SOCIAL MEDIA AND PERSONAL DATA PROTECTION

Radi Romansky

Technical University of Sofia
e-mail: rrom@tu-sofia.bg
Bulgaria

Abstract: The Information Society has created different possibilities for remote access to distributed information resources and communications between users (virtual environments, cloud services, social media, etc.). All these aspects of the globalization make users create their own profile with personal data and publish personal information. Are this data protected in a reliable way? This is an important question that every user should ask oneself. The answer is related to the privacy and the principles of personal data protection. The main goal of this article is to discuss the challenges of social media for data protection as a component of privacy. In this reason a brief review of social media is made and a formal description of global communications by using discrete structures is proposed. Main principles of personal data protection are presented on the base of organizational scheme, life cycle and data protection policy in the frame of security policy and in particular related to the information and communication security policy.

Key words: Information Society, Social Media, Data Protection, Security Policy, Privacy

1. INTRODUCTION

The modernization and the improvement of the Information Society (IS) determine new requirements to the contemporary information and communication technologies (ICT) for solution of different problems with globalization [1], remote access to information resources and cloud computing [2], distributed information servicing and virtual environments [3] and determination of adequate information security policy in enterprises [4]. Social networks and social media should be included in this group too because the contemporary ICT permit extension of social relationships and confirmation of the field “social computing” connected with building of networks of web sites (MySpace, Facebook, Twitter, XING, LinkedIn, YouTube, Pinterest, Foursquare, Newshub, e-Britanica, etc.) [5]. Users are not only passive participants in social media and consumers of information, but they could
realize different forms of direct communications, including companies and other participants.

The new ICT and distributed environments make users create their own profile with personal data and publish personal information, available to other users via the global network. This is an opportunity to extend the social contacts but it could cause undesirable effects on privacy of the registered user [6]. In this respect the protection of personal information is an important problem at the distributed services and network communications [7], including cloud services [8]. Therefore, special technological and organization measures for personal data protection (PDP) should be applied. These measures must protect every user’s profile with personal data against illegal access, dissemination and using of information for other goals different from the defined. A new point of view about the rules for the authorization and authentication in social environments must be formulated.

The need of PDP is determined by the fact that the privacy is an important human right that combines a complex of separate individual rights – correct and adequate processing of personal data, different form of personal communication (by post, by Internet, etc.), secure maintain of personal profiles in the social forums and groups, etc. The traditional definition of the term “privacy” is “the right to be alone” and this sense should be preserved in all social contacts via global network. This could be realized on the base of strong security policy defined for each special purpose of personal communications and support of profiles of individuals, because the using of contemporary ICT puts new requirements to the PDP policy and changes the understanding of privacy in the global society. This policy should increase the effectiveness of the means and tools for PDP in the new network society and in particular in social media and networks for keeping the privacy [9, 10, 11]. The main problem is that the existing procedures for PDP are not adequate to the real communications in the global society and they must be actualized. Some challenges of network communications including cloud services and distributed information service are discussed in [12, 13].

Privacy rights are connected to the personal information that could be gathered in the process of using social media. It is possible to collect and capture any personal information without user knowing and, moreover the personal data could be disseminated legally or illegal to any third person. More companies and institutions prefer to use social media to promote their services and products and the employees could post their personal data. A survey from different universities shows that the users are inclined to misrepresenting themselves online [14]. For that reason, the employers must have a strong policy for using social networks and sites by employees to protect personal data and must draw a line between personal and professional life. In other hand, the employers could use the social network and social media profiles to select promising employees and this issue raises different ethical questions.
2. A BRIEF REVIEW OF SOCIAL MEDIA (RELATED WORK)

