A Careful Design for a Tool to Detect Child Pornography in P2P Networks

Iván Pau de la Cruz
EUITT – UPM (Madrid)
ipau@diatel.upm.es

Celia Fernández Aller
EUI – UPM (Madrid)
cfaller@eui.upm.es

Sergio Sánchez García
EUITT – UPM (Madrid)
sergio@diatel.upm.es

Justo Carracedo Gallardo
EUITT – UPM
justo@diatel.upm.es

Abstract

This paper addresses the social problem of child pornography on peer-to-peer (P2P) networks on the Internet and presents an automated system with effective computer and telematic tools for seeking out and identifying data exchanges with pedophilic content on the Internet. The paper analyzes the social and legal context in which the system must operate and describes the processes by which the system respects the rights of the persons investigated and prevents these tools from being used to establish processes of surveillance and attacks on the privacy of Internet users.

1. Introduction

In recent years, Internet has undergone considerable growth both in the number of people accessing the services it offers and the volume of information being exchanged, in addition to the speed of data transfers and the geographic areas from which one can access Internet. This growth has meant that interpersonal activities or citizens' relationships with other social actors – which were performed until recently with conventional communications methods, often in person – are now increasingly being conducted on the Internet or through other computer networks.

These new technological opportunities, with their extremely powerful tools for communication between Internet users, offer considerable advantages but also pose certain risks that must be addressed. Thus, we are seeing changes that are both swift and far-reaching in the behavior of social actors. However, we often lack either ethical or moral standards on what is licit and what is illicit, or legal regulations to adapt laws to new situations and specify certain actions as unlawful.

In some cases, even though both law and social consensus exist on the propriety of punishing certain criminal acts, the obstacles posed by the multi-national nature of the Internet make effective prosecution of crime extremely difficult. One example of this is bank robberies in which the theft of the money is carried out in Spain, but the thief is in a faraway country. In such cases, the differences between legal systems, disagreements between states and differences between policing bodies make it quite difficult to carry out effective action against the commission of this sort of crime.

In cases that lack pre-established ethical and moral standards, debates arise in society on the advisability of different restrictions when initiatives arise to condition and legally regulate certain exchanges on the Internet (music, films, literary texts, etc). The controversy occurs because these cases involve conflicts between different previously established civic rights in democratic societies, such as intellectual property, freedom of speech and privacy of communications, among others. Depending upon which of these rights is considered a priority, often irreconcilable conflicts emerge between different actors in society.

Unfortunately, there has been a substantial increase in the number of files being shared by members of peer to peer (P2P) networks containing child pornography [1]. Perhaps owing to the execrable nature of this behavior, we have found no societal debate on the advisability of police prosecution, elimination of pedophilic contents from the Internet and punishment for the guilty. Behind every picture or video with pedophilic content is a child who is being exploited and abused by unscrupulous people. According to the information available, it would seem that the police of different countries – who are quite hesitant in other cases – are quite willing to collaborate and share information in persecuting pedophiles [2].

Given this context, the authors of this paper, along with other members of a multidisciplinary research team, are participating in a joint project between the Polytechnic University of Madrid and the Spanish police aimed at developing a system with the computer and telematic tools needed to seek out and identify exchanges of data with pedophilic content, with a view to obtaining evidence that can be used by judges in formulating legal rulings.

Regardless of how repugnant and reprehensible pedophilic conduct may be, the law requires that any such system establishes safeguards to ensure respect for the presumption of innocence for all people who communicate on the Internet, and respect for the rights of the persons being investigated as well as to prevent this tool from being used to perpetrate processes of

surveillance and attacks on the privacy of the immense majority of users of P2P networks exchanging lawful data.

Thus, as described below, individuals managing the system, including persons with the status of administrators, may not engage in certain actions without the express permission – in the form of information pieces secured with robust cryptography – of persons authorized to grant such permission, presumably the judge in charge of the investigation.

