Parental Control and Children's Internet Safety: The Good, the Bad and the Ugly

Emmanouil Magkos, Eleni Kleisiari, Panagiotis Chanias, Viktor Giannakouris-Salalidis

Ionian University,

Department of Informatics

7 Tsirigoti Sq., 49100, Corfu, Greece,

E-mail: {emagos, p11klei, p11chan, p12gian1@ionio.gr

Abstract. In this paper we assess the threats and risks that children are exposed to as a by-product of their Internet experience. We assess good and bad strategies and practices for increasing children's online safety, from a technological, legal and ethical point of view, and explore some of the challenges that law, ethos, technology must overcome towards Internet safety for children. For example we pose the question whether a parent could ever become, intentionally or not, a threat source for a child's privacy loss. At the technical field, we run an experiment that demonstrates why parental control software has a long road ahead in meeting some minimum goals for filtering effectiveness.

1 Introduction

Children are the heart of our society. As Internet natives [Gui & Argentin, 2011], they are born and raised inside an environment where Internet and digital technologies are omnipresent. The universal broadband penetration in most countries, in concert with the advent of smart, mobile devices with touch-screen and networking capabilities, have also changed the cyber society our young children live in. A characteristic example is the exponential growth rate of online social networking (OSN) penetration among children, starting from early adolescence [Quinn & Oldmeadow, 2013].

The potential beneficial impact of (balanced) use of the Internet and digital technologies into the psychosocial well being, creativity, cognitive skills and academic performance of children has already been noted in the literature [Jackson et al, 2006, Fiorini et al, 2010]. This is reflected on the fact that most OECD countries support, starting from primary education, the development of digital skills in early childhood, while less developed countries engage initiatives such as the "one laptop per child" project[1]. Not surprisingly, the majority of parents support their young children's acquaintance with the computers and the Internet [Holloway et al, 2013].

As most things in life have dual aspects, children's exposure to the Internet can also be seen from a different, more negative theoresis. Specifically, children may be exposed

[1] http://one.laptop.org/

to a number of threats such as, among others, inadequate content and/or contacts [Marinos et al, 2011], Internet addiction [Andreou et al, 2013], other psychosocial deviation [Fiorini et al, 2010, Wang et al, 2013], loss of personal/sensitive data, etc.

Our contribution. In this paper we assess the threats and risks that online children are exposed to as well as most typical practices and strategies for reducing those risks. In our assessment we examine the problem from a ethical, legal and technological points of view, and discuss most of the challenges involved. We also run an experiment showing the low effectiveness of current parental control programs in filtering non-English (Greek) content and discussed their limitations.

This paper is organized as follows. Section 2 reviews the threats and risks pertaining to Internet use by children. Sections 3 and 4 assess good and bad security strategies related to the above risks. In Section 5 we discuss why technical solutions are not fullly ready to solve the problem. Section 6 concludes this paper.

2 Threats and Risks

2.1 Threat events

We will outline a set of the most significant threats children are facing when using the Internet. Obviously, risks related to the threats of this section can be assessed in (totally) different ways depending on the age group of children[2]; In Fig. 1 we extend the categorization of [Valcke et al, 2011] to also include typical threats related to computer/Internet information security (*e.g.*, malware, phishing, identity and data theft/loss), but also Internet addiction. Besides that, we believe that *content & contact* threat categories are sufficient containers for most threats; for example, the threats referred to as *commercial risks* in [Valcke et al, 2011] are, essentially, threats related to content, as shown in the analysis below.

Inappropriate content. Typical threats involve adult (pornographic) content, but also other inappropriate content types such as: hate, violence, racism, gambling, anorexia/bulimia, suicides, drugs etc [Livingstone et al, 2010]. A second risk is related to children consuming information which is not properly verified. This threat is exacerbated by the fact that it is not easy for children (let alone adults) to develop an

[2] A typical categorization involves children under nine (0-8 years old) [Holloway et al, 2013], pre-teenagers (9-13) and teenagers (13-18) [Blaya et al, 2012]. Such age-specific assessment is out of scope in this paper.

academic information-seeking behavior, and thus they often tend to believe bad, untruthful or unverified assertions [Livingstone & Bober, 2004, Valcke et al, 2011].

Commercial threats and spam. This category involves children treated, or in fact, manipulated to act as active consumers [Livingstone et al, 2006], for example in order to place unwanted orders or to visit unwanted commercial pages in a browser. Risks related to commercial threats are essentially content risks. Typical scenarios involve a child clicking or liking a page (as a condition to accessing some information), unwanted tabs or pop-up browser windows. Or, a child may receive a targeted advertisement, while connected in her/his OSN account, but the advertisement may involve transactions or Web visits that were done when the child was disconnected from her/his account [Cubrilovic, 2011]. Spam mail (in-)security is another well-studied topic, at least in the information security literature [Wang et al, 2013]. With children's use of e-mail increasing overtime, and given the inefficacy of spam detecting/preventing technologies, the risk related to this threat may also be high.

