popstojanova, K., Vaidyanathan, K., Trivedi, K.S., 2002 *Modeling and quantification of security attributes of software systems, Dependable Systems and Networks*, Proceedings of International Conference on Dependable Systems and Networks vol., no.pp. 505-514

Modarres, M., Krivtsov, V., Kaminskiy, M.,1999 *Reliability Engineering and Risk Analysis*, Taylor & Francis Ltd New York 1999 ISBN 0824720008

Moore, A., 2009 *Statistical Data Mining Tutorials* http://www.autonlab.org/tutorials/ accessed October 2009

Murphy, K., 2009 *An Introduction to graphical models* http://people.cs.ubc.ca/~murphyk/Bayes/bayes_tutorial.pdf accessed October 2009

O'Connell, R., (2003) A typology of child cybersexploitation and online grooming practices, Cyberspace Research Unit, University of Central Lancashire (UK), http://image.guardian.co.uk/sys-files/Society/documents/2003/07/24/Netpaedoreport.pdf accessed October 2009

Rabiner LR and Juang BH., 1986 *An Introduction to Hidden Markov Models*, º IEEE ASSP Magazine – Citeceer

Refsdal, A. and Stolen, K., 2009 *Employing key indicators to provide a dynamic risk picture with a notion of confidence*", Springer, ISBN: 978-3-642-02055-1

Rish, I., 2001 *An empirical study of the naive Bayes classifier* IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence

Schneier, B., *Attack Trees* http://www.schneier.com/paper-attacktrees-ddj-ft.html accessed October 2009

Stanley, J., (2001) *Child abuse and the Internet*, Australian Institute of Family Studies, Melbourne, http://aifs.org.au/nch/pubs/issues/issues15/issues15.pdf accessed October 2009

Subrahmanyam, K., Smhel, K. and Greenfield, P., 2006 *Connecting Developmental Constructions to the Internet: Identity presentation and Sexual Exploration in Online Teen Chat Rooms*, National Science Foundation Grant

Taylor, C., Krings, A., Alves-Foss, J.,2009 *Risk Analysis and Probabilistic Survivability Assessment (RAPSA):An Assessment Approach for Power Substation Hardening*, http://www.csds.uidaho.edu/papers/Taylor02a.pdf accessed October 2009

Weisstein, E. W. 2009 *Markov Chain*. From MathWorld A Wolfram Web Resource. http://mathworld.wolfram.com/MarkovChain.html accessed October 2009