ECONOMICS & Sociology

Yarovenko, H., Lopatka, A., Vasilyeva, T., & Vida, I. (2023). Socio-economic profiles of countries - cybercrime victims. *Economics and Sociology*, *16*(2), 167-194. doi:10.14254/2071-789X.2023/16-2/11

SOCIO-ECONOMIC PROFILES OF COUNTRIES - CYBERCRIME VICTIMS

Hanna Yarovenko
University Carlos III of Madrid, Madrid, Spain;
Sumy State University, Sumy, Ukraine
E-mail: hyaroven@inf.uc3m.es
ORCID 0000-0002-8760-6835

Agnieszka Lopatka
Department of Economics, Institute of Economics and Finance, University of Szczecin, Poland
E-mail: agnieszka.lopatka@usz.edu.pl
ORCID *0000-0001-9775-640X*

Tetyana Vasilyeva
Sumy State University, Sumy, Ukraine
E-mail: tavasilyeva@biem.sumdu.edu.ua
ORCID 0000-0003-0635-7978

Imre Vida*
Széchenyi István University, Győr, Hungary
info@vidaimre.com
ORCID 0000-0001-8089-9703
** Corresponding author*

ABSTRACT. The article analyses socio-economic profiles of countries that are victims of cybercrimes due to attacks by malicious programs and viruses spread through email applications, vulnerabilities of information systems and computer networks. The study is based on two hypotheses. The first is that powerful countries with significant global influence are both the cybercrime initiators and cybercrime victims to a greater extent than those with weak leverage. The second hypothesis is based on the fact that the level of socio-economic development of countries can be an indirect motivation for cyber criminals to commit mass cyberattacks. The proposed hypotheses were proved using cluster analysis based on the k-means and silhouette methods for the data from 93 countries. It formed 12 groups of countries based on the cyberattack volume on email applications and networks. Using the Farrar-Glauber test, the research revealed that identified vulnerabilities in information systems highly correlated with other factors. Thus, this factor was eliminated from the data set. An associative analysis was used to form a profile of the victim countries. It identified common socio-economic characteristics for each group and developed the rules of cause-and-effect relationships for them. The cluster analysis results confirm the first hypothesis that the most powerful countries, such as the USA, China, Germany, France, and others, are both victims of cyberattacks and their initiators. The analysis of profiles of countries' clusters based on the associative rules fully confirmed the second hypothesis.

Received: January, 2023
1st Revision: March, 2023
Accepted: June, 2023

DOI: 10.14254/2071-789X.2023/16-2/11

JEL Classification: C38, F01, H56

Keywords: cybercrime, cyberattack, cyberwar, socio-economic development, cluster analysis, silhouette method, association analysis.

Introduction

The Fourth Industrial Revolution led to the introduction of computer technologies into all areas of life. The development of Smart factories, powerful cyber-physical systems, and the Internet of Things have been contributing to the active economic and social development of many countries. On the other hand, mass computerization and digitalization stimulated cybercrime, which has become rampant across the world over the last decade. Nowadays, mass cybercrimes are committed not only to obtain financial benefits for individuals but also to exert non-violent influence on specific groups of people, companies, governments, and entire states. The mass nature and impact on a country's vital infrastructure facilities to disrupt their functioning or render them inactive can serve as signs to identify cyberattacks as cyberwar. Although Smith (2013) denies such an identification, Lucas (2016) considers the implementation of massive cyberattacks to be one of the major features of cyberwars. Despite the differences in views on this phenomenon, the undisputed fact is that the most powerful countries in global cyberspace, such as the USA, China, Great Britain, Russia, the Netherlands, France, Germany, Canada, Japan, and Australia, sometimes use its tools to fulfil completely non-peaceful goals (Voo et al., 2022).

In 2016, Russia interfered in the presidential elections in the United States, which was confirmed by the U.S. Department of Homeland Security and the Office of the Director of National Intelligence (U.S. Department of Homeland Security, 2016). In 2017, there was a large-scale cyberattack using the malware Petya (NotPetya) and WannaCry, which targeted various companies in Ukraine. Then the virus spread to other countries around the world, affecting large companies such as the American pharmaceutical corporation Merck, the Danish shipping company Maersk, the UK National Health Service, the German logistics company DHL, the Australian chocolate factory Cadbury, and many others (Perlroth, 2017). In 2019, the United Arab Emirates carried out a series of cyberattacks against its political opponents who were participating in a project on organising activities to track militants and terrorists (Bing & Schectman, 2019). In 2020, India launched a series of large-scale cyberattacks against Pakistani government services, as reported by the Tribune (2020). Due to vulnerabilities in Microsoft software, a Chinese cyber espionage unit hacked 30,000 American organizations, significantly affecting their operations (Krebs, 2021). In 2022, Ukraine became the object of military aggression by Russia. It was preceded by massive DoS and ransomware attacks against the Ukrainian government on January 13-14, 2022 (Deutsche Welle, 2022). Many other examples of cybercrimes can be given, but despite the difference in their goals and means of achievement, their impact on events and processes in different countries is significant.