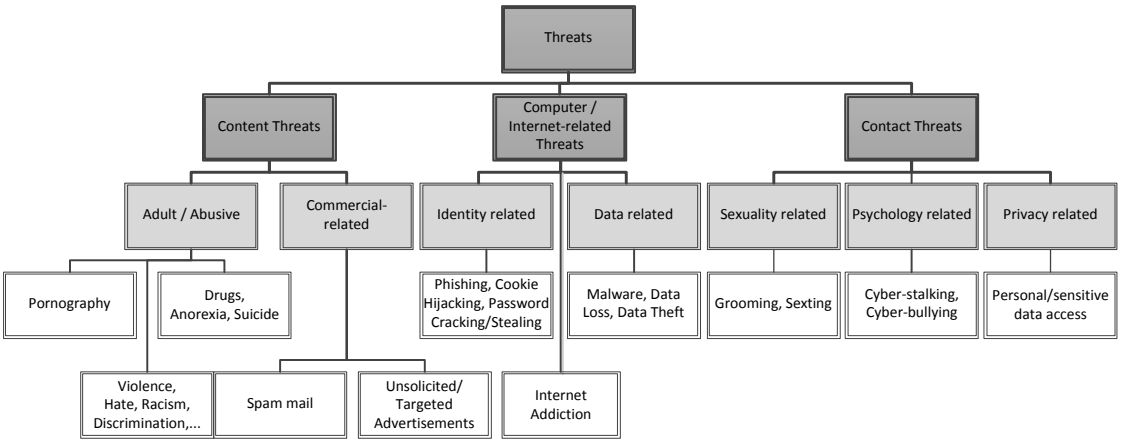

Figure 1. A categorization of cyber threats against children

Grooming. We could use the *trojan horse* allegory: Typically, online groomers attempt to establish an emotional connection by maliciously befriending their victims, offering them a pseudaisthesis of trust and confidence, with the goal to sexually abuse them during an offline meeting, and even convince them to keep this a secret afterwards [Craven et al, 2006, Marinos et al, 2011].

Sexting. A popular trend among teenagers [Blaya & Alava, 2012] involves exchanging sexually related messages or other content (*e.g.*, photos, videos) using the Internet (mostly, through an OSN site), cell phones or other digital communication equipment. The sexting threat can also be seen from other, mostly sociological and legal,

dimensions: By committing to such behaviour, children may involuntarily not only possess and process personal and sensitive data of their friends (95/46/EC, Art. 2), but also, be involved in processing exchanging explicit child pornographic content.

Cyber-bullying. A children victim of online bullying is the target of intimidating, harassing, discriminatory, provocative or other similar behaviour, exhibited though an Internet connection or other digital communication equipment, with an immediate impact to its psychism [Valcke et al, 2011, Blaya, & Alava, 2012]. The anonymity, often provided (or easily established) in a communication network can be considered as an amplifying factor for this threat.

Privacy loss. Children often divulge private information to third parties, either during a transaction with an information/service provider, or during their contacts with other persons [Livingstone et al, 2006, Byron, 2008, Blaya & Alava, 2012]. Note that in Section 4, we will also discuss a scenario where the parents themselves may be the threat source of their child's privacy loss threat.

A vulnerability related to the privacy loss threat (but also to most of the threats mentioned so far) is that children are not aware of the risks related to the inappropriate use of their personal and sensitive data.

We should note that privacy loss could be a result of, or a means to carry out, threats related to both content-related and information security-related threats; we choose the contact category to include this threat mainly because children actively (though usually involuntarily) participate in the privacy leakage, whereas, in the other two categories, children are, more or less, passively involved.

Internet addiction. Excessive exposure computer/Internet may also result in what is known as Internet Addictive Behaviour (IAB), where a child may have poor ability to control his/her self over the Internet use with detrimental effects (Whang et al, 2003, Wójcik, 2013) to the child's psychosocial well being.

2.2 Risk assessment and management

Generally speaking, and by borrowing the terminology of a typical, threat-oriented, risk assessment model[3] [NIST, 2011], every risk can be seen, at high level, as the outcome of a function that takes a number of risk parameters. More specifically, the level of a risk is reflected in the *overall likelihood* that a *threat event*, initialized by a

[3] Note that the risk assessment model in [NIST, 2011] is concerned with information-related risks.

threat source with certain characteristics (*i.e.*, capability, intent), exploits one of more *vulnerabilities* in the defence of a system and causes an adverse *impact*.

Our analysis in Section 2.1 (depicted in Fig. 1) concerned most typical threats, *i.e.,* events or circumstances our children may face during their online experience. Informally speaking, the risk level for each threat can be determined by theorizing some of the risk parameters related to a threat. For example, while the threat sources related to most content threats are individuals or groups, not directly targeting a specific child, contact-related threat sources such as grooming may be sophisticated individuals (*e.g.*, pederasts), being very concerned about minimizing attack detection, and targeting a particular child. On the other hand, most threat sources related to commercial-related threats, computer/Internet-related threats and some content-related threats (*e.g.*, the adult content subcategory) are typically well resourced organizations with sophisticated level of expertise, however not directly targeting specific children[4]. The characteristics of a threat source typically affect the *likelihood of initiation*, *i.e.*, the level of certainty we have that an adversary will initiate a threat event. For non-adversarial threat sources (*e.g.*, children omissions or errors resulting to a privacy loss, or to an installation of malicious code), such likelihood is also known as *likelihood of occurrence* [NIST, 2011]. Likelihood of initiation/occurrence can also be estimated by other factors, such as historical evidence, empirical data, expert judgment etc.