The term ‘social media’ describes a complex of different web-based and mobile technologies that permits to transform the communication in an interactive conversation and it is used to share pictures, audio and video information, experience, etc. The social media ensure access of people to network resources for creating, editing and complementing content, but this content will be accessed if it respects several standards. Here are some popular social media: the social networks (Google+, Facebook, LinkedIn, Pinterest, Friendster, MySpace, Black Planet, etc), blogs (Twitter, etc.), web sites for sharing video contents (YouTube, VBOX7, Flickr, etc.), Internet forums, wiki-applications, virtual social sites (Second Life) and virtual sites for games (World of Warcraft), etc [15]. The social media have changed the understanding of communication via Internet and have created new dialog methods. The expansion of social media has created new opportunities for marketing specialists for direct communication with millions potential consumers. Different services (e-mail, voice over IP, crowd-sourcing, music sharing, banking, etc.) could be integrated in the social media by using aggregation platforms.

The traditional definition defines the social media as a group of Internet applications based on Web 2.0 that permits developing and exchanging content generated by user. Software engineers and users use the global network as a platform for sharing published content and new type of social interactions between organizations, communities, business companies and individuals. Contemporary technologies make the process for developing and publishing content very simple – each person could use any application at server and it is not necessary to know special program languages. In other hand, some IT specialists affirm that Web 2.0 is not effective and it will be transformed to next version Web 3.0 (a modern platform for media developing). This platform will be oriented to knowledge accumulation, semantic structures and ontology.

Seven functional blocks of the framework of social media are determined in [16]: identity (about the extent of disclosing information in a social media settings); conversations (about the extent of communications between users in a social media); sharing (about the extent of exchanging, distribution and receiving content by users); presence (about the extent to which users can know if other users are accessible); relationships (it represents the extent to which users can be related to other users); reputation (it represents the extent to which users can identify the standing of others, including themselves, in a social media setting); groups (it represents the extent to which users can form communities and sub communities). These functionalities permit each user to regulate your identity, relationships and reputation on the base of basic elements.

The interest of users and their activity to access and use different social media is investigated by Pew Research Center in the survey for year 2013 [17] – Fig. 1.
The results determine Facebook as the most used social media and the sum of all statistics shows that some adults use multiple social networks. The authors affirm that the margin of error is ±2.9%. Other interesting conclusion of the research is that Facebook is popular for different users (“mix of demographic groups”), Pinterest is preferred by female users, LinkedIn is especially popular among college graduates and internet users in higher income households, Twitter and Instagram are accessed by younger adults, urban dwellers, and non-whites.

The frequency of social media using obtained by the research in [17] is shown on Fig. 2. The assessments present the part of users (in %) who visit the investigated social media in three categories – daily using, weekly using and less often using. The leader is Facebook with 63% of users that access the social media site on a daily basis. The second place is occupied by Instagram (57% for daily using). LinkedIn and Pinterest are visited less than once per week.

“Growing Social Media” presents some statistical data about social media using. For example, Facebook is the biggest social network with 665 millions daily
users and 1.1 milliard monthly users. Google+ takes second place with 359 million active monthly users and the number of users of Google+ increases with 33% from June 2012 to March 2013. The growth of Twitter is 44% for the same period and it has about 288 million active monthly users (21% of the World population). This is the fastest growing network.

Another investigation carried out in the middle of year 2012 is presented in [18] and consists of some different statistics about social media. An interesting statistic is about the geographic distribution of the social media access (Fig. 3). It shows that most of the social media users are located in Asia Pacific.

![Fig. 3. Using social media in different zones of the World in 2012 (source [18])](image)

The social networks are very popular and give useful opportunities for contacts and exchange of information between different users by the resources of global network and web environment. This is valid no only for the individuals but also for business organizations, managers, traders, etc. For example, the business and commercial relations (including the e-commerce models B2B and B2C) determine a specific direction of the social media using. It is a fact that more dealers search their potential clients by using social networks. The most widespread social networks as YouTube, Google+, Facebook, LinkedIn and Twitter are preferred by online traders working on B2C and B2B models. Different statistical data obtained by investigations of social media show that the retail dealers prefer Facebook for their contacts but the most part of specialized tradesmen in USA (about 90%) are oriented use Pinterest. Another investigation shows that near 50% of the technological companies find their clients on Twitter.

In other hand, the individuals upload a profile in different social network sites and this information could be used by traders, managers, employers, etc. to select potential clients or job candidates. In this respect the social communication and profiles created and stored in the social media could create a problem for privacy of individuals. The popularity of social media makes easier the access to personal data. For example, Facebook is one of the most popular social networks preferred
by traditional college-age students. The young people communicate with friends, family, colleagues and upload different information (personal information, photos, video, etc.). In regard to this the number of employers that use this social network to assess job candidates increases [19]. This fact raises the ethical side of the relations because the employers use the global network that is a public forum as an instrument to decide their private problem without the individuals knowing.

Confidentiality is an important problem for social networking and adequate organizational and technical measures should be applied. The European Commission informs that three-quarters of Europeans think that the disclosure of personal data is an increasing problem of global IS [20] and 72% of Internet users consider that too much personal data are collected for online registration.

3. FORMALIZATION OF GLOBAL COMMUNICATIONS

Communications in the global environment (particularly in the web space) could be described by using discrete structure of elements (nodes) \( V = \{V_1, \ldots, V_n\} \), \( V \neq \emptyset \) and relations between them \( R_{ij} : V_i \to V_j \). Each node \( V_i \) presents a physical participant in global communication and it could be regarded as an independent distributed unit with own internal functionality. An abstract model built on the base of this concept is presented in Fig. 4.

The participants in the communications form two groups:

(a) individuals determined as a set of users \( U = \{U_1, U_2, \ldots, U_N\} \), \( U \neq \emptyset \), that could initialize the remote access via Internet;

(b) information environments \( IE = \{IE_1, \ldots, IE_M\} \), \( IE \neq \emptyset \), with specific technological components for determining the specific structure and purpose of the space, including a module for preliminary registration and personal data collecting in own database (DBforPD).

Organization of all processes for global communication is made by the resources of ‘communication medium’ that could be described as transmitters \( T = \{T_1, \ldots, T_K\} \), \( T \neq \emptyset \). Each transmitter consists of hardware and software tools for
distribution of requests for access to the information objects \( \text{req}: \{U\} \rightarrow \{IE\} \) and returning information content \( \text{inf}: \{IE\} \rightarrow \{U\} \) to the clients.

The formalization permits to describe the elements of global communications by ordered triplet \((U, \text{IE}, T)\) with two types relations \( \text{req}: \{U\} \rightarrow \{IE\} \) and \( \text{inf}: \{IE\} \rightarrow \{U\} \) for \( \forall U_i \in U; \forall IE_j \in \text{IE} \).

Let it us define a distance \( d_{ij} (i \neq j; i,j \in \{1 \cdots n\}) \) between each couple of nodes \((V_i, V_j)\). This permits to construct a matrix of distances \( DM \) with dimension \( n \) and elements \( d_{ii} = 0 \) \((i=1, \ldots, n)\) and to determine the minimum length of paths between nodes in the structure.

Let us present two binary parameters \( u_{ik} \in \{0,1\} \) (user \( U_i \in U \) location in node \( V_k \in V \)) and \( r_{jk} \in \{0,1\} \) (environment \( IE_j \in \text{IE} \) location in node \( V_k \in V \)):

\[
u_{ik} = \begin{cases} 1, & \text{if } U_i \in U \text{ is located in the node } V_k \in V \\ 0, & \text{otherwise} \end{cases} \quad (1)
\]

\[
r_{jk} = \begin{cases} 1, & \text{if } IE_j \in \text{IE} \text{ is located in the node } V_k \in V \\ 0, & \text{otherwise} \end{cases} \quad (2)
\]

In the general case, it is possible that two participants from different type (user and information environment) are physically allocated together in common node \( V_k \) and this could be described by the expressions (Fig.5):

(Type 1) \( \exists V_k \in V \Rightarrow (V_k \in U) \& (V_k \in \text{IE}) \) \( (3) \)

(Type 2) \( \exists V_k \in V \Rightarrow [(V_k \in U) \& (V_k \notin \text{IE})] \lor [(V_k \notin U) \& (V_k \in \text{IE})] \) \( (4) \)

This assumption determines power (number of elements) \( n \leq N+M \) of the set \( V \). Two binary matrixes with components \( u_{ik} (N \times n) \) and \( r_{jk} (M \times n) \) could be constructed for investigation physical allocation of participants in global communications. The logical processing based on

\[
\text{IF} \left\{ \sum_{i=1}^{N} UL [i,k] \geq 1 \right\} \& \left\{ \sum_{j=1}^{M} RL [j,k] \geq 1 \right\} \text{ THEN } v_k = 1 \text{ ELSE } v_k = 0 \quad (5)
\]

![Fig. 5. Two possible types of communication](image-url)
(for \( k=1 \div n \)) will define a new vector for determining common nodes for allocation of participants with different type. This formalization permits to make a deterministic describing of communications by couples of two elements (the first from \( U \), the second from \( IE \)) and to carry out an investigation of processes in global environment.

### 4. MAIN PRINCIPLES OF PERSONAL DATA PROCESSING

The term “personal data” determines the information that permits to identify a person directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity. Any operation or set of operations with personal data (using automatic or not-automatic means) is called “processing of personal data”. The main principles of personal data processing require strong rules for personal data protection (PDP).

The organization of personal data processing is summarized in Fig. 6. The participants in this process are “data subject” (the owner of personal data), “data controller” (determines the purpose and the means of processing and it is responsible for all procedures with personal data), “data processor” (real processing of personal data on the base of agreement with the data controller), “receiver of personal data” (the giving of personal data could be on the base of lawful reason only).

![Fig. 6. Organization of personal data processing](image)

A graphical interpretation of the life cycle of personal data processing is proposed in Fig. 7. This model of life cycle describes the traditional processing of personal data by a sequence of phases beginning from giving of personal data by individual and finishing with personal data destroying (by the data controller) after the goal realization. The purpose of the phases is listed below:

- The collection of personal data must be made based on a legitimate reason only and with the consent of the individual.
The preserving of collected data should be realized in the registers based on preliminary defined goal and criteria.

The using must be made by legitimate persons on the base of principles of information security: authentication (by using username, password, digital certificate, personal identification number, and biometric means); authorization (on the base of developed digital right management system); accountability (personalisation of the access to the data structures and registration of users’ activities).

- Actualization – the personal data must be correct, full and actual;
- The transfer to other country and the giving to other person must be realized on the base of strong rules only;
- Archiving could be made if it is required by law but for a limited period only;
- Destroying of personal data must be made after realization of the goal.

![Fig. 7. Life cycle of personal data processing](image)

Some different models of PDP exist in the World [13] – model of centralized legislation, model of joint regulation, model of sector legislation, model of self regulation and model of personal (individual) protection. The last model shows that most users of information services via global network have no confidence in the applied policy for information security and PDP at distributed information servicing. This requires the PDP policy to be coordinated with the measures of ICT Security Policy in the frame of the general Security policy (Fig. 8).

The first standard for Security Policy titled “Department of Defence Trusted Computer System Evaluation Criteria (TCSEC)” is accepted at 1985 in USA. TCSEC describes the security policy as a collection of rules, standards, procedures and practical instructions for regulation of the management, protection and dissemination of the information. This document gives rules for a control of access to the information resources. In this reason, the ICT Security Policy is an important component of the security policy and defines some important levels that must be discussed at the process of Data Protection Policy realization.
The users of social media give their personal data to make registration and upload personal information about your private life. This information is accessible and could be used by other persons (including dissemination to third party) without the consent of the owner. This requires a general harmonization of Data Protection Policy with the principles of the ICT Security Policy on all levels and to build useful Personal Data Security System (PDSS) in the frame of Personal Data Protection Policy (Fig.8).