Why are some countries more likely to become victims of cybercrimes while others are not of any interest in mass cyberattacks, espionage, terrorism, or other forms of cyberwar? What factors reduce the interest of cybercriminals and increasing protective reserves to counter this phenomenon? This study is aimed at obtaining answers to these questions. Thus, we will form several hypotheses for proving or rejecting which analytical calculations will be carried out in this article. They will allow us to develop profiles of countries that are cybercrime victims based on the most important indicators of socio-economic development. The first hypothesis is that countries that are the most powerful in the world and that initiate cybercrime are also more victims than those with a weak influence on the world stage. Another hypothesis is that the socio-economic development of countries can indirectly motivate cybercriminals to mass cyberattacks. Proving the proposed statements require various analytical methods. To solve the first question, it is advisable to group countries depending on the impact of different numbers of cyberattacks directed at them. According to the second hypothesis, it is possible to form

conclusions only if profiles are formed based on the key indicators that characterize the socio-economic development of countries.

1. Literature review

Nowadays, it is not easy to imagine implementing any processes in society without modern digital and computer technologies. Global challenges also contribute to their development. For example, such crises as the COVID-19 pandemic impose severe restrictions on societal changes, which can cause destructive processes in society and the economy (Dečman et al., 2022; Voronenko et al., 2022). On the other hand, such challenges cause an even more rapid development of digitalization, which becomes a favourable factor for minimizing the negative consequences for the countries' socio-economic development (Dluhopolskyi et al., 2023). As an opposing side of such events, there is an increase in the level of cybercrimes that occur in various spheres of society's life.

But rapid digitization and its reasons have a colossal impact on the socioeconomic development of any country (Melnyk et al., 2021; Millia et al., 2022) and even determine the interaction between it and the sustainability of ecosystems (Melnyk et al. et al., 2019). It promotes the convergence of countries, which was substantiated by Mačiulytė-Šniukienė et al. (2022) on the example of the EU. At the same time, its total integration into various state management processes occurs (Chen et al., 2023). The rapid development of technologies creates many risks that turn them into threats, especially for "leader states" (Barabashev et al., 2022). Among them, there is a risk of the spread of cyber war, which can be carried out covertly against other countries and lead to their stabilization on the world stage (Yarovenko, 2020). In this regard, the role of cyber security in ensuring and strengthening the state's national security is growing (Chen et al., 2023). There is also a need to develop ethical aspects in developing national strategies, which receive little attention from countries' governments (Fobel & Kuzior, 2019; Pakhnenko & Kuan, 2023).

It should be considered that several factors can directly or indirectly affect the level of cybercrime directed towards certain countries. Thus, Tiutiunyk et al. (2022a) analytically confirmed the existence of links between the shadow sector and socio-economic development. And in conditions of active development of cyberspace, it has become easier to conduct illegal operations and launder criminal proceeds. The study by Glova et al. (2020) determined the significance of corruption control as one of the indicators of the impact on the country's risk. Therefore, the corruption component can be an indicator of creating favourable conditions for the implementation of cybercriminal activity (Bozhenko, 2022). Remeikienė et al. (2022) proved the existence of a direct relationship between the level of crime and the country's economic development. Orlov et al. (2021) confirmed the importance of several economic and social factors for strengthening the country's security in general and cyber security in particular. Yarovenko & Rogkova (2022) found that recent research in the field of cybercrime is shifting towards its integration with the processes of corruption, financial cyberfraud, financial terrorism, cryptocurrencies and cyberwarfare.

Ensuring the cyber security of the financial system is one of the directions of preserving the economic and national security of the country. It becomes especially relevant in conditions of severe crises. For example, the intensive digitalization of the financial sector allowed Ukraine to withstand a period associated with active military aggression in February 2022 (Shkolnyk et al., 2022). Cryptocurrencies' creation and dynamic introduction into payment systems have led to an increase in cyber fraud using them (Pakhnenko et al., 2022; Gontareva et al., 2020). Bitcoin is a digital alternative to money, but despite its advantages, its active use is associated with risks to the financial system's stability and massive cyber fraud

(Mnohoghitnei et al., 2022). Today, the level of cyber security in banking institutions does not meet their urgent needs to protect transactions; this is evidenced by the unsatisfactory level of their digitalization and the dependence of many systems and processes on human influence (Tran et al., 2022; Vitvitskiy et al., 2021). This can be a source of mass cybercrimes.

