DOI: 10.5604/01.3001.0013.0905

ALGORITHM OF USER'S PERSONAL DATA PROTECTION AGAINST DATA LEAKS IN WINDOWS 10 OS

Olexander V. Zadereyko, Olena G. Trofymenko, Nataliia I. Loginova
National University "Odessa Law Academy", Department of Information technology, Odessa, Ukraine

Abstract. *In the European Union, in the first half of 2018, the General Data Protection Regulation came into force, which established the new rules for processing users' personal data for IT companies. The operating systems (OS) are the dominant software that is responsible for collecting and processing data in computer systems. The most common OS is the Windows OS family. The authors identified Windows 10 operating systems, that collect and accumulate user's personal data; developed and tested practically an algorithm, the application of which localizes and blocks the transfer of user's personal data to official servers of the Microsoft company.*

Keywords: Windows 10, personal data collection, personal data leakage, operating system, operating system telemetry, personal data protection

ALGORYTM OCHRONY DANYCH OSOBOWYCH PRZED WYCIEKAMI DANYCH W OS WINDOWS 10

Streszczenie. *W Unii Europejskiej w pierwszej połowie 2018 r. weszło w życie ogólne rozporządzenie o ochronie danych, które ustanowiło nowe zasady przetwarzania danych osobowych użytkowników dla firm informatycznych. Systemy operacyjne są dominującym oprogramowaniem odpowiedzialnym za zbieranie i przetwarzanie danych w systemach komputerowych. Najpopularniejsza obecnie jest rodzina systemów operacyjnych Windows. W artykule autorzy zidentyfikowali systemy operacyjne Windows 10, jako zbierające i gromadzące dane osobowe użytkowników; opracowali i przetestowali w praktyce algorytm, którego zastosowanie lokuje i blokuje transfer danych osobowych użytkownika na oficjalne serwery firmy Microsoft.*

Słowa kluczowe: Windows 10, gromadzenie danych osobowych, wyciek danych osobowych, system operacyjny, telemetria systemu operacyjnego, ochrona danych osobowych

Introduction

In May 2018, in the European Union the General Data Protection Regulation came into force, which established the new rules for processing users' personal data for IT companies in the European Union. A distinctive feature of the new regulation is the provision of tools for residents and citizens of the European Union to ensure the full control of their personal data, i.e. an individual acquires the right to demand termination of the processing and deletion of their personal data upon request [7].

Typically, the collection of user's personal data is performed by the operating system, which manages the user's workplace The operating system (OS) Windows 10 is among the most common operating systems. Taking that into account, the authors identified the system services and services responsible for collecting personal data. The IP addresses of Internet resources were also established, to which personal data collected by the OS is sent without notifying the user.

Most users of personal computers (PCs) running Windows 10 do not even know what a huge amount of data is collected and sent to the server of its official developer, Microsoft. This fact raises concerns about the leakage of important data from the user's point of view [13].

The working group on the protection of individuals with regard to processing of personal data, more commonly known as the working group of the Article 29, twice sent letters to Microsoft expressing displeasure with the company's policy regarding the collection of private data in Windows 10. The last letter, dated February 2017, states that the European Union considers that the changes made by the developers of Windows 10 to the system during this period are not enough to fully protect the users' personal data [2].

The purpose of the article is to analyze the user's personal data, which is collected in Windows 10, and the development of practical recommendations, the use of which will prevent their transfer to the official servers of Microsoft.