Vulnerabilities is another crucial risk parameter that needs to be assessed, as their severity can affect the *likelihood of impact*, *i.e.*, the likelihood that a particular threat event, which exploits one or more vulnerabilities, results in a severe impact, thus affecting the level of risk related to this threat. For example, vulnerabilities related to most of the threat events of Fig. 1 are (from high severity to low): Poor awareness, education at school, lack of parental control, insufficient legislation, use of insecure computer systems and digital equipment, lack of parental control software, immaturity of filtering technologies etc. The more severe is a vulnerability related to a threat event, the more higher is the likelihood of the event resulting in adverse impact(s).

Given, that the overall likelihood of a threat event resulting in adverse impact is determined by the two aforementioned likelihoods, what remains before the final risk determination is to determine the severity of impact resulted from a particular threat

[4] Observe that the targeting scores per threat may also influence the possible measures that need to be taken in order to reduce the risk related to that threat. For example, the use of parental control software (Section 3.2) may have poor results against high-targeted attacks, while increasing the security awareness (*e.g.*, through discussions and guidance from parents or at school) could be more beneficial.

event. Any assessment of impacts related to children's online safety should focus on harm made to children, and particularly (from high to low severity): Injury or loss of life, physical or psychological mistreatment [Cho, & Cheon, 2005], privacy loss etc. The *overall risk* can be determined as a function of the overall likelihood and the resulted impact, should the event occur. A full, typical, risk assessment related to risks children face when being online, is out of our scope and is left for future work.

Risk management. Any risk reduce[5] strategy related to the above risks should aim at either: a) deter/remove the threat sources (*e.g.*, a strict regulation/legislation would reduce the likelihood of initiation of a threat event), b) to remove or reduce the severity of vulnerabilities (*e.g.*, increase awareness, education, parental mediation, adopt technical measures, or finally c) to lessen the severity of impacts (*e.g.*, youth psychological support). In Section 3 we review, at a high level, some of the positive strategies/initiatives/measures related to reducing the online risks of children.

3 The Good Security

3.1 Non-technical solutions

A protecting Europe. A safer Internet for children has been one of the main goals of the Digital Agenda for Europe: the European Commission, starting from 2009, has been strongly supporting and promoting initiatives such as self-regulation[6], awareness campaigns[7a], national reporting points[7a], harmful content alert hotlines[7b] as well as measures against child sexual abuse[7c]. The proposed draft regulation on data privacy [European Commission, 2012], replacing Directive 95/46/EC, is expected to impose more strict requirements for individual consent related to the processing of children personal data by data controllers [De Hert, & Papakonstantinou, 2012].

The sensitive state and society. In most European states, a few public and nonpublic organizations have started and/or support awareness[8] initiatives both for children,

[5] Of course, we could *avoid* any risk by ensuring that a child will not have any online experience, or, at least, not be left alone during any such experience. While this could be seen, to an extent, as reasonable decision for children *e.g.*, under eight years old, it would be unrealistic in most other circumstances.
[6] European Commission. "Creating a Better Internet for Kids". Retrieved April 23, 2014. http://ec.europa.eu/digital-agenda/en/creating-better-internet-kids
[7] European Commission, Digital Agenda for Europe, Pillar III: Trust & Security. a) "Action 36: Support reporting of illegal content online and awareness campaigns on online safety for children". b) "Action 40: Member States to implement harmful content alert hotlines". c) "Action 125: Expand the global alliance against child sexual abuse online". Retrieved April 23, 2014, http://ec.europa.eu/digital-agenda/en/pillar-iii-trust-security.
[8] In Greece, popular awareness promoting places are (retrieved: April 23, 2014):

teenagers, parents and educators, promote preventative campaigns, but also detection and response strategies against the phenomenon.

Successful legislation. Most countries in the civilized world have signed and ratified the Convention on the Rights of the Child[9] [United Nations, 1990]. In addition, member states have adopted the "Lanzarote Convention" [Council, 2007] on protection of children against sexual exploitation and sexual abuse, as well as Directive 95/46/EC on the protection of individuals with regard to the processing of personal data. Generally, in most European countries laws have been passed and/or updated, to take into account the cyber environment surrounding most of the threats of Section 2.1. For example, in the Greek legal system, at the time of writing, the following behaviors, related to Internet safety for children, are penalized: Grooming and child abuse[10], processing of children's personal and sensitive data[11], offering of products/services dangerous for the psychological health of children[12].

The cautious parent. Parents need to adopt effective strategies, with an emphasis to preventative measures that will reduce most of the risks related to their children Internet safety [Blaya & Alava 2012, Lee, 2013]. Two main directions involve *restrictive mediation* and *pedagogical* (or, instructive) *mediation*. Restrictive parental mediation sets limits on what a child does/sees on the Internet. Enforcing or (monitoring the fulfillment of) such limits typically involve a combination of technical measures such as (see Section 3.2): use of parental control software, secure online services, client side security, etc. Pedagogical mediation involves risk-aware parents who, discuss with their children about those risks, and, during the early ages, may even co-surf with their children. Depending on the child's age and other (mostly, psychological) parameters [Lee, 2013], the right mixture of technical filtering/monitoring with pedagogical tactics should be pursued for an effective security. At any case, parents also need to develop better technological capabilities that will allow them to better understand most common risks, but also, if possible, to help towards reducing some of the vulnerabilities related to those risks.