Shao et al. (2022) proved that Internet technologies affect the development of territories and can give an impetus to the socio-economic development of the least developed countries. Thus, "Education 4.0" has been implemented for them, transforming the traditional education system based on the study of modern information technologies (Caballero-Morales et al., 2020). Yoshimori & Yoshimor (2022) found the influence of the development of cognitive skills in children on the digital phase of the economy. Some countries, such as China, understand the importance of digital education and are increasing spending on it (Yu et al., 2023). Also, within the framework of EU support, the concept of smart specialization began to be implemented to reform education and scientific research strategies aimed at the active implementation of information and digital technologies (Șavga, 2019; Șavga & Baran, 2022). But will such innovations backfire, where computer-literate users from less developed countries can enrich themselves through cybercrime or become flexible tools in cyberwars between the world's most powerful nations? The answer to this question is open.

Kuzior & Kwilinski (2022) have noted that artificial intelligence systems and cognitive technologies are actively being implemented in various spheres of society's life and are trying to replace humans in performing many processes. For example, cognitive production technologies significantly increase the efficiency of complex operational production processes, positively affecting the company's profit (Lăzăroiu et al., 2022). But on the other hand, digitalisation can cause an imbalance in management processes (Straková et al., 2022; Vasudevan, 2022). According to the study by Gurbanov et al. (2022), 42% of respondents felt a positive impact of digitalisation in companies in crisis conditions, 38% did not feel any changes, and 20% had a negative experience. That is, digitisation has a negative aspect, and it must be considered. This leads to the generation of cybercriminal insiders, whose activities can be aimed at breaching security systems within the company and creating vulnerabilities for third-party intervention. Orientation of a business to its promotion in online networks helps to increase the efficiency of many of its processes (Kurniawati et al., 2022). But it also leads to a magnification in its vulnerabilities due to the emergence of new sources of cyberattacks aimed at the profiles of customers and company employees in social networks (Bozhenko et al., 2022b). Personal security issues arise, as many users neglect its basics and disclose their personal data when making transactions (Zimaitis et al., 2022).

Using software robots helps increase companies' product competitiveness (Sobczak, 2022). In the near future, integrating intelligent computers with humans and complex network systems is planned (Kumar & Kumar, 2019). At the same time, it can make it much easier to conduct cyberattacks on fully robotic production, which can lead to serious financial consequences for such a business. Also, capturing strategically important enterprises by other countries in the framework of cyber wars will become a straightforward matter. The results of Industry 4.0 enable the practical implementation of complex business concepts such as the Internet of Things and Big Data Analytics (Ćwiklicki & Wojnarowska, 2020). On the other hand, its consequences can cause many threats to competition between businesses (Tvaronaviciene & Burinskas, 2020). This can increase cyber threats and increase cyber espionage to obtain trade secrets and use them against competitors.

The problem of cybercrime directed in the form of aggression against countries is related to various aspects. Therefore, this research requires the use of appropriate mathematical apparatus. To solve these problems, scientists have actively used statistical methods (Adeyemo et al., 2020), regression models (Safarov et al., 2022), lag econometric models (Tiutiunyk et al.,

2022b), vector error correction (Lyulyov et al., 2021), mar-spline approach (Bozhenko et al., 2022a), DEA-model and its extension (Wang et al., 2022), Fuzzy Logic (Bayram & Akat, 2019), game theory (Stehel et al., 2019), Artificial Intelligence (Gupta & Mishra, 2022) and Data Mining tools (Kuzmenko et al., 2020). In this study, Data Mining tools will be the most effective, allowing you to use different types and arrays of data, conduct analysis according to various parameters, and easily interpret the obtained results.

2. Preliminary analysis of input data

Two data sets were chosen for the study. One of them formed clusters of countries depending on the level of detected cybercrimes directed at them. The data source is a Kaspersky Lab resource (Kaspersky, 2023). The second set of data was formed by indicators characterizing the socio-economic development of countries, considering their influence on macro- and global processes. It enabled to conduct an analysis of potential attractiveness for cyber fraudsters and identify areas that require attention from international organizations and the government to combat cybercrime.

The first data set was generated for 93 countries, representing three types of cybercrimes in one month of 2022-2023. The first set contained malware and viruses detected using Mail Anti-Virus (MAV) software. This choice is justified because phishing cyber-attacks using email applications take first place among other types of crime in 2022 (Statista, 2023). *Figure 1* presents a map of the distribution of this cybercrime type based on the analysed indicators. The most attacked countries are Spain, Mexico, Turkey, Vietnam, Italy, United Arab Emirates, Germany, Brazil, Colombia and Malaysia. The least attacked are Norway, Mongolia, Kyrgyzstan, Luxembourg, Nicaragua, Rwanda, New Zealand, Sweden, Denmark and Ethiopia.