1. Collection and recording of personal data

In Windows OS, the protocoling of all user actions is performed in the logbook, in which records of all events in the OS are recorded. Such data, in particular, includes: prefetch data; recent documents; autocomplete; all entered commands; network passwords; list of deleted files; temporary files; memory dumps; ChkDsk memory fragments; OS log files; DNS cache; list of running software; Thumbs.db file. Conducted practical research performed by debugging tools and network monitors, allowed to establish what kind of information Windows 10 extracts from the user's PC and sends it to the Microsoft servers [1]:

1) typed text. It does not matter with the help of what software the test is typed. Keystrokes are intercepted, them they collected in packets and sent once every 30 minutes to the servers:
   - oca.telemetry.microsoft.com.nsatc.net,
   - pre.footprintpredict.com,
   - reports.wes.df.telemetry.microsoft.com.
2) geolocation data. The names of Wi-Fi networks in which the PC adapter is registered are transmitted every 30 minutes, which allows to track the physical relocation of the user.
3) recording from a microphone. The most active "collector" of data in the OS is Cortana – a virtual voice service with artificial intelligence elements for OS Windows Phone 8.1, Microsoft Band, Windows 10. The information is sent to the specified servers every 15 minutes (approximately 80 megabytes). Cortana software converts voice records to text, and then also sends wav-files to Microsoft servers, including records of conversations via SIP-telephony. The collected text information is stored inside the built-in notebook program [5]:
   - vortex.data.microsoft.com,
   - vortex-win.data.microsoft.com,
   - telecommand.telemetry.microsoft.com
   - telecommand.telemetry.microsoft.com.nsatc.net,
   - oca.telemetry.microsoft.com,
   - oca.telemetry.microsoft.com.nsatc.net,
   - sqm.telemetry.microsoft.com,
   - sqm.telemetry.microsoft.com.nsatc.net,
   - watson.telemetry.microsoft.com,
   - watson.telemetry.microsoft.com.nsatc.net,
   - redir.metaservices.microsoft.com,
   - choice.microsoft.com,
   - choice.microsoft.com.nsatc.net,
   - df.telemetry.microsoft.com,
   - reports.wes.df.telemetry.microsoft.com,
   - wes.df.telemetry.microsoft.com,
   - services.wes.df.telemetry.microsoft.com,
   - sqm.df.telemetry.microsoft.com,
   - telemetry.microsoft.com,
   - watson.ppe.telemetry.microsoft.com,

- telemetry.appex.bing.net,
- telemetry.urs.microsoft.com,
- telemetry.appex.bing.net:443,
- settings-sandbox.data.microsoft.com,
- vortex-sandbox.data.microsoft.com,
- survey.watson.microsoft.com,
- watson.live.com. watson.microsoft.com,
- statsfe2.ws.microsoft.com,
- corpext.msitadfs.glbdns2.microsoft.com,
- compatexchange.cloudapp.net,
- cs1.wpc.v0cdn.net,
- a-0001.a-msedge.net,
- statsfe2.update.microsoft.com.akadns.net,
- sls.update.microsoft.com.akadns.net,
- fe2.update.microsoft.com.akadns.net,
- diagnostics.support.microsoft.com,
- corp.sts.microsoft.com,
- statsfe1.ws.microsoft.com,
- pre.footprintpredict.com,
- i1.services.social.microsoft.com,
- i1.services.social.microsoft.com.nsatc.net,
- feedback.windows.com,
- feedback.microsoft-hohm.com,
- feedback.search.microsoft.com,
- ad.msn.com, preview.msn.com,
- ad.doubleclick.net,
- ads.msn.com, ads1.msads.net,
- ads1.msn.com,
- a.ads1.msn.com,
- a.ads2.msn.com, adnexus.net,
- adnxs.com, az361816.vo.msecnd.net,
- az512334.vo.msecnd.net.