---

a) http://www.saferinternet.gr/, b) http://www.0-18.gr/, c) http://internet-safety.sch.gr/,
d) http://www.unicef.gr/safety-online/, e) http://www.antibullyingnetwork.gr/.

[9] In Greece, the Convention was ratified with Greek Law 2101/1992.

[10] Art. 3 of Greek Law 3727/2008 on child abuse (implementing the Lanzarote Convention), and Art 24, Law 3500/2006 on protection against domestic violence.

[11] Greek Law 2472/1997 on data privacy (implementing 95/46/EC).

[12] Art. 7a of Greek Law 2251/1994 on consumer protection.

Child educandus. The role of educators towards children's online safety should be pedagogical and act complementarily to parental preventative behavior. [Anastasiades & Vitalaki, 2011]. It seems that younger children could be educated about Internet safety through interactive modules, *e.g.*, game-based learning [Juhari & Zin, 2013]. It is also natural to expect that children's developing of Internet skills may be reversely analogous to the risks they may be exposed to in the future [Sonck & de Haan, 2013].

3.2 Technical solutions

In the technical field of defence, we distinguish between the (more focused) mediation through parental control software, and the (more general and inclusive) *holistic approach* which pinpoints a set of tools and methods to increase children's safety.

Parental control software. Such applications are either installed on the Internet client[13], or they are part of the client's operating system[14], or they supply a network service that acts as a proxy[15] to which Internet clients are connected to. Such applications allow for controlling, *i.e.,* automated filtering and/or monitoring of information, whose superset (ideally) include: adult or inappropriate content contained in web pages and/or videos, contacts (social network, mail, chat, etc), locally executed programs, time spent on the client and/or on the Web, sending of private/sensitive data, use of microphone/camera, etc. In Section 5 we assess the filtering accuracy, related to non-English (Greek) content of several off-the shelf parental control programs. An analytic presentation and assessment of parental control software[16] is out of scope in this paper (instead, the reader may refer[17] to (SIP-Bench II, 2011, Zwaan et al, 2014).

A holistic approach to (technological) security. Filters to increase children's safety, could also be operated at several complementary levels: a) At the *service provider* level, *e.g.*, through the safe search filters employed in some popular Web search[18] or video sharing[19] engines, or through managing the privacy settings of the SN provider[20] b) At the *Web browser* level, *e.g.,* through browsing history for monitoring, cookies

[13] For example, Norton Family (https://onlinefamily.norton.com /family safety/).
[14] For example, Windows family safety (https://familysafety.live.com/).
[15] For example, the OpenDNS Parental Control proxy service (http://www.opendns.com/).
[16] Parental Controls Product Guide, 2010, http://filteringfacts.files.wordpress.com/2010/03/productguide2010.pdf
[17] For an up-to-date assessment: SIP-Bench III, 2013, Safer Internet Programme Consortium. Benchmarking of Parental Control Tools (http://sipbench.eu/index.cfm).
[18] For example, Google's SafeSearch filter (https://support.google.com/websearch/answer/510).
[19] For example, YouTube safety mode (https://support.google.com/youtube/answer/174084).
[20] For example, http://www.internetsafetyproject.org/wiki/how-create-safe-facebook-account

management for privacy configurations, Web security settings etc; c) At the *application* level, *e.g.*, through managing the Safe Senders and Safe Recipients lists in Outlook Mail[21], managing privacy and security settings in Skype[22], safe content management in iTunes[23]; d) at the *network equipment* level, *e.g.*, establish filters at the rooter level[24] so that they can be shared by all clients connecting from a specific place.

In addition, it should be noted that children may use a number of devices for communication and Internet connection. As a result, the above considerations are also relevant for mobile devices[25], game consoles[26], and even digital TV equipment[27].

Ideally, we believe that any good security strategy should also be complemented with a risk aware, security-conscious day-to-day attitude to using and managing the personal client devices children may use to connect to the Internet. For example, minimal precautions for a secure workstation should include, among other, keeping operating systems and applications up-to-date, regularly installing software patches, installing (and keeping updated) an antivirus and a firewall application, using strong passwords and managing them correctly, backing-up critical personal data, blocking third-party cookies etc. Most of these precautions should be exercised by a member of the family who has developed the necessary skills to follow them.

4 The Bad Security

Unsuccessful legislation. At the European level, the proposed new data protection Regulation [European Commission, 2012] has been criticized [De Hert, & Papakonstantinou, 2012] for allowing *further processing* for a purpose *that is not compatible with the one for which the personal data have been collected;* this could be seen as a lost opportunity to reduce the risks related to commercial-related threats against children safety. At the national level there are often some failures and inconsistencies too. For example, an obsolete law (Greek Penal Code, Art. 339, par. 1 & par. 2), dictates that if someone, younger than seventeen, performs a lewd act upon a

[21] http://office.microsoft.com/en-001/outlook-help/add-a-name-to-your-safe-senders-or-safe-recipients-list-HP005243357.aspx
[22] https://support.skype.com/en/faq/FA140/how-do-i-manage-my-privacy-settings-in-skype-for-windows-desktop
[23] http://support.apple.com/kb/ht1904
[24] http://digedags.bplaced.net/web/
[25] For example, parental control for iOS-based devices: http://support.apple.com/kb/HT4213
[26] For example, https://support.us.playstation.com/app/answers/detail/a_id/5097/~/ps4-parental-controls
[27] For example, http://www.comcast.com/Corporate/Customers/ParentalControls.html

minor younger than fifteen, the offender should be sent to a house of correction, or be treated. Or, no prosecution is to be taken, if the persons involved get married[28] (par. 3). Furthermore, legal systems need to be constantly and rapidly evolving so that risks related to the current and future threats are dealt with. For example, unfortunately there is still no specific legislation in the Greek legal system regarding cyber-bullying.