The second type of cybercrime included network cyberattacks, detected by the "Intrusion Detection Scan" (IDS) system. If the first type of cybercrime is aimed specifically at the target user, the second type has more harmful consequences since the entire network is affected, leading to the disruption of the entire company's work. The purpose of such a crime is to inflict mass damage on as many public and corporate sector users as possible. Network attacks lead to downtime in companies, loss of large volumes of data, and, as a result, increased financial losses. China, the United States, Brazil, Mexico, Vietnam, France, India, Indonesia, Germany, and Spain became the target countries for this type of cybercrime (*Figure 2*). The least attacked are Norway, Rwanda, Albania, Zimbabwe, Georgia, New Zealand, Guatemala, Montenegro, Zambia and Cyprus.

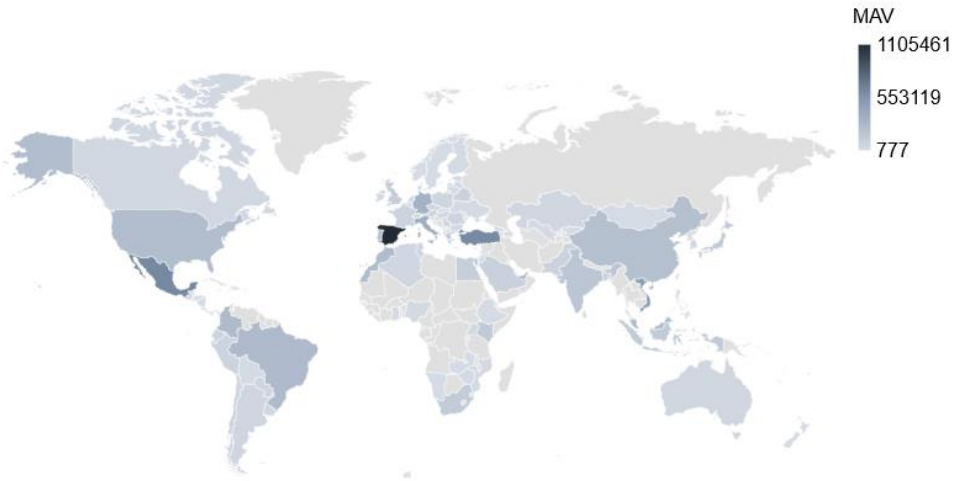


Figure 1. Map of detection of malware and viruses distributed through email applications
Source: *own compilation based on Kaspersky (2023)*

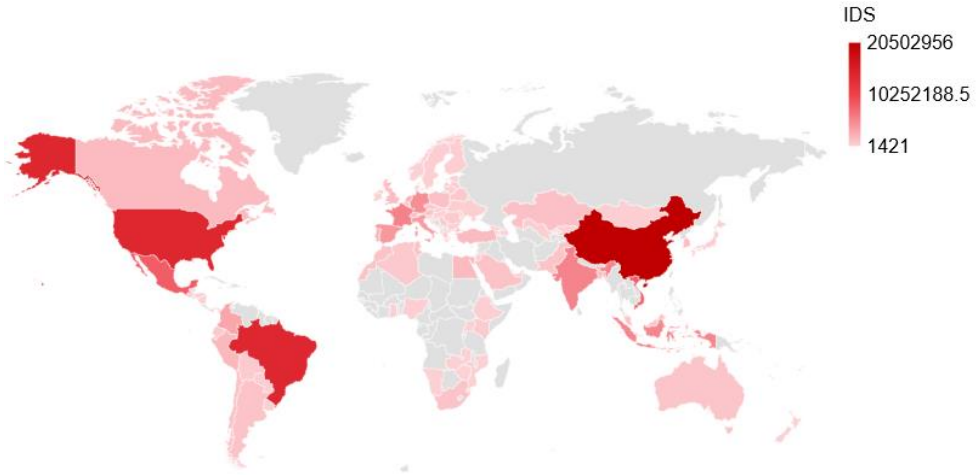


Figure 2. Map of network attacks detection
Source: *own compilation based on Kaspersky (2023)*

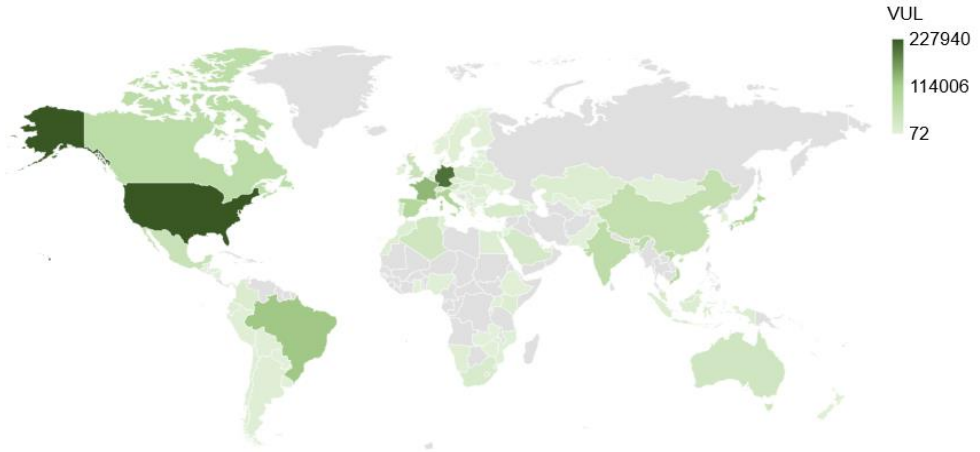


Figure 3. Map of system vulnerabilities detection
Source: *own compilation based on Kaspersky (2023)*