4) telemetry. Traditionally, the term "telemetry data" refers to information about the state of a user's PC and its activity: what software is installed or running, the amount of used memory, software logs, and fragments of RAM. The telemetry service collects information about the geographical location, the IP address of the user's OS. It uses not only information entered directly in the web browser, but also information taken from the chats, which is intercepted from any installed application software (including when typing from the keyboard). The collected telemetry data is transmitted to the following servers:
   - telecommand.telemetry.microsoft.com,
   - telecommand.telemetry.microsoft.com.nsatc.net,
   - oca.telemetry.microsoft.com,
   - oca.telemetry.microsoft.com.nsatc.net,
   - sqm.telemetry.microsoft.com,
   - sqm.telemetry.microsoft.com.nsatc.net.
5) the fight against piracy. Windows 10 searches the names of the user files. If the name of the popular media content is printed in any OS application software, the OS will search for it on the PC hard disk and index the media files found. The created index file will contain a variety of information about files and documents stored on disk. Indexing occurs constantly while adding, deleting or changing files in folders. Indexing is used to find pirated media content on the user's PC and the creation of targeted advertising focused on their interests. The index file information is sent to the servers:
   - df.telemetry.microsoft.com,
   - cs1.wpc.v0cdn.net,
   - vortex-sandbox.data.microsoft.com,
   - pre.footprintpredict.com.
6) search history. If the user disables the search function on the Internet from the Start menu, performing such a search will still result in sending a request to the Bing server. As a result, the threshold.appcache file will be generated, which will also contain some Cortana information, including the device ID. The threshold.appcache file is saved when the OS reboots. The OS tracks the user's search queries, information about which is transmitted to the Bing search service. The collected information is stored in temporary files, which are transmitted every 30 minutes to the following servers:
   - oca.telemetry.microsoft.com.nsatc.net,
   - pre.footprintpredict.com,
   - reports.wes.df.telemetry.microsoft.com.
7) Taking into account that OS Windows 10 is a system that is used in many of the largest international companies, such functionality can potentially lead to leakage of important confidential data. All data transmitted to Microsoft's servers is encrypted, so it is not possible to perform a detailed analysis of information sent by the OS without the need for decryption algorithms.
8) SmartScreen filter, constantly checks the installed software for a list of phishing sites from which it is downloaded. This filter sends information about each application software that a user has downloaded from the Internet, tried to install or installed in the OS to the Microsoft server [4]. Data is sent when the user launches the installer. When Microsoft receives the data, it checks whether the software has the required certificate. If it is absent, a message is issued that the launch of this software can cause damage, after that the user is prompted to abandon the installation. This information can be used by law enforcement agencies to identify PC owners who use pirated copies of software. Microsoft can use this data itself – the company regularly conducts raids against pirates.
9) Query Formulation Via Task Continuum" technology is based on the use of an intermediate component – Mediator, which monitors user actions by transferring the collected data to the rest of the application software and user devices. In particular, Mediator monitors text input, search for media content on the network, work with images stored on a PC hard disk, collects data from such application software, for example: MS Office, Skype, Edge, etc. [6].

Thus, OS Windows 10, playing the role of an intermediate component between the user and the servers of the Microsoft company, monitors all the text and voice data entered in any software, collects them and sends a data dump in real time.

2. Methods of protection of user's personal data

The following measures can be used as measures to protect the OS from leaks of personal data: freezing of the system; portable software (PPO); anti-virus software; firewalls; OS encryption using cryptographic software. As a result of the application of the combination of the above measures, the probability of personal data leakage in the OS will be minimized.

2.1. Organization of the isolated OS environment

Using software to organize an isolated program environment in the OS is performed by "freezing" the OS. This approach allows you to constantly maintain the system in a "clean" state and not accumulate personal user data [9]. It should be noted that there are two main approaches to the organization of an isolated "frozen" software environment in the OS:

1) During one session the user of the OS uses both software running in an isolated environment (on a hard disk or its partition), and software running in normal mode.
2) The user creates an OS image after which he begins to work with it. All actions performed by the user are saved only until the OS is restarted. After the OS reboots, it returns to the original "clean" state.

In most cases, it is preferable to use the second approach in practice. It allows you to maintain the performance of the OS at the original level, since the OS does not accumulate personal user data that slows down its work.