PDSS is a collection of technical and organizational means and tools for realization of the protection of the personal data structures by data controller. All procedures for personal data processing that use ICT instruments should be analysed in details during the determination of Data Protection Policy for social networks and protecting the personal data.

5. PRIVACY IN SOCIAL MEDIA

PDP is an important component of the privacy of individuals and this is determined in different international agreements and documents. The growth of possibilities for automated information processing, the using of remote access to
different information resources and the extension of network communications (including social networking) change the technological aspects of personal data processing. The new forms of communication in the global IS as information sharing, social media, cloud services, etc., create barriers for enforcing basic directives for PDP. It is needed to apply stronger requirements to data protection policy and information security for all Internet communications and using of social media. The initial step in this direction is the proposal of European Commission to change the old paradigm “to leave alone” with a new one “to be forgotten” [20] – this means the personal data of the individuals should be processed only for the short period and must be removed after finishing the legitimate reason for processing. The main responsibility of data controllers is to ensure reliable protection of collected and processed personal data in social media and providers must apply the principle “privacy by default” because only a quarter of social networks users feel in complete control of their personal data.

Privacy in social media concerns with protection of user’s information and securing the user’s rights. The media must try to prevent different incidents with user’s data as unauthorized access, viruses, illegal transfer to third party, etc. Some challenges of social media to PDP are summarized below.

- An important obligation of data controllers is to build a reliable PDSS for realization the principles of all structural levels (see Fig. 8) and to inform clearly, understandably and transparently each user for using of personal data. This requires as an initial step to identify the roles of “data controller”, “data processor” and “data subject” in the social media and to determine the responsibility for data protection procedures (rules, measures, privacy and rights of data subjects, etc.). According to the definitions of Directive 95/46/EC the data controller determines purposes and means of the processing of personal data. The problem in social media is that the functions of customer, vendor and provider and the relation between them could be defined for concrete case only. The service providers have no legal obligation to protect personal data if they are not defined as controllers or processors. The different opportunities to determine the provider’s status (and the possibility to change it) make very hard the solution of this problem. This characterization will permit to ignore the data protection obligations at the cases of personal data outsourced or transferred to a third party for processing. In this respect European Commission has proposed a new regulation that will extend the frame of the Directive 95/46/EC and will respond to the new challenges of Internet society [21].

- Another problem that could be determined in social networks is the data subject’s right to be informed. This is an integrated problem because the individuals have different rights during the personal data processing. As a first one there is a risk for user’s privacy during the registration (more personal data could be required for registration and identification) and using resources of social media. The level of privacy in social networking is very different – some social
networking sites collect limited personal data in the page known as a “profile” (names, birth date, address, phone number), but other sites require additional information about social life, gender, country, hobbies, relationships, etc. These pieces of data personalize the users in major level and the individuals must know the purpose of these data and reason for processing.

- An obligation of the controllers is to guarantee easy access to the own personal data of users. This will permit to realize the user’s rights to revise, access, block or delete their personal data in the profile (which is a fundamental right guaranteed by data protection laws). Another side of the problem is the access to the information in the profile – the controller must guarantee that each user could define restriction for the own profile accessing. This will prevent unauthorized access and incorrect dissemination of personal information. This action could be realized by making the profile private from the user by selection of these who can visit the page. Traditional manner for authentication at the access to the profile is by username and password and this will ignores any operation with stored information (adding, removing, modification, changing, including pictures, etc.). In this case the privacy in social media will be assured for the user.

- International data transfer could be determined as the next eventual problem of the privacy in social media. According the main principles of PDP personal data could be transferred to another country if their level of PDP is adequate. The data transfer between different service providers is typical procedure in the social networks because the nodes (servers, clients, storages, etc,) could be located anywhere in the world. If any personal information is uploaded to social media it must be protected according to the rules of PDSS and the individual (social media user) must be informed for all transfers of their data from one service provider to another in the frame of the country or outside.