Vulnerabilities in software, computers and networks often allow cybercriminals to conduct more active cyberattacks and compromise the security of various users. Such threats arise due to imperfect programming and misconfiguration of routers, application servers, web servers, firewalls, and other hardware and software. For their analysis, some vulnerabilities in information systems ("Vulnerability" VUL), detected with the help of scanner programs, were selected. The USA, Germany, France, Brazil, Italy, Japan, Spain, Vietnam, Canada and India are the top 10 countries that have experienced the largest number of attacks to identify weak points in information systems (*Figure 3*). Nicaragua, Montenegro, Georgia, Armenia, Uruguay, Finland, Pakistan, Cyprus, Kyrgyzstan and Mongolia experienced the least amount of cybercrime of this type.

The analysis of basic statistics of selected cybercrimes, presented in *Table 1*, shows a big difference between the volume of cyberattacks for different countries. Thus, users of some countries may not experience cyber intrusions through email applications, network attacks or system vulnerabilities, as evidenced by their minimum value, which is thousands to tens of thousands of times less than the maximum. On the other hand, some countries are subjected to brutal cyberattacks, as evidenced by the colossal maximum values of the number of incidents. The calculated median, mean, asymmetry and kurtosis indicators indicate the uneven

distribution of cybercrimes. It can be concluded that many countries form a direct target audience for cybercriminals, and there are countries that either have significant resources to counter or are not attractive to cyberwars.

Table 1. Basic statistics defined for three types of cybercrime

Statistical Indicator	MAV	IDS	VUL
Mean	92015.30	1500464.72	20193.98
Mode	–	–	–
Standard Error	16935.15	330862.77	4210.08
Standard Deviation	163316.68	3190725.02	40600.58
Kurtosis	17.30	17.53	11.64
Skewness	3.68	3.92	3.23
Minimum	777.00	1421.00	72.00
First quartile	7920.00	110523.00	1340.00
Median	31154.00	363228.00	3315.00
Third quartile	91051.00	1099658.00	15018.00
Maximum	1105461.00	20502956.00	227940.00
Sum	8557423.00	139543219.00	1878040.00
Count	93.00	93.00	93.00
Confidence Level (95.0%)	33634.67	657122.02	8361.59

Source: *own calculations*

Several indicators of socio-economic development for 2022 for 93 countries were selected to identify characteristics of the country's attractiveness to cybercriminals. The National Cyber Security Index (NCSI) was chosen first. It allows for assessing the country's level against cyber threats (E-Governance Academy, 2023). A high level of national cyber security allows for forming a powerful defence base regarding legal, informational, technical, software, organizational and other security system support. It should contribute to reducing the country's attractiveness to massive cyberattacks. Since cybercrimes are nowadays used to spread cyberterrorism, selecting the Global Terrorism Index (GTI) indicator will reveal the country's impact on global terrorism as a whole (Institute for Economics and Peace, 2022). Thus, the countries with the highest level of terrorism can be the biggest victims of massive hacking attacks or their initiators. In contrast, the countries with the lowest impact, on the contrary, will not become their target. Understanding the circumstances surrounding the state of crime within the country, measured by the Crime Index (Numbeo, 2023), is important to form a profile regarding potential attractiveness for cybercriminals. It assesses the internal situation of the favourable conditions to develop and support various types of crimes in a specific country. Since the popularity of the Darknet is growing today and most criminal actions are carried out using computer technologies, the analysis of this indicator will allow us to assess how the internal environment contributes to the formation of conditions for the support of cybercrime. It will turn the country not only into its victim but also into an active cyberterrorist. The country's image in cyberspace can also be affected by the level of corruption, forming a suitable plane for the legalization of funds, illegal redistribution of cash flows, violation of legislation, etc. The Corruption Perceptions Index (CRI) (Transparency International, 2023) was chosen to analyse this feature.