The following section of this paper deals with the social and legal context in which the tool is to operate, and this shall be entirely separate from issues of technical feasibility. As a system conceived to be managed by Spanish police and judges, the discussion of legal and political conditions will center mainly on the situation in Spain and the European Union. Nevertheless, analysis is provided of the fact that much of the information in question comes from countries in which Spanish police can not act. In the event the tool should prove useful in another state, the legal context in which it would operate must first be subject to analysis.

Section 3, in contrast, is eminently technical in nature and describes in highly generic fashion the main features of the system in which the computer and telematic tools operate in searching for and identifying pedophilic contents in P2P networks. Because the system has a high degree of technical complexity, the description provided in section 3 takes a necessarily generic and summary form. Thus, its technical descriptions are comprehensible to readers lacking specialized knowledge of computing and communications. Section 4 offers a summary discussion on the performance of the tool with regard to respect for the rights of individuals to be investigated, and section 5 presents some conclusions.

2. Social and legal context

2.1. Overview

The emergence of Internet, along with other technological advances, has precipitated a huge leap in the volume and nature of child pornography available. Internet acts not only as mechanism for creating, displaying, exhibiting, commercializing or distributing child pornography, but also as vehicle for pedophiles to make contact with new victims and deceive them. Internet has brought a revolution in the methods of access, operation and dissemination [3].

The typological of criminal activities on the Internet [4] is quite wide: swindles, crimes against intellectual and industrial property, damage to computers – viruses and so on – crimes against privacy, disclosure of business secretes, telecommunications fraud – phreaking and wardriving – among others [5]. This paper presents some of the main features of a research project being conducted jointly by the Polytechnic University of Madrid and the Spanish police, having chosen prosecution of the crime of distributing child pornography through P2P networks because this practice has generated some of the highest levels of alarm in society and it most seriously affects the dignity of minors.

This crime is addressed in article 189 of the Spanish Penal Code. It penalizes the use of minors in pornographic spectacles, the distribution or commercialization of child pornography material and the personal use of said material.

At this time, the Spanish Ombudsman has proposed the inclusion of two further types of conduct in the wording of the Penal Code: simply viewing child pornography without storing or distributing it and attendance at spectacles of child pornography.

First of all, the Convention on Children's Rights recommends defining as a crime access to child pornography through information and communications technologies: that is, conduct consisting of the simple viewing of child pornography without storing or distributing it. In short, the idea is to provide a penal reaction against the "potential stimulus of sexual exploitation of children this conduct implies" (in this regard, the proposed decision-making framework of the Council of the European Union on combating the sexual exploitation of children and child pornography, 2001/0025). As stated in Query 3/2006 of the State Attorney of Spain, lawmakers sought to protect the sexual indemnity of minors by setting up penal protections. In accordance with this argument, it is understood that Spanish Penal Code, just as it penalizes the mere possession of child pornography, should contain provisions that punish the mere access to child pornography on the Internet.

In like manner, the Council of Europe agreement recommends, while allowing for the possible presentation of reserves on this point, punishment as a criminal offence the knowing attendance at pornographic spectacles in which minors participate. On this point, the Spanish Penal Code penalizes, among other forms of conduct, the financing of exhibitionist or pornographic spectacles in which minors are used but, in a manner that is similar to the consumption child prostitution, the mere attendance as a spectator at such displays is not expressly and specifically penalized. In view of the seriousness of the current problem and the alarm it has aroused, it would seem appropriate for Spanish law to classify as a knowing attendance of spectacles of child pornography as a criminal offence. When placing such conduct in the context of Internet, it would seem reasonable to believe that the mere conscious viewing of pedophilic contents constitutes an improper and morally reprehensible act that could also be considered criminal; although, if they were

classified as such, the nature of the Internet would make it highly difficult to obtain proof of an offence. Hence the importance of persecuting the production and distribution of pedophilic contents.

The table below contains the texts of laws now in force domestically in Spain, in the EU and internationally. We might say that minors enjoy a sufficient degree of protection in the exercise or enforcement of their rights.

Table 1. Related spanish legislation

DOMESTIC LAW
Law for the legal protection of minors Penal Code (art. 189) and the Law on Criminal Justice Law on Information Society Services

Table 2. Related international legislation

INTERNATIONAL LAW
Convention on the Rights of the Child Optional protocol on the sale of children, child prostitution and child pornography Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services. Council framework Decision on combating the sexual exploitation of children and child pornography (2004/68/JHA). Council of Europe Committee of Ministers Recommendation R (91) 11 concerning sexual exploitation, pornography and prostitution of, and trafficking in, children and young adults Recommendation (2001) 16, on protection of children from sexual exploitation Convention on Cybercrime of 23 November 2001 Lanzarote Convention of 2007

2.2. Conflicts of rights

The system for detecting child pornography under development, which is briefly described in the following section, is designed to safeguard the rights of minors. However, this must be achieved without infringing upon the rights of Internet users. In the event of a conflict between these rights, the jurisprudence of the courts to date must be observed.

We shall now address some of the conflicts of rights that will appear throughout the investigation.

2.2.1. Secrecy of communications – minor rights

In article 18.3, the Spanish constitution enshrines the right of secrecy in communications; article 8.2 of the European Convention for the Protection of Human Rights and Fundamental Freedoms of 1950 and Convention 185 of the Council of Europe on Cybercrime – which is a domestic Spanish law – clearly define secrecy of communications, including "any computer data related to communication that indicate the origin, destination, route, time, date, size and duration of the communication or the type of underlying service." This also includes any subscriber data in the possession of the service provider: data on traffic, content, communication type, identity, physical address, billing, payment, the location of communication equipment, time of service and other aspects. The European Court of Human Rights has held that approval of an intervention requires sufficient certainty and foreseeableness that the intervention will be successful in proving a specific crime (Malone case, 2 August 1984), criticizing lack of specificity in the cause for intervention and the lack of evidence of criminal conduct as elements intrinsic to abuse and arbitrariness by public authorities.

This must be taken into account when considering the use of the technology analyzed herein.

2.2.2. Data protection – minor rights

A number of recent rulings of the Spanish Supreme Court (Sentences 292/2008 of 28 May, 279/2008 of 29 October, 873/2009 of 23 July, 795/2009 of 28 May) have held that the localization of IP addresses of computers used to commit crimes against art. 189 of the Penal Code is lawful and does not violate the secrecy of communications. Among other reasons, because IP addresses are easily localizable for any person, as they are in the public domain.

However, with regard to protection of privacy and personal data, identification of the owner of a given telephone terminal or user of Internet can be obtained legally only on the basis of consent or a court order. In this regard, it must be recalled that law 25/2007 of 18 October on the preservation of data on electronic communications is a huge step forward in achieving authorization of these provisions of data to the state security services.

Here, the technology being developed will adhere to these legal standards, so that any identification of users of P2P networks shall be done following issuance of the required court order.

2.2.3. Freedom of speech – minor rights

This freedom also has limits. Thus, article 20.4 of the Constitution sets forth some of these, such as the right to honor, privacy, own image and the protection of children and young adults [6]. The latter point is highly germane to the issue addressed herein, and it mandates protection of minors. The Spanish Constitutional Court, in its sentence 66/82, reminds us of this and aligns itself with the

European Convention for the Protection of Human Rights and Fundamental Freedoms.

On the basis of the above, pedophiles may seek to claim their right to freedom of speech, but this freedom can never be used against the dignity of minors, as in the distribution of child pornography.

2.3. Legal challenges

With respect to legal implications, major difficulties exist in *determining liability*, as different operators are involved in the communications chain, such as the telecommunications operator, the Internet service provider, the service provider, etc. Spanish law, such as the Law on Information Society Services, establish the principle of exempting services providers until they are actually cognizant of the unlawful nature of the content [7].

These difficulties mount when links in the chain are located in different countries. This is the second challenge: *legal jurisdiction*. Although the rule of thumb in penal matters is to apply the law of the place of commission of an offence, the jurisprudence of the Supreme Court clearly opts for ubiquity. At first, it accepted the theory of results, but it has slowly moved in the direction of ubiquity. This is clear in ruling 3.2.97, when it states that "…the facts must be considered to have been committed in all places in which the action has occurred and in places where results have arisen …" Without a doubt, this is the only theory that enables prosecution of cybercrime with a certain level of assurance.

3. System Architecture

The architecture of a system must reflect the needs of users and meet their expectations. Thus, a first level of abstraction is defined that provides both the functional blocks comprising the system and the relationship between these blocks and users. Cooperation among all functional units in accordance with previously established rules and restrictions gives rise to the ultimate functionality of the system.

As part of the diagram for a high-level architecture, users can be considered an ordinary functional component. Thus, and due to the intelligence and flexibility of trained users, users often have total control of the architecture of the system and are even assigned a part of the functionality the system is designed to implement. This means that the proper use of the system largely depends on how users make appropriate use of it.

This user-attributed capacity to control systems is welcome in many contexts, as the user becomes an intelligent supervisor. However, applications in which citizens' secrecy of communications must be ensured, administrators' capacity for total control can become a threat.

Therefore, this paper will now propose a high-level architecture in a multi-protocol system for searching for pedophilic content. The proposed system safeguards user rights to secrecy of communications until an authority with the necessary powers legally allows an invasion of the privacy of suspect users following an analysis of the digital evidence provided.

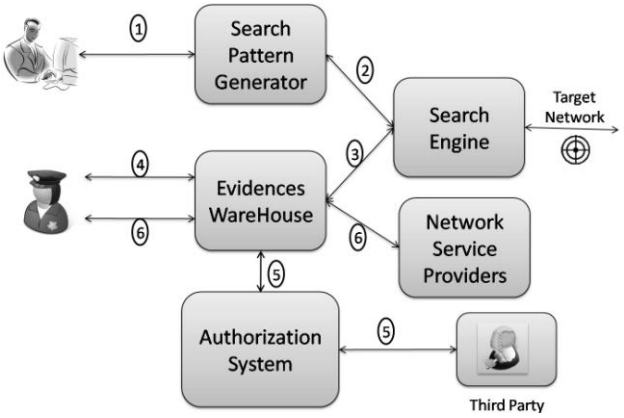

Figure 1. General Architecture

The design of the application is oriented towards safeguarding the rights of users that are under investigation (potentially, any user in a targeted network). As shown in figure 1, the system is composed of several functional elements and several steps are required to conclude an investigation. To describe the architecture, a summary explanation will be provided of the steps enumerated in the figure.

1. Preparation of the search. A system operator communicates with the search pattern generator to prepare the search to be performed. The system searches for resources – i.e., files in different formats – that have previously been classified as pedophilic. The operator can select what resources, or files to include in the search – it would not be feasible to include all possible resources – the networks in which the search will be performed (for example, in the Ares, BitTorrent, eDonkey networks, etc.), the minimum number of matches for just one user, etc. The search pattern generator also generates evidence in a protected format (better if read only) to allow for subsequent editing. This will ensure that the system is used only for its initial purpose: to search for pedophilic material online.

2. Once the search has been configured and recorded for subsequent analysis, it is sent to the search engine. The search engine connects to each of the networks and searches for the content configured. The search engine is designed to allow for a proper distribution of the load in different machines, running the process in phases and other functionalities not discussed herein, as they are beyond the scope of this paper.
3. As the search engine finds occurrences, it stores them in an evidence warehouse. An occurrence is the localization of a user that possesses a certain number of resources of pedophilic content. The configuration of the search engine (phase 1) contains a number of coincidence thresholds that will add a user to the evidence warehouse and the type of information to be stored on a user who possesses pedophilic content.
4. When the search process is complete, other system operators can make a query to ascertain if there are other users that store content that is potentially pedophilic, according to the criteria established in the search. These operators have access to the general parameters of the users, indicating the number of resources with pedophilic content in their possession, but not information that can uniquely identify them. If identification of users is seen as necessary to pursue the investigation, authorization from the relevant authorities must be requested, presumably the judge in charge of the investigation. Until this authorization is granted, the system must not allow any operators to gain access to this data, regardless of the operator's role. To attain this objective, information on users found must be protected even from system administrators themselves (this role has the greatest level of control over the system). Thus a mechanism based on robust cryptography has been designed that enables this protection. The implementation of this mechanism is one of the principal challenges of this work.
5. Upon receiving a request from an appropriate operator, the system summarizes the evidence obtained from a set of users and requests authorization from the proper authorities, namely the judicial system. The system must provide these entities with the pertinent tools in order to execute this process in a manner that is efficient and that guarantees due process. Further, the system stores proof of all requests and concessions of authorizations made so as to allow for subsequent analysis, if necessary.

6. Finally, and upon reception of the necessary authorization, the system discloses information on users who are in the possession of pedophilic content. First, it accesses the Internet services providers and requests information needed to identify the citizens under investigation. After gathering this information, the system provides access to the information to uses with sufficient authority. The entire process is recorded for any subsequent auditing.

3.1. Technology challenges

Technologically, the proposed architecture poses challenges that must be overcome in order to meet the requirements of the system. These are the most important ones in terms of user guarantees.

- System audit. One of the foundations of the guarantees offered by the system is its capacity to be audited by third parties. Any accredited organization may audit the use of the system. Making the system sufficiently self-descriptive, inalterable and unable to disclose sensitive information either on users or the operations of police officials working on investigations is an objective that will be difficult to achieve. The sub-system is being defined by a multi-disciplinary team that is taking account not only of purely technological aspects, but also political and legal ones.
- Use of formats protected from manipulation. Most present-day storage systems can be manipulated with certain tools. Although some are designed to prevent manipulation, they are not usually capable of storing large amounts of information. Creating storage systems that meet these two conditions – high capacity and resistant to illicit manipulation – is another of the challenges system developers will have to face.
- Reversible anonymization of evidence stored. Reversible anonymization of information has been dealt with in previous research papers in which the authors of this paper [8]. In this system, the specific problem has been studied of anonymizing information stored centrally – with total control of hardware and software by one of the actors in the process – and enabling the theoretically multiple authorizing entities to break that anonymity.
- Creation of viable systems of interaction for managing authorizations. Authorization entities are the foundation for safeguarding users' rights with respect to the secrecy of communications. Only these entities, on the basis of evidence that is proper in both form and content, can decide whether to disclose the identity of a user to the

agents involved. Designing a system that will enable them to efficiently perform this task in terms of usability without infringing on the due process rights required throughout the process is another of the challenges faced.

4. Discussion

As mentioned in section 2, which describes the social and legal context, there are conflicts of rights in the structure and methodology of persecuting child pornography that must be taken into account. These conflicts can be summarized as follows:

1. Conflict between the duty to persecute the criminal offence of producing, storing and distributing child pornography and the citizens' right to secrecy in their communications.
2. Conflict between the right to honor and the image of children and young people and the right of freedom of speech.
3. Conflict over the jurisdiction for criminal offences committed in different legal regions.

The system presented herein adequately manages the first of these conflicts. Under Spanish law, proper resolution of the conflict demands requires the involvement of the courts and network access providers on the basis of the proper terms and time frames.

This collaboration exists at present, but it is not managed in a comprehensive manner by a single system working as an interface between the different entities involved. Instead, each of the entities possesses information to which it is not entitled under the law on users suspected of handling content with child pornography.

This architecture has been designed following a study of the methodology used by institutions responsible for persecuting this type of criminal offence in Spain. The architecture fulfills two principles that are fundamental for resolving the conflict:

1. The tool can be used only for the detection of pedophilic content.
2. Information identifying users will remain anonymous until authorization has been received from a judicial authority.