2.2. Application of portable software

Portable software (PS) is a software that does not require installation in the OS and is run by the user from an external drive with USB interface (or another). The main difference of portable software from the standard software is the independence of the PS from the standard version and configuration of the OS, as well as the preservation of personal data and user settings in the software after the session in the OS [12]. Signs on which the applied software can be carried to the PS class are:

- the independence of the application software from the OS version installed on the PC, which is provided by the shell program or the virtual environment,
- exclusion of the OS reboot during the installation of the application software,
- the ability to run and save all the software application settings on an external data storage device and their immutability in subsequent work sessions, regardless of the OS version,
- the reversibility of the computing environment of the OS after its configuration, initiated by the launch of the application software, so after its completion, all temporary directories and registry keys which were added to the OS are deleted. Some shells ensure that this condition is met even when the device is removed from the USB port while the application software is running.

2.3. Application of firewalls

The problem of localizing leaks in the OS when connected to any external network is effectively solved by firewalls that protect the user's personal data from unauthorized access using vulnerabilities in the OSI network model protocols or in the installed software.

A firewall is a set of software or hardware tools that allow filtering and monitoring of packets passing through it in accordance with the parameters set in advance [11]. The firewall monitors the information that enters the OS from the external network (Internet) and back, blocks attempts to introduce malicious codes into the OS and does not allow the OS and application software to send any data to the external network unconditionally. Using a firewall in the OS is one of the key requirements for ensuring user privacy.

The use of firewalls can significantly reduce the amount of user network traffic by filtering the access of the OS and application software to the external network. To protect the user's personal data during the installation of the Windows OS family, it is recommended to block access to the Microsoft servers listed in the work using the firewall tools [2].

2.4. The use of cryptographic software to encrypt the OS

Now, the mobility of users is rapidly growing, and, consequently, the number of possible network connections is increasing. It greatly increases the likelihood of leakage of personal user data. One of the effective ways to counteract this is to use encryption of the system partition of the OS and working partitions of the PC hard disk (if necessary).

In practice, the cross-platform cryptographic software TrueCrypt [3] or its open source analogue Veracrypt is used to solve this problem. Its distinctive feature is the ability to encrypt data on the fly (On-the-fly encryption). At the same time, the data is encrypted in full volume, including file headers, their contents, metadata, etc. This approach minimizes leakage of confidential user information.

2.5. Deactivation of personal data collection and transmission services

The use of application software allows you to deactivate the process of collecting and transmitting personal user data [8]. In general, the features of the above listed software are reduced to the following functions:

- stopping of services that enable user shadowing and preventing them from restarting,
- disabling the service of recording characters entered on the keyboard,
- disabling the Cortana module,
- disabling or blocking the tracking components in the Media Center, Customer Experience Improvement Program, Power Efficiency Diagnostics, Family Safety Monitors, Office ClickToRun Service Monitor, Application Experi-ence, Office Telemetry, Disk Diagnostic, Media Center, Windows Search, and others,
- disabling services that record the history of visits to Internet resources and installed software,
- disabling services for the collection and transmission of information about the location of the user.
- disabling the ability to save and transmit information about access points and their passwords,
- blocking the service of sending an encryption key to the telemetry server encrypted by the OS,
- disabling the Update service,
- cancel spyware updates in the OS,
- blocking of a network screen embedded in the OS to send data to telemetry servers,
- creating an OS recovery point before making changes to the OS registry.

To protect the user's personal data from leaks in Windows 10, the authors suggested and used the following algorithm (Fig. 1):

1) Installing the OS on the user's PC in the mode "without connecting to the Internet", which makes it possible to exclude its identification by Microsoft Corporation. Disable built-in data collection tools of Windows 10.
2) Processing of the OS by specialized software that disables the built-in services for collecting and transmitting personal user data [8].
3) Creation of an isolated software environment in the OS (installation of portable software).
4) Filtering outgoing OS network traffic using a firewall. The use of a firewall prevents the transfer of any data from the OS to the external environment by blocking connections to Microsoft servers. Firewall configuration is performed in the blocking mode of all outgoing connections with servers.
5) Encryption of the system partition of the OS is performed using specialized cryptographic software Truecrypt or Veracrypt in the mode of pre-load authentication. This measure implements encryption/decryption of data of the system partition of the OS in the process of writing/reading.