The country's economic freedom level affects the formation of its socio-economic profile. It allows measuring the relationships between various economic spheres: public finances, business, taxes, investment and tax spheres, trade, government honesty and the effectiveness of the judicial system, etc. As a rule, the economic component drives scientific

and technical progress development, affecting a suitable cyber environment in a separate country. Therefore, the Index of Economic Freedom, IEF (The Heritage Foundation, 2023) was chosen for analysis in this area. In addition to economic well-being, the population's satisfaction with health, education, art, culture, the environment, employment opportunities, and psychological support are essential. These aspects can be evaluated using the Happiness Index, NO (World Happiness Report, 2023). Compared to the less fortunate, the luckiest countries may attract cybercriminals precisely to obtain financial benefits from this crime. Life Expectancy at Birth (LE) is an indicator that characterizes countries' socio-economic development levels. Its highest values correspond to economically developed countries and the lowest - to least-developed countries (The World Bank, 2023). The last characteristic selected is the level of democracy, which allows for assessing the civil and political freedoms the country's government respects. The Democracy Index (DI) is used (Economist Intelligence, 2023). The presence or absence of such rights and freedoms seriously affects the formation of an unfavourable environment for sustainable socio-economic development of the country, which can also cause a particular interest for cybercriminals.

The listed indicators were selected for 93 countries for 2022. The analysis results of their basic statistics are shown in *Table 2*.

Table 2. Basic statistics defined for socio-economic development indicators

Statistical Indicator	NCSI	CPI	DI	HI	LE	GTI	IEF	CI
Mean	59.5870	49.3548	6.1411	5.8568	74.1943	2.1476	64.0323	43.3312
Mode	59.7400	36.0000	7.9700	4.5160	–	0.0000	74.4000	46.1000
Standard Error	2.2611	1.9753	0.2248	0.1076	0.7185	0.2437	1.0954	1.4213
Standard Deviation	21.8056	19.0495	2.1680	1.0375	6.9287	2.3506	10.5637	13.7063
Kurtosis	-0.7821	-0.9355	-0.8826	-0.0889	0.0707	-0.3215	-0.3272	-0.7468
Skewness	-0.3059	0.4706	-0.3728	-0.4704	-0.6294	0.8727	-0.2389	0.0851
Minimum	9.0900	19.0000	1.9400	2.9950	52.6760	0.0000	33.1000	15.1000
First quartile	41.5600	36.0000	4.5500	5.1730	70.2300	0.0000	55.7000	32.1000
Median	62.3400	45.0000	6.4500	6.0220	74.2560	1.2430	65.1000	45.4000
Third quartile	76.6200	63.0000	7.9500	6.4800	80.8756	4.1060	71.8000	53.7000
Maximum	94.8100	90.0000	9.8100	7.8210	84.4456	8.2330	84.4000	76.1000
Count	93	93	93	93	93	93	93	93
Confidence Level (95.0%)	4.4908	3.9232	0.4465	0.2137	1.4270	0.4841	2.1756	2.8228

Source: *own calculations*

Generally, there is a slight imbalance among the data, as evidenced by indicators such as Mean, Minimum, Maximum, First quartile, Third quartile, and Median. But this is explained by the fact that the selected set included countries with different socio-economic development. The kurtosis and skewness values indicate that most of the data are close to a normal distribution and are acceptable for further analysis. There are several outliers by the country for the Happiness Index and Life Expectancy at Birth, which will not critically affect the further results of the analysis. For such indicators as Crime Index, Global Terrorism Index and Corruption Perceptions Index, it is observed that most countries are included in the 3rd and 4th quartiles,

as for others, on the contrary, in the 1st and 2nd. It is due to the fact that the listed indicators are in their content destimulators, and others are stimulants, which must be taken into account in the process of normalization and assigning a rating.

Thus, two sets of data have been created to analyse the profiles of countries that are victims of cyberattacks through email applications, networks and vulnerabilities of information systems.

3. Methodological approach

The methodology for researching socio-economic profiles of countries that are victims of cybercrimes was carried out in three stages. The implementation of the first is related to the pre-processing of the data, the determination of the presence or absence of multicollinearity between the three types of cybercrimes, and the standardization of the observation values. The second stage is related to the clustering of countries based on the quantitative value of those types of cybercrimes that are not multicollinear. Cluster consistency is checked using the Silhouette method. The third stage is necessary to identify socio-economic patterns peculiar to certain countries, implemented using associative analysis.

The first stage of the research consisted in pre-processing the input data. Since they were collected manually, handling missing values was unnecessary. Also, the data did not require an abnormality study since, in our case, such observations indicate an excess of cyberattacks towards this country.

For cluster analysis, it is necessary to check for multicollinearity and standardize the data. Standardization allows you to remove the mean value and increase the scale to the value of the variance. This procedure was carried out according to formula (1):

$$x_{ij}^{scaled} = \frac{x_{ij} - \overline{x_j}}{\sigma_j}, \quad (1)$$

where x_{ij}^{scaled} – standardized value of j–th cybercrime in i–th country, x_{ij} – actual value of j–th cybercrime in i–th country, $\overline{x_j}$ – sample mean for j–th cybercrime, σ_j – sample standard deviation for the j–th type of cybercrime.