The first principle is provided by the recording in a non-manipulable form of the search to be performed and the possibility of external audits. Thus, the search cannot include content not related to the handling of child pornography resources, as this would be detected.

The second principle is resolved by cryptography-based solutions that can keep information anonymous until the proper conditions have been met. Information on the user can be retrieved from the network itself in which the occurrence arose or by accessing the information stored in the network access provider once necessary authorization has been received.

Further, a judge may make a summary of the evidence found and grant authorization at his or her own discretion. Both reception of the authorization request with the attached evidence and the granting or denial of authorization shall be recorded in a form that is auditable by third parties, also including all security mechanisms needed for the process to uphold the due process rights and obligations of all involved parties. One traditional challenge in this type of system, which involves different participants with different technological capabilities and major security requirements, is system usability. For the system to be accepted by all the actors involved, security mechanisms must implemented that not only support the safeguards and obligations mentioned above, but also manage the trust of all actors in their transactions. To reinforce the usability of a system like the one proposed herein, with so many security requirements, personal cryptographic devices will be used, such as smart cards, cryptographic tokens and so on. In terms of usability, the effectiveness of these devices seems to have been proven for a broad spectrum of the public, as shown by the massive implementation by some states – such as Spain – of mandatory identification documents for citizens [9].

Police access to the databases of Spanish network access providers is not automated at present. The project of which the system described herein forms a part is also working to determine what type of information access providers must make available and under what conditions, so as to allow an investigation to move forward without infringing on the rights of the users of their services.

5. Conclusions

The exchange and commercialization of content with child pornography through the Internet has increased recently with the use of P2P networks. Significant alarm has been generated in society and a broad social consensus exists favoring their elimination from the Internet and the punishment of individuals committing these criminal offences.

The use of automated computer and telematic tools capable of operating day and night on different computers is a major step forward in searching for and eliminating pedophilic content on the Internet, if compared to the limited actions of which a single human being is capable when operating alone.

It is both technologically feasible and socially imperative for such tools to respect the civic rights of

persons being investigated, and to not be used to set up processes of surveillance and attacks on the privacy of the overwhelming majority of users of P2P networks who exchange lawful information. Therefore, the automatic tool must be fitted with mechanisms ensuring that certain technical actions can be performed only if it has been authorized by the proper judicial entities. The design of the automated system presented in this paper meets these conditions and is an effective tool in detecting pedophilic content.

6. References

[1] Chad M.S. Steel. "Child pornography in peer-to-peer networks". Child Abuse & Neglect, Volume 33, Issue 8, August 2009, Pages 560-568, ISSN 0145-2134, DOI: 10.1016/j.chiabu.2008.12.011.

[2] European Commission. "EU-level co-operation crucial for national police". forceshttp://ec.europa.eu/justice_home/fsj/police/fsj_police_intro_en.htm

[3] Villagrasa Alcalde (coord). "Nuevas Tecnologías de la Información y Derechos humanos". CEDECS, Barcelona, 2003, p. 87.

[4] Vid. Fernández Teruelo. "Los delitos cometidos a través de internet". Cibercrimen. CCC, 2007.

[5] Nabarro Nathanson. "The Laws of the Internet". Butterworths, UK, 1997, pag. 247.

[6] Vid. Díaz Barrado. "La protección del niño en el ámbito europeo". Ministerio de Asuntos Sociales, 2001, p.187.

[7] Further información in Cavanillas Múgica. "Responsabilidad de los proveedores de información en internet". COMARES, 2007.

[8] Carracedo Gallardo, Justo and Pérez Belleboni, Emilia. "Use of the New Smart Identify Card to Reinforce Electronic Voting Guarantees". The 4th International Conference for Internet Technology ans Secured Transaction. Published by Infonomics Society, UK, November 9-12, 2009, London, pp.439-444. UK ISBN 978-0-9564263-1-4.

[9] European Commission services for the professional community of eGovernment, eInclusion and eHealth practitioners. "Government Factsheet - Spain - Legal framework". http://www.epractice.eu/en/document/288370