- Data deletion. If any user wants to delete data in his/her profile he/she must be sure that these data will be really deleted. In some cases, data could be transferred to other service provider and a copy of data could be stored in different place(s). This will be a problem of privacy for the individual. Another case is when the information that was deleted or removed by the user is passed to third party before deletion. Data protection legislation gives strong rules for deletion of personal data in the traditional cases, but for the social media this is not clearly determined.

- Shared information – the social networking is realized on Internet and all information resources could be accessible from different points in the world. This provokes the traditional danger in the global network (loss of data, destroying the integrity, problems with accountability, hackers’ attacks, etc.). Each user uploads information that will be shared between a set of users of social network and it could be disseminated to different locations. In this case the data subject does not know what policy and measures are used for counteraction to eventual attacks.
• Technical and organizational measures for data protection – to implement appropriate measures for information security is an important obligation for the data controllers. These measures should be a counteraction to all forms of destruction or loss of personal data, to an unauthorized access (during the personal data processing or transmission via communication links), and to all illegal forms of processing. The service providers should guarantee an effective protection of data integrity and data availability. It is known that more data security measures will increase the cost of PDP procedures and this will be a reason for ignoring some of the measures.

6. CONCLUSION

Privacy in social media can be undermined by many factors determined by incorrect using of personal data. Fact is that the young people prefer to communicate by social networks and these forums are easy to access. However, the access to the users’ personal data should be restricted.

After summarizing the main challenges of social media for the privacy in the previous section, we should note the interesting point of view that is presented in [22]. The author reveals 12 myths for privacy in social networks concerning rights of individual users to own their data, treating the user’s privacy by social media companies and technologies, enforcement of law and paradigm “right to be forgotten”, anonymous using, etc. Yes, some of the examples and conclusions about information posted in the social media platforms by users could be accepted after discussion, but the problem of privacy (and the obligations of social media companies) is related to the profiles that the users must create at the registration. In this respect, an activity for modernization of data protection rules on European level is realized. An example is the document “Proposed Regulation” of the European Commission (January 2012) that proposes new rules to strengthen online data protection rights. The reason for these draft amendments is the fact “that rapid technological development and globalization have profoundly changed the world and brought new challenges to the protection of personal data…” [21]. This document discusses the paradigm “right to be forgotten” (Article 17) and the data subject rights to data portability (Article 18) – transfer between different electronic processing systems.

The European Parliament has determined (12 March 2014) architecture and fundamental principles of the data protection reform for improving user protection and security in cyberspace [23]. The conclusion is that the further development and exploitation of cyber space could not be realized without an adequate and strong protection of the rights of individual users [24]. Four pillars have been determined:

Pillar (1): “One continents one law” – a requirement about the regulation and sanctions in private and public sectors.
Pillar (2): “Strong regulation of European digital industry” – a requirement for the non-European companies, when offering services to European consumers, to apply the European rules and level of data protection.

Pillar (3): “The right to be forgotten / The right to be erased” – this is the right of an individual to remove own personal data from the system if she/he no longer want to use the online services or there is no legitimate reason for keeping it in this online system. This regulation will permit the individuals to control own online identify and to require the personal profile to be removed from the system (including social media platforms).

Pillar (4): A "One-stop-shop" for businesses and citizens – a regulation for the personal data processing by controller or processor established in more that one country of European Union.

The new principles of regulation must extend the PDP frame determined by the previously directives, and to propose adequate solutions for all problems of PDP in social environments.

REFERENCES


Information about the author:

Radi Romansky – Doctor of Science in Informatics and Computer Science; Full Professor on Computer Systems, and Networks at the Department of Electronics, Computer Systems and Technologies (College of Energy and Electronics, Technical University of Sofia); Head of the Department; Vice Chairman of the General Assembly and Member of the Academic Council of the Technical University of Sofia. Areas of scientific interests: information and computer technologies, computer architectures, data protection and information security, e-governance, etc.

Manuscript received on 10 September 2014