The Farrar–Glauber algorithm was used to test the data for multicollinearity. It involves the calculation of Chi-squared by formula (2):

$$\mathrm{X}^2 = -\left((n-1) - \frac{2m+5}{6}\right) \times ln|R|, \quad (2)$$

where X^2 – calculated Chi-squared value, n – the number of observations in the array of variables, which is equal to 93 countries, m – the number of explanatory variables, which is equal to 3 types of cybercrimes, $|R|$ – the determinant of the matrix, formed from pairwise correlation coefficients, i.e:

$$R = \begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{pmatrix}, \quad (3)$$

where r_{kj} – the value of correlation coefficients between pairs of explanatory variables that correspond to the investigated cybercrimes ($k = 1 \div 3$; $j = 1 \div 3$), calculated by (4):

$$r_{kj} = \frac{\sum_{i=1}^{n}\left(\left(x_i^k - \overline{x^k}\right)\left(x_i^j - \overline{x^j}\right)\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i^k - \overline{x^k}\right)^2 \sum_{i=1}^{n}\left(x_i^j - \overline{x^j}\right)^2}}. \quad (4)$$

The calculated Chi-squared value is compared with critical ones at $\frac{1}{2}m(m-1)$ – degrees of freedom and the corresponding significance level α. If $X^2 > X_{cr}^2$, there is multicollinearity in the array of variables. They should be checked, otherwise there is no multicollinearity and the check is not performed.

For further research, the Fisher test is calculated by (5), which allows for determining the correlation of a separate factor with others:

$$F_k = (a_{kk} - 1) \times \frac{(n-m)}{(m-1)}, \quad (5)$$

where F_k – Fisher test value, calculated separately for each of the three variables, a_{kk} – diagonal element of the matrix inverse of the matrix R. The calculated Fisher test value is compared with the critical value $F_{cr}(\alpha, k_1, k_2)$, where α – appropriate level of significance, $k_1 = n - m$ та $k_2 = m - 1$. If $F_k > F_{cr}$, the relevant variable is correlated with the others. In contrast, it does not correlate with others.

Then partial correlation coefficients, which show the closeness of the relationship between two variables without considering the influence of other variables, are calculated by formula (6):

$$r_{kj}^* = \frac{-a_{kj}}{\sqrt{a_{kk}a_{jj}}}, \quad (6)$$

where r_{kj}^* – values of partial correlation coefficients between pairs of explanatory variables that correspond to the investigated cybercrime ($k = 1 \div 3$; $j = 1 \div 3$), a_{kj}, a_{jj} – the corresponding elements of the matrix inverse of the R matrix . If the calculated values are close to 1 or -1, it indicates a strong correlation between the variables. Critical values $r_{cr}(\alpha, v)$ from Table Fisher–Yates can be used to obtain a refined conclusion, where α – appropriate significance level, $v = n - l - 2$, where l – the number of excluded values in the case of partial correlation, n – the number of observations.

During the last step, the Student's criterion is calculated by formula (7) to check the statistical significance of partial correlation coefficients:

$$t_{kj} = \frac{r_{kj}^*\sqrt{n-m}}{\sqrt{1 - r_{kj}^{*\,2}}}. \quad (7)$$

The obtained values of Student's criterion are compared with its critical value $t_{cr}(\alpha, k)$, where α – appropriate significance level, $k = n - m$. If $|t_{kj}| > t_{cr}$, the correlation dependence between the variables is statistically significant, otherwise the dependence is not statistically significant.

In the second stage of the research, a cluster analysis was carried out using the k-means method and a check of the cluster consistency using the Silhouette technique. K-means clustering is one of the Data Mining methods, allowing to divide the data set into a certain number of groups (clusters), provided that each observation is close to the corresponding cluster centroid (mean).

It means that the goal of cluster analysis is to minimize the variance within the cluster and find the optimal observation distance to the middle of the group, which can be represented by formula (8)

$$\arg\min_{C} \sum_{i=1}^{k} \sum_{x_p \in C_i} \left|\left|x_p - \mu_i\right|\right|^2 = \arg\min_{C} \sum_{i=1}^{k} |C_i|\, Var C_i, \quad (8)$$

where $(x_1, x_2, \dots, x_p)$ – a set of variables, each of which represents d – measuring vector with n – observations in each, μ_i – centroid of i-th cluster, calculated by formula (9):

$$\mu_i = \frac{1}{|C_i|} \sum_{p \in C_i} x_p, \quad (9)$$

where $C = \{C_1, C_2, \dots, C_k\}$ – sets of variables corresponding i-th cluster at $i = 1 \div k$. At the same time, the appropriateness of the i-th cluster observations is shown by formula (10):

$$\mathrm{C}_i = \{p | if\ x_p\ belongs\ to\ the\ i^{th}\ cluster\}. \quad (10)$$