3. Conclusion

Practical application of the algorithm proposed by the authors will allow users:

- to hide the list of actions that he had performed in any case, including if they receive to have access to the physical media OS,
- to avoid the accumulation of service and personal data in the OS,
- to perform deep cryptographic encryption of the OS system partition using public and private keys,
- to block the transfer of personal data to Microsoft servers,
- to block the transfer of data from third-party software installed in the OS to the servers of its developers,
- to prevent external network attacks on the OS, carried out in order to obtain unauthorized access to confidential data,
- to ensure the highest possible level of privacy when working on a computer running Windows 10 while maintaining its integrity and ease of use.

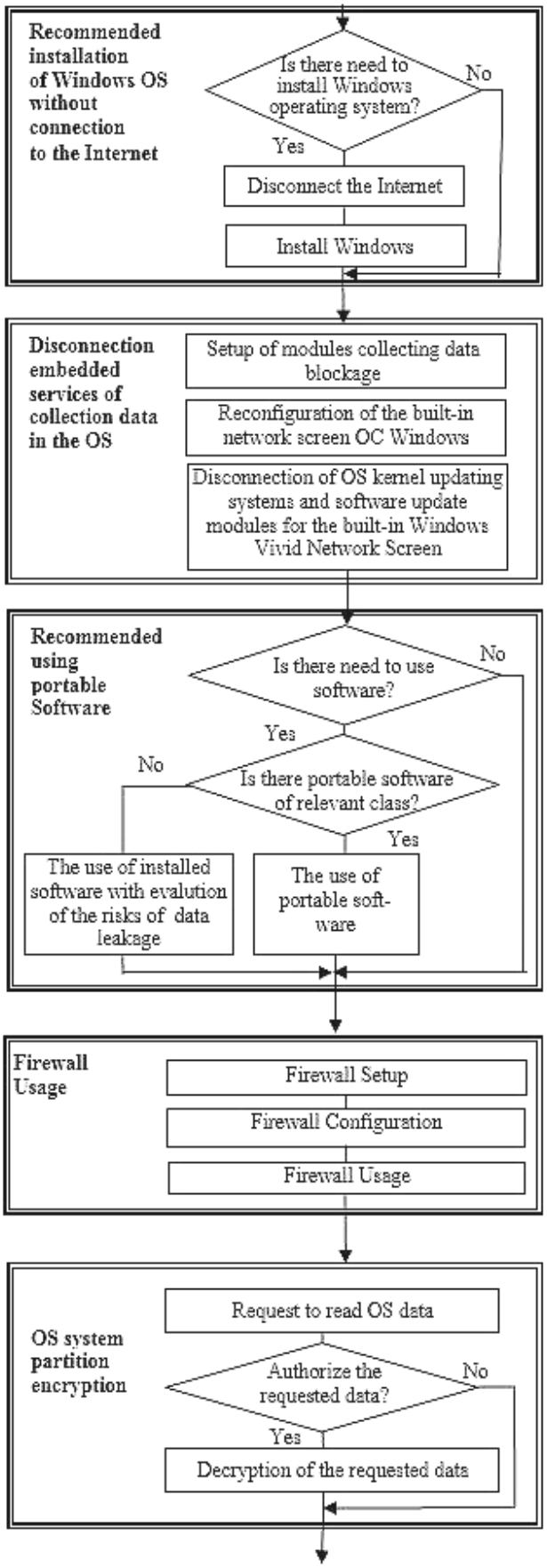

Fig. 1. The algorithm for protecting users' personal data from leakage in Windows 10

References

[1] Analýza Windows 10: Ve svém principu jde o pouhý terminál na sběr informací o uživateli, jeho prstech, očích a hlasu! URL: https://aeronet.cz/news/analyza-windows-10-ve-svem-principu-jde-o-pouhy-terminal-na-sber-informaci-o-uzivateli-jeho-prstech-ocich-a-hlasu. (Available: 29.09.2018).

[2] ARTICLE 29 Data Protection Working Party. https. URL: http://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/20160112_microsoft_privacy_policy_letterpdf_0.pdf. (Available: 29.09.2018).

[3] Loginova N., Trofimenko E., Zadereyko O., Chanyshev R.: Program-Technical Aspects of Encryption Protection of Users' Data. XIIIth 2016 International Conference Modern Problems of Radio Engineering, Telecommunications, and Computer Science. Lviv 2016, 443–445.

[4] Microsoft denies Windows 8 app spying via SmartScreen. URL: https://www.theregister.co.uk/2012/08/25/windows8_smartscreen_spying/. (Available: 29.09.2018).

[5] Microsoft Monday: Leaked Windows 10 Changes, Cortana Suggested Reminders, Visual Studio 2017 Details. URL: https://www.forbes.com/sites/amitchowdhry/2017/02/13/microsoft-monday-leaked-windows-10-changes-cortana-suggested-reminders-visual-studio-2017-details/#156efd7d2dda. (Available: 29.09.2018).

[6] Microsoft Patents Big Brother AI To Monitor Everything You Do In Windows And Feed It To Bing For Search Results. URL: https://hothardware.com/news/microsoft-patents-big-brother-ai-to-monitor-everything-you-do-in-windows. (Available: 29.09.2018).

[7] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679#ntr4-L_2016119EN.01000101-E0004. (Available: 29.09.2018).

[8] Tools to tweak Windows 10 Privacy settings and fix privacy issues. URL: https://www.thewindowsclub.com/tools-tweak-privacy-settings-windows-10. (Available: 29.09.2018).

[9] Top 10 Reasons for Computer Freezing. URL: https://www.stellarinfo.com/blog/top-10-reasons-computer-freezing/. (Available: 29.09.2018).

[10] Veracrypt. URL: https://www.veracrypt.fr/. (Available: 29.09.2018).

[11] What Is a Firewall? URL: https://www.cisco.com/c/en/us/products/security/firewalls/what-is-a-firewall.html. (Available: 29.09.2018).

[12] What is a portable app? URL: https://portableapps.com/about/what_is_a_portable_app. (Available: 29.09.2018).

[13] Windows 10: Microsoft under attack over privacy. URL: https://www.theguardian.com/technology/2015/jul/31/windows-10-microsoft-faces-criticism-over-privacy-default-settings. (Available: 29.09.2018).

Ph.D. Olexander Zadreyko
e-mail: zadereyko@onua.edu.ua

Associate professor of the Department of Information technology of the National University "Odessa Law Academy". His researches interests include security of information systems and microelectronics. Author of nearly 50 scientific and methodical works on protection of information and office technology.

ORCID ID: 0000-0003-0497-9861

Ph.D. Olena Trofymenko
e-mail: trofymenko@onat.edu.ua

Associate professor of the Department of Information technology of the O.S. Popov Odessa National Academy of Telecommunications. Her researches interests include modern information technologies and developments of computer sciences. Author of nearly 120 scientific and educational works, co-author of patents and textbooks on programming, database, web-design, office technology.

ORCID ID: 0000-0001-7626-0886

Ph.D. Nataliia Loginova
e-mail: loginova@onua.edu.ua

Associate professor of the Department of Information technology of the National University "Odessa Law Academy". Her researches interests include modern information technologies and security of information systems. Author of nearly 80 scientific and educational works and textbooks on protection of information, database and office technology.

ORCID ID: 0000-0002-9475-6188

otrzymano/received: 02.12.2018 *przyjęto do druku/accepted: 28.02.2019*