Unwilling society. Sometimes it is not enough to establish a rule, there have to exist mechanisms to enforce it, but also people who will embrace and adopt it; otherwise, the rule will be, de facto, nullified. In Greece for example, school students are not allowed to carry cell phones in school[29]. However, this rule is seldom enforced.

Confused parents. At its excess, restrictive parental mediation strategies (Section 3.1) may involve parents deciding to avoid most risks by ensuring that a child will not have any online experience, or, at least, not be left alone during any such experience. While this could be seen, to an extent, as a reasonable decision for very young children *e.g.*, under eight years, it would be unrealistic in most other circumstances. In Greece for example, with 52% of children using the Internet in their own bedroom, and with 70% of teenagers aged 13-16 years using an OSN platform [Haddon & Livingstone, 2012], such view would be a chimera. Indeed, most teenagers create an OSN profile despite parent restrictions [Holloway et al, 2013].

In another bad strategy, less restrictive parents may totally resort to parental control software for managing the risks involved. Beyond the fact that parental control technology is still immature on some aspects related to content filtering (Section 5), it should be made clear that inherently, parental control programs mostly focus to content, not to contacts. As a result, it is evident that parents should not be exclusively rely on such programs for reducing non-content related (*e.g.*, contact) risks.

The Big Brother(s). Parental control software (Section 3.2) may include capabilities that vary from keyword (or category-) based filtering and detection, where the monitoring component typically involves daily or weekly reports and/or alarms according to a pre-specified set of criteria, to *fully monitoring* all of a child's computer and Internet activities. Some programs may even perform keyword sniffing[30] or

[28] V. Sotiropoulos (2006). "About sexuality, below 15, in Greece" (in Greek). E-lawer.blog, http://elawyer.blogspot.gr/2006/05/15.html
[29] http://blogs.sch.gr/12dimch/files/2013/11/kinita.pdf (in Greek).
[30] For example, Spyrix Monitor (http://www.spyrix.com/).

complete monitoring of OSN activity[31]. Many would argue that restrictive strategies such as the above establish for our children a cyber world very close to the world envisaged by G. Orwell (1949) in his famous novel, where *the parent would play the role of the Big Brother*.

A related scenario involves *the parent's circle of OSN friends as the Big Brother*. With the proliferation of the social networks, children typically acquire an early digital footprint, (typically) without their consent, sometimes even before they are born: for example, about one out of four mothers have uploaded antenatal scans, one out of three have uploaded images of their newborn, while two out of three uploaded images of their child under two years old [Holloway et al, 2013]. With the unprecedented advent of capabilities for digitizing, collecting, storing and communicating information, in conjunction with the convergence of OSN services with Location-based Services, a massive collection of personal, sensitive and other context information related to children is posted everyday *e.g.*, on Facebook walls by their parents.

The 95/46/EC Directive (Art, 7 & Art.8, par. c) and the relevant laws, allow personal/sensitive data to be processed even without the data subject having given his consent, in order to protect the vital interests of the data subject, or when the data subject is physically or legally incapable of giving his consent. So, while the law gives legal grounds on using parental control software to filter a child's online traffic, the following two questions should be at the heart of further research on the legal and ethical grounds of parents acting as data controllers:

- Is full-scale parental monitoring ethically acceptable?
- How could children be protected from inappropriate use of personal/sensitive data from their parents?

Careless providers. Providers of social networking services have been strongly criticized for their poor or non-transparent policies concerning the privacy of their customers, in general **[**Anthonysamy et al, 2012]. Particularly OSN providers have been heavily criticized for very weak default privacy/security settings concerning their youngest members [Holloway et al, 2013]. For example, very recently, Facebook has decided to loosen privacy restrictions for teenage members [Hern, 2013]. Such policies, in conjunctions with the fact that relatively high percentages of children of 9-12 and teenagers of 13-16 years have a public profile, and/or display personal

[31] For example, Minor Monitor (http://www.minormonitor.com/).

information on their profile [Haddon & Livingstone, 2012], naturally cause much controversy. Furthermore, while typically no child under age 13 is allowed to create an OSN profile, in practice this is not enforced: In Greece, 2013, for example, more than 30% of children aged 9-12 had a profile on an OSN network [Livingstone et al, 2013].

5 The Ugly Security: The limits of the technology

5.1 Parental controls: Not ready yet

We run an experiment in order to assess the effectiveness of some popular parental control programs against abusive, non-English (in our case, Greek) content threats.