The silhouette is a visualization method for testing data consistency in clusters proposed by Rousseeuw (1987). This technique involves determining the silhouette coefficient for all samples, considering the average distance to the centre of the cluster and the average distance to the nearest cluster by formula (11):

$$\begin{cases} s(i) = \dfrac{b(i) - a(i)}{max\{a(i), b(i)\}}, if\ |C_I| > 1, \\ s(i) = 0, if\ |C_I| = 1 \end{cases} \quad (11)$$

where $s(i)$ – a silhouette value for i-th observation from the data set, $a(i)$ – average distance between i-th and other observations in the cluster, calculated by formula (12), $b(i)$ – average distance from i-th observations in the cluster to other observations of other clusters, calculated by formula (13), $|C_I|$– set of observations of one cluster:

$$a(i) = \frac{1}{|C_I| - 1} \sum_{j \in C_I, i \neq j} d(i, j), \quad (12)$$

$$b(i) = \min_{J \neq I} \frac{1}{|C_J|} \sum_{j \in C_J} d(i, j), \quad (13)$$

where $d(i, j)$ – distance from i-th observation to j-th.

It is necessary that $-1 \leq s(i) \leq 1$ to identify the results. If the silhouette value approaches 1, data is appropriately grouped. If its value is close to 0, then the data is on the border of two clusters and it is difficult to attribute it to a specific cluster. If the silhouette approaches -1, the data belongs to another cluster.

The third stage of the research requires pre-processing of data that correspond to socio-economic factors. First, they need to be normalised according to formula (14) if the indicator is a stimulant, and formula (15) if the indicator is a destimulator:

$$x'_{ij} = \frac{x_{ij} - x_j^{min}}{x_j^{max} - x_j^{min}}, \tag{14}$$

$$x'_{ij} = \frac{x_j^{max} - x_j}{x_j^{max} - x_j^{min}}, \tag{15}$$

where x'_{ij} – normalized value of i-th observations for j-th variable, x_j^{min} and x_j^{max} – accordingly, minimum and maximum value for j-th variable.

It is necessary to replace data with rating groups to implement the associative analysis. It is due to a small number of observations and a large variation in their values. For this, we will use the formula (16):

$$x_{ij}^* = \begin{cases} x_{ij} = 1, if\ 0 < x_{ij} \leq 0.25 \\ x_{ij} = 2, if\ 0.25 < x_{ij} \leq 0.50 \\ x_{ij} = 3, if\ 0.5 < x_{ij} \leq 0.75 \\ x_{ij} = 4, if\ 0.75 < x_{ij} \leq 1 \end{cases}, \tag{16}$$

where x_{ij}^* – rating value of i-th observations for j-th variable; 1 – the rating value that corresponds to the low value of the indicator included in the first 25%; 2 – the ranking value that corresponds below the mean value of the indicator, included in the second quartile of the sample values; 3 – the ranking value that corresponds above the mean value of the indicator, included in the third quartile of the sample values; 4 – the ranking value that corresponds to the high value of the indicator, included in the fourth quartile of the sample values.

This stage of the research consists of conducting an associative analysis, which will reveal the reasons for connecting the analysed indicators for a particular cluster of countries. It will contribute to forming the countries' profiles of cybercrime victims based on their socio-economic development factors. It will help to understand the motivation of criminals to carry out targeted cyberattacks. The Apriori algorithm, based on the detection of frequency sets of data in the set, is used to implement this type of analysis. It will form a list of typical factors for clusters of countries. Also, its construction of associations and correlations will contribute to the identification of causal relationships within a separate group of countries. The following indicators are determined by formula (17) to identify associative rules:

$$supp(X \Longrightarrow Y) = \frac{F(X,Y)}{N},$$

$$conf(X \Longrightarrow Y) = \frac{F(X,Y)}{F(X)}, \tag{17}$$

$$lift(X \Longrightarrow Y) = \frac{S(X \Rightarrow Y)}{S(X) \times S(Y)},$$

where $supp\ (support)$ – an indicator characterizing the frequency of appearance of X and Y elements; $conf\ (confidence)$ – an indicator allowing to determine the percentage of elements that satisfy the element condition X, which also satisfy the element condition Y; $lift$ – an indicator that shows the interest level in an item Y provided there is an interest in the item X; if $lift(X \Rightarrow Y) = 1$ – there is no correlation in the data set; if $lift(X \Rightarrow Y) > 1$ – the correlation is positive, i.e., the probability of compatible implementation of X and Y elements is high; if $lift(X \Rightarrow Y) < 1$ – the correlation is negative, i.e., i.e. compatible implementation of X and Y elements is unlikely. Since the calculations of the associative analysis were performed in the analytical package STATISTICA, the indicator $lift$ was designated as $correlation$.

4. Empirical results

The first stage of the methodology regarding data standardization and checking for multicollinearity using the Farrar-Glauber test was calculated using MS Excel software. The results of the test are presented in *Table 3*, where multicollinearity in the dataset is formed based on three types of cybercrimes, since $X^2 > X^2_{cr}$.

Table 3. The results of Farrar-Glauