УДК 004

Обзор технологии OSINT для анализа социальных сетей

Ксения Антоновна Щербакова[1], Диляра Ильшатовна Губина[2]

[1, 2] Уральский федеральный университет

имени первого Президента России Б. Н. Ельцина, Екатеринбург, Россия

[1] ksehcherbakova@gmail.com

[2] diliara.kurmanova@urfu.ru

Аннотация. В данной статье речь пойдет о методах поиска информации с помощью разведданных из открытых источников. Open-source intelligence (OSINT) - это разведданные, собранные из общедоступных источников для удовлетворения конкретных требований разведки. Анализ полученных OSINT данных постов в социальных сетях от целевых пользователей показывает, что можно извлечь любую информацию о человеке, данные которого есть в открытых источниках, таких как Facebook, Instagram, Twitter и другие источники.

Ключевые слова: социальная сеть, пользовательский интерфейс, база данных, оценка угроз, сбор информации, OSINT

Application of Open Source Intelligence (OSINT) technology for social media analysis

Ksenia A. Shcherbakova[1], Dinara I. Gubina[2]

[1, 2] Ural Federal University named after the First President of Russia B. N. Yeltsin, Ekaterinburg, Russia

[1] ksehcherbakova@gmail.com

[2] diliara.kurmanova@urfu.ru

Annotation. This article will discuss ways to search information using special information gathering method. Open-source intelligence (OSINT) is intelligence gathered from publicly available sources to meet specific intelligence requirements. Analysis of OSINT data from social media posts from target users shows that it is possible to extract any information about a person whose data is available in public sources, such as Facebook, Instagram, Twitter and other sources.

Key words: social media, user interface, database, treat assessment, information gathering, OSINT

Introduction

Earlier social media was used only in a quality of service but in today's modern era it is more than just that. In itself, a social network refers to a way in which people interact and share information via virtual communities. Now it is just like our pseudo-online biodata whenever we meet any new person, we visit their social media profiles and get an idea about almost everything like hobbies, interests, likes, dislikes, etc. Nowadays there are various tools which do some data analysis using Artificial Intelligence or Machine Learning that create one's visual profile. This information is quite essential and this tools automatically gather this information since it is a tedious and time-consuming process if done manually. Hence, there is a need for a system which focuses on automated data gathering from various social media platforms and makes it easy to access all in one place.

Open Source Intelligence has been widely acknowledged as a critical source of valuable and cost efficient intelligence that is derived from publicly available sources. With the rise of prominent social media platforms such as Facebook and Twitter that record and expose a multitude of different datasets. Some major obstacles that OSINT analysts often face are privacy and platform restrictions that serve both to protect the privacy of individuals and to protect the economic livelihood of the social media platform.

1 Architecture

This method is composed of two elements in charge of collecting data coming from OSINT sources and the infrastructure to centralize all collected and generated data. Its main objective is to collect, clean, and aggregate data to feed the operational module with composed indicators of compromise and other threat related data for further data processing and analysis. The remainder of this section details the components of this module.

1.1 Implementation

The flow graph in Figure 1. shows the process of data gathering and displaying the results. It is explained in brief below: (Fig. 1)

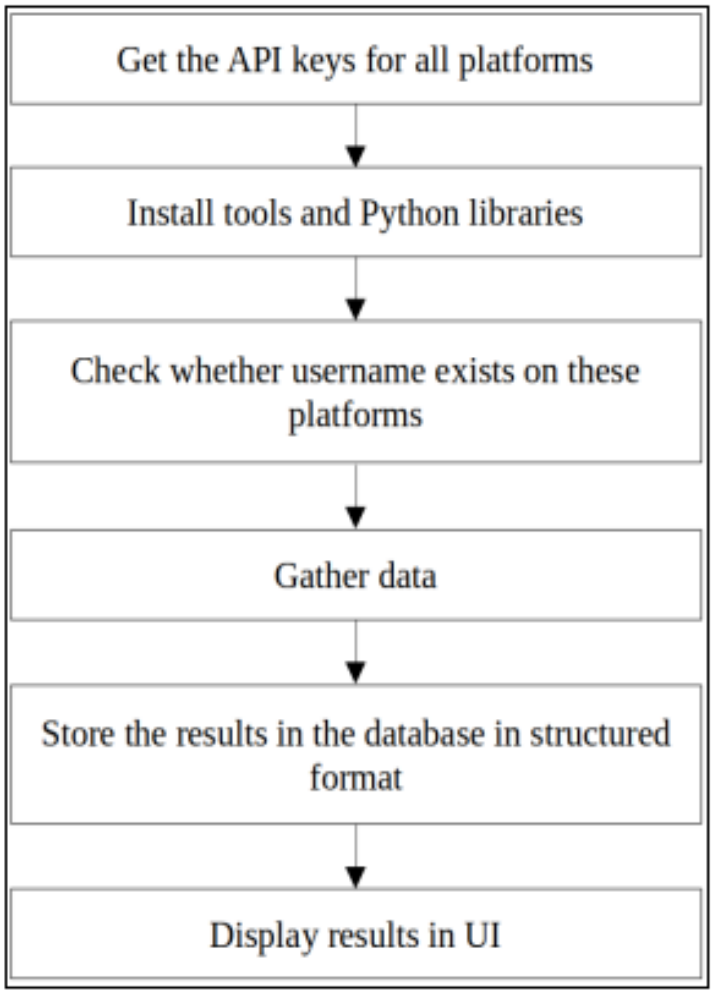

Figure 1. Flowgraph of the OSINT System

The main parts are the following:

a. *Getting the API keys*

This stage involves connection to the social media platforms, where the users need to get the API keys in order to gather information from their websites. This step

is essential since websites do not allow direct access to user data for security purposes. This can be done by creating a developer account on these platforms and authenticating using the API keys.

b. *Installation of tools and Python libraries*

At this point, the widely used open-source tool Sherlock, which is also written in Python, is used to verify the existence of the username. After installing these tools and libraries, you can start collecting data.

c. *Check whether a username exists on these platforms*

In this section there is a request to enter a username, this user name will be passed as an argument to the Sherlock tool. Subsequently, the tool will run scripts to verify the existence of the username on each of the platforms and return the results.

d. *Gathering data*

The next step, from the results returned by Sherlock, using API keys to access all the public user data, is to collect data from the platforms where the username exists. This data will then be structured into JSON format for ease of use.

e. *Storing the results in database*

The collected results will then be cached in the database to reduce the overhead of retrieving the same data again. This will help to reduce the required processing time and increase the efficiency of our software.

f. *Displaying the results in UI*

After the results are stored in the database, they are fetched and displayed in the software UI in proper format. The results can then be exported to various file formats such as PDF, Docx, CSV, etc.

2 Integration process between selected OSINT techniques and regulatory standard

The complexity of protecting the information sphere is that every day technology makes a big leap in development, and, consequently, the number of methods of attack and defense grows. Therefore, there is a problem in regulating and providing protection, because not all methods will be allowed in terms of law. To that end, there are the benchmark ISO 27001, which will be used as it is a reputable, internationally

recognized standard, the “ISO/IEC 27001:2013 standard specifically provides requirements for an information security management system”, once risks are identified in the initial assessment, controls are selected and implemented to mitigate them noting that there are 114 controls in place under the ISO 27002 umbrella. Moreover, complying with ISO 27001 requirements helps the organization in meeting other standards and regulations in the future. As mentioned before, ISO 27001 (formally known as ISO/IEC 27001:2013) is an international standard for an information security management system (ISMS), which can be defined as “the process of identifying, evaluating, and treating risks around the organization’s valuable information” 4. The standard contains a set of policies and procedures including technical, physical, and legal controls involved in an organization’s risk management program, Table 1 shows that standard consists of 114 controls in 14 domains. The proposed OSINT integration process will focus on the following domains

Table 1

ISO 27001 Main Domains

Annex A#	Domain Name	# of controls
A.5	Information security policies	2 controls
A.6	Organization of information security	7 controls
A.7	Human resource security	6 controls
A.8	Asset management	10 controls
A.9	Access control	14 controls
A.10	Cryptography	2 controls
A.11	Physical and environmental security	15 controls
A.12	Operations security	14 controls
A.13	Communications security	7 controls
A.14	System acquisition, development and maintenance	13 controls
A.15	Supplier relationships.	5 controls
A.16	Information security incident management	7 controls
A.17	Information security aspects of business continuity management.	4 controls
A.18	Compliance with internal requirements such as policies and with external requirements such as laws.	8 controls

But despite the complexity of legal regulation, OSINT intelligence techniques not only help investigators' cases, but also allow them to identify existing or potential factors that could destabilize a crisis situation. We will look at this issue in more detail in the next paragraph.