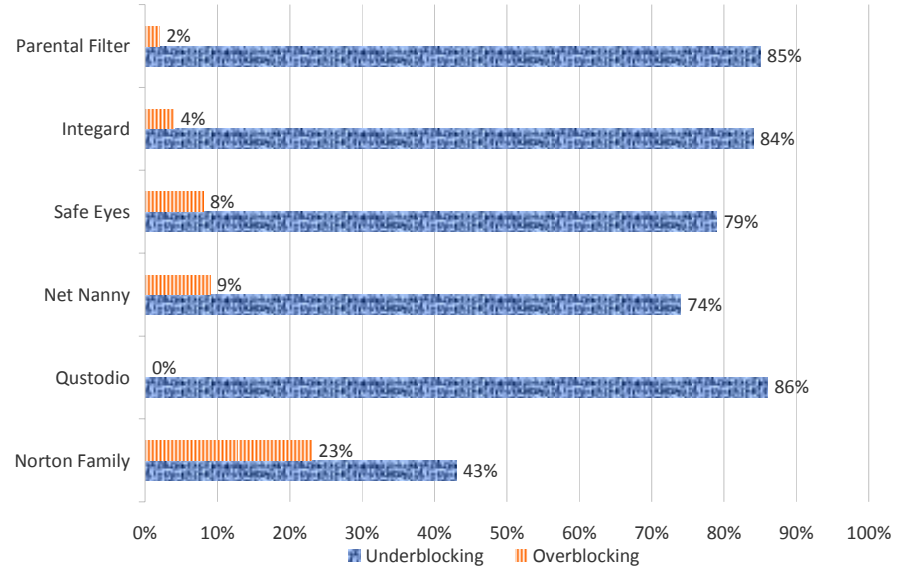

Figure 2. Underblocking and overblocking for Greek content: Overall results

The experiment was run on a Windows 7 machine, with Google Chrome browser installed. We followed the following process: we installed (one at a time) the free-trial versions of the following programs: Norton Family, Qustodio, Net Nanny, Safe Eyes, Integard, Parental Filter, trying to keep configurations the same across the programs. While under the protection of each program, we attempted to open URLs from a list of "good" and "bad" URLs. Specifically we selected 100 "good" (non-abusive) and 100 "bad" (abusive) URLs with content in Greek language[32]. Non-abusive content was selected from categories such as education, hobbies and interests, while abusive content was organized according to themes into four areas, *i.e.*, violence, drugs, pornography and other. Fig. 2 depicts the overall results, with scores for: a) underblocking (*i.e.*, permitting sites they should not be permitted) and b) overblocking

[32] For the full list, see http://di.ionio.gr/~emagos/Safety/Results.pdf.

(blocking content it should not be blocked). Fig. 3 shows the overall results for underblocking of abusive content.

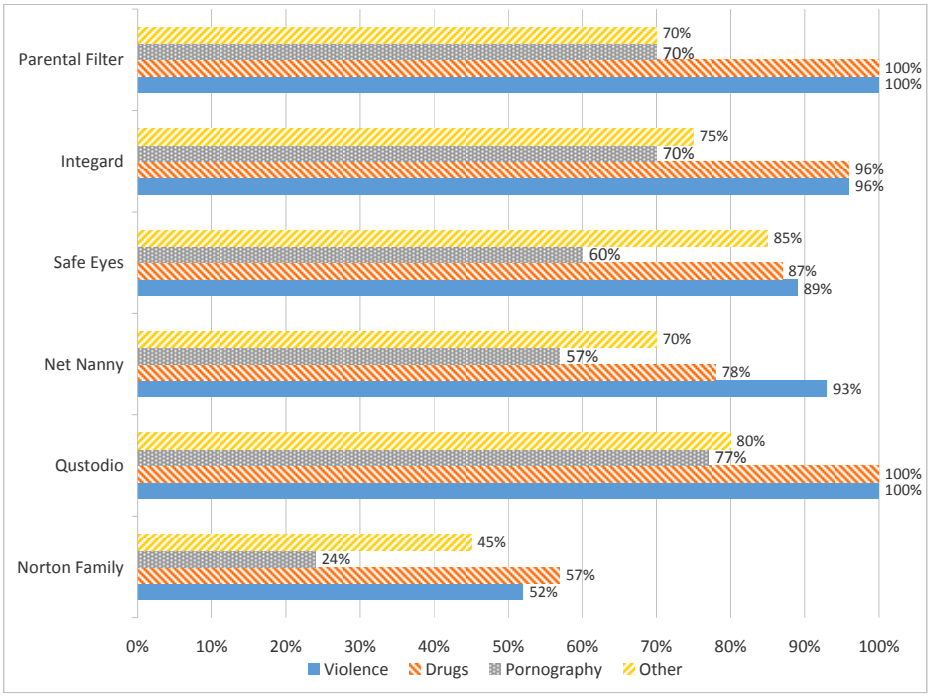

Figure 3. Underblocking of "bad" content (in Greek): Overall results per category

The results confirm what we expected, namely that:

- There is a tradeoff between underblocking and overblocking. That is, a very high level of security typically implies many false positives, while a less conservative blocking policy increases false negatives (and thus, the risk of exposure to abusive content). This trade-off is an inherent weakness of all content filtering platforms.

- Most content blocking subsystems are based on blocking URL keywords, black-listed sites etc. This is a design principle that naturally disfavors content blocking for non-English content. For the same reasons as above, and especially for filtering adult content, results are worse than what the programs scored in English-based content (SIP-Bench II, 2011).

- Abusive content non-related to porn/adult is even more difficult to detect and block. Particularly, context information about a specific site, such as semantics, syntactics and phraseology, image or other multimedia content. The underblocking results for non-adult "bad" content categories (in Greek) were discouraging, showing that much work needs to be done in the field. As a characteristic example, none of the installed platforms were able to block a site which contained photos

with dead children, amputated and tortured people: Scenes, that may stigmatize a child, much more that a typical adult scene.

Regarding the usability of the parental control programs, while there has been some progress during recent years, much has to be done: For the majority of the programs we have tested, it was relatively difficult to customize protection, while the interfaces are typically not easy to use for novice/inexperienced users.

5.2 Technologies for adult content detection: Still immature

Most typical systems performing adult content filtering, such as parental control programs, employ techniques such as contextual keyword pattern matching, which checks the context of the website (e.g., URL filters, HTTP content filtering) or blocking black-listed websites. Yet, such strategies do not seem to be adequate enough for websites that may escape the contextual filters, though containing explicit content, *e.g.*, pornographic images that should be blocked.

Today, filtering of images can be achieved through using computer vision techniques [Forsyth & Ponce, 2002], a computer science field that studies automated methods for acquiring, processing and understanding images. Computer vision can be combined with machine learning and pattern recognition to develop algorithms and methods for naked image recognition. For example, naked image detection can be achieved using a learning-based chromatic distribution-matching scheme that consists of the online sampling method and the one-class-one neural network [Lee et al, 2007]. This system uses several representative features from the naked images to verify the skin areas and the roughness feature is applied to reject confusion coming from non-skin objects. As a result, the skin area can be detected more efficiently. Another line of works aims at detecting the Region-of-Interest (ROI)[33] *i.e.*, a selected subset of samples identified for a particular purpose, for example the erotic parts of an image. In [Yizhi et al, 2013] a novel approach for ROI detection was proposed that regards the intersection of skin-color, salient and no-face regions as the ROI of the pornographic images. Support Vector Machines (SVMs) are also powerful supervised learning models for classification and regression analysis, having possible application to adult video detection [Behrad et al, 2013].

[33] Region of interest, (n.d.). In Wikipedia. Retrieved April 20, 2014, from http://en.wikipedia.org/wiki/Region_of_interest

Pornographic content detection is an important factor for children's Internet safety. While there are some algorithms and techniques that can detect adult content such as pictures and videos with a good precision, the accuracy and effectiveness results are still relatively low. Every algorithm has a small deviation from the fully precise detection, so there is always a probability of false detection (overblocking) or lack of identification (underblocking). Much research still needs to be done in the field.

6 Conclusions

In this paper we discussed most typical threats and risks that online children are exposed to today. In addition we analyzed some good and bad practices for reducing those risks and discussed several challenges related to those practices, from the points of view of law, ethos and technologies. Furthermore, we run an experiment showing that off-the-shelf parental control programs do not work well with non-English (in our case, Greek) content, and discussed their general limitations.

References

Anastasiades, P. S., & Vitalaki, E. (2011). Promoting Internet Safety in Greek Primary Schools: the Teacher's Role. Journal of Educational Technology & Society, 14(2).

Andreou, E., & Svoli, H. (2013). The association between internet user characteristics and dimensions of internet addiction among Greek adolescents. International Journal of Mental Health and Addiction, 11(2), 139-148.

Anthonysamy, P., Greenwood, P., & Rashid, A. (2013). Social networking privacy: understanding the disconnect from policy to controls. Computer, 46(6), 60-67.

Behrad, A., Salehpour, M., Saiedi, M., & Barati, M. N. (2013). Obscene Video Recognition Using Fuzzy SVM and New Sets of Features. International Journal of Advanced Robotic Systems, 10.

Blaya, C. & Alava, S. (2012). Risks and safety for children on the internet: the UK report: full findings from the EU Kids Online survey of 9-16 year olds and their parents in France. EU Kids Online, London School of Economics & Political Science, London, UK.

Byron, T. (2008). Safer children in a digital world: the report of the Byron Review: be safe, be aware, have fun.

Cho, C. H., & Cheon, H. J. (2005). Children's exposure to negative Internet content: effects of family context. Journal of Broadcasting & Electronic Media, 49(4), 488-509.

Council of Europe. (2007) Text of the Council of Europe convention on the protection of children against sexual exploitation and sexual abuse. http://www.coe.int/t/dghl/standardsetting/children/Source/Text_en.doc

Craven, S., Brown, S., & Gilchrist, E. (2006). Sexual grooming of children: Review of literature and theoretical considerations. Journal of sexual aggression, 12(3), 287-299.

Cubrilovic, N. (2011). Logging out of Facebook is not enough. New Web Order, 25.

European Commission (2012). Proposal for a Regulation on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:52012PC0011.

De Hert, P., & Papakonstantinou, V. (2012). The proposed data protection Regulation replacing Directive 95/46/EC: A sound system for the protection of individuals. Computer Law & Security Review, 28(2), 130-142.

Forsyth, D. A., & Ponce, J. (2002). Computer vision: a modern approach. Prentice Hall Professional Technical Reference.

Fiorini, M. (2010). The effect of home computer use on children's cognitive and non-cognitive skills. Economics of Education Review, 29(1), 55-72.

Gui, M., & Argentin, G. (2011). Digital skills of internet natives: Different forms of digital literacy in a random sample of northern Italian high school students. New Media & Society, 13(6), 963-980.

Haddon, L., & Livingstone, S. (2012). EU Kids Online: national perspectives.

Hern. A. (2013) Facebook defends looser restrictions on teen usage. The Guardian (Oct.18, 2013), http://www.theguardian.com/technology/2013/oct/18/facebook-defends-looser-restrictions-on-teen-usage.

Holloway, D., Green, L., & Livingstone, S. (2013). Zero to eight: young children and their internet use. EU Kids Online, EU Kids Online Network, London, UK.

Jackson, L. A., Von Eye, A., Biocca, F. A., Barbatsis, G., Zhao, Y., & Fitzgerald, H. E. (2006). Does home internet use influence the academic performance of low-income children? Developmental psychology, 42(3), 429.

Juhari, S. F., & Zin, N. A. M. (2013). Educating Children about Internet Safety through Digital Game Based Learning. International Journal of Interactive Digital Media, 1(1). 65-70.

Lee, J. S., Kuo, Y. M., Chung, P. C., & Chen, E. (2007). Naked image detection based on adaptive and extensible skin color model. Pattern recognition, 40(8), 2261-2270.

Lee, S. J. (2013). Parental restrictive mediation of children's internet use: Effective for what and for whom?. New Media & Society, 15(4), 466-481.

Livingstone, S., & Bober, M. (2004). UK Children Go Online: Surveying the experiences of young people and their parents.

Livingstone, S., van Couvering, E., & Thumim, N. (2006). Children's privacy online. Computers, Phones, & the Internet: Domesticating Information Technology, 2, 128.

Livingstone, S., Haddon, L., Görzig, A., & Ólafsson, K. (2010). Risks and safety for children on the internet: the UK report. Politics, 6, 1.

Livingstone, S., Ólafsson, K., & Staksrud, E. (2013). Risky Social Networking Practices Among "Underage" Users: Lessons for Evidence‐Based Policy. Journal of Computer‐Mediated Communication, 18(3), 303-320.

Marinos, L., Acquisti, A., Anderson, P., Cadzow, S., Carr, J., Dickman, P., ... & Wiench, P. (2011). Cyber-bullying and online grooming: Helping to protect against the risks. European Network and Information Security Agency (ENISA), Greece. Available at: http://www.enisa.europa.eu/activities/risk-management/emerging-and-future-risk/deliverables/Cyber-Bullying and Online Grooming/

NIST, SP. (2011). 800-30. Guide for Conducting Risk Assessments.

Orwell, G. (1949). Nineteen Eighty-Four. The Complete Novels, 743-925.

Quinn, S., & Oldmeadow, J. (2013). The Martini Effect and Social Networking Sites: Early adolescents, mobile social networking and connectedness to friends. Mobile Media & Communication, 1(2), 237-247.

SIP-Bench II (2011). Benchmarking of Parental Control Tools for the Online Protection of Children. Safer Internet Programme, http://ec.europa.eu/information_society/activities/sip/docs/sip_bench2_results/report_feb11.pdf

Sonck, N., & de Haan, J. (2013). How The Internet Skills Of European 11-To 16-Year-Olds Mediate Between Online Risk And Harm. Journal of Children and Media, 7(1), 79-95.

United Nations Human Rights. (1990). Convention on the Rights of the Child. http://www.ohchr.org/Documents/ProfessionalInterest/crc.pdf

Valcke, M., De Wever, B., Van Keer, H., & Schellens, T. (2011). Long-term study of safe Internet use of young children. Computers & Education, 57(1), 1292-1305.

Wang, D., Irani, D., & Pu, C. (2013, October). A study on evolution of email spam over fifteen years. In Collaborative Computing: Networking, Applications and Worksharing (Collaboratecom), 2013 9th International Conference Conference on (pp. 1-10). IEEE.

Wang, L., Luo, J., Bai, Y., Kong, J., Luo, J., Gao, W., & Sun, X. (2013). Internet addiction of adolescents in China: Prevalence, predictors, and association with well-being. Addiction Research & Theory, 21(1), 62-69.

Whang, L. S. M., Lee, S., & Chang, G. (2003). Internet over-users' psychological profiles: a behavior sampling analysis on internet addiction. CyberPsychology & Behavior, 6(2), 143-150.

Wójcik, S. (2013). Research on internet addictive behaviour among European adolescents. 13th ISPCAN European Regional Conference on Child Abuse & Neglect, Dublin, 2013.

Yizhi, L., Dong, Z., Jianxun, L., Hongtao, X., & Ying, Y. (2013, November). A Novel Approach for Region-of-Interest Detection in Pornographic Images. In 3rd International Conference on Multimedia Technology (ICMT-13). Atlantis Press.

Zwaan, J. M., Dignum, V., Jonker, C. M., & van der Hof, S. (2014). On Technology Against Cyberbullying. In Minding Minors Wandering the Web: Regulating Online Child Safety (pp. 211-228). TMC Asser Press.