3 A Real-Time Social Media Monitoring System as an OSINT Platform for Early Warning in Crisis Situations

I would like to emphasize that the field of early warning early warning should be approached with moderate expectations, because early warning does not mean prophecy. Early warning means providing timely information that allows decision makers to analyze the data in detail and take interventions to avoid or reduce its unintended consequences and prepare for an effective response.

In recent years, the technological revolution on the one hand has contributed to the increasing complexity of the space in which current crises occur, and on the other hand. it continues to have a significant impact on the development of specific crisis management activities. First of all, the factor that increases the complexity of the space in which current crises occur is due to the extensive use of specific web-technologies by people and organizations, especially social media platforms, for disinformation and radicalization purposes.

Second, the IT revolution manifests itself in the process of crisis management. The transition from peace to distrust and then to confrontation or armed conflict can be detected by carefully defining and monitoring certain indicators using appropriate IT tools. These indicators should identify key characteristics of the crisis (actors and relations between actors) in order to assess the risk of triggers that degrade the situation.

Analyst Sherman Kent estimated that in peacetime, 80% of the information needed by decision makers is public 5. A platform for collecting and processing data from public sources has as a common goal is to provide decision makers with OSINT products in real time. In this regard, the platform must support all stages of the OSINT process. To do this, OSINT relies on sources such as:

- media;
- unobserved literature;
- data and informational commercial sources;
- online communities;
- user-generated digital content.

Conclusion

To summarize, the methods for collecting data from various social media networks have been reviewed. And also, an analysis was made of how public data can be accessed using various tools and interfaces, how the method can be legally regulated, and it was examined how OSINT acts in crisis situations. Although there are few options for changing the way public resources are accessed. Currently, open-source information is critical to decision makers. The number of active users on social media sites and information in general has increased exponentially. Therefore, there is a need to constantly adapt our methods and ways of gathering information, process and produce information in a timely manner to disseminate intelligence to decision makers in a timely manner to mitigate the effects of a crisis. This software has tremendous value in the field of digital investigation and forensics. People working in the cyberspace and in the field of crime investigation can use this software to gather information quickly and easily. The software can also be used by regular users to easily find contact information.

References

1. USING OSINT TO GATHER INFORMATION ABOUT A USER FROM MULTIPLE SOCIAL NETWORKS. – 2021. – Text: electronic.

2. OSINT Techniques Integration with Risk Assessment ISO/IEC 27001. – 2021. – Text: electronic.

3. A REAL-TIME SOCIAL MEDIA MONITORING SYSTEM AS AN OPEN SOURCE INTELLIGENCE (OSINT) PLATFORM FOR EARLY WARNING IN CRISIS SITUATIONS. – 2018. – Text: electronic.

4. Alan Calder and Steve G Watkins. 2010. Information security risk management for ISO27001/ISO27002. It Governance Ltd.

5. Allen Dulles, Memorandum Respecting Section 202, Central Intelligence Agency, April 25, 1947

Информация об авторах

Ксения Антоновна Щербакова — студентка кафедры информационной безопасности института радиоэлектроники и информационных технологий Уральского федерального университета имени первого Президента России Б. Н. Ельцина (Екатеринбург, Россия). E-mail: ksehcherbakova@gmail.com, https://orcid.org/ 0000-0003-2920-7980.

Губина Диляра Ильшатовна - старший преподаватель кафедры иностранных языков и перевода Уральского гуманитарного института Уральского федерального университета (Екатеринбург, Россия). E-mail: diliara.kurmanova@urfu.ru, https://orcid.org/0000-0002-1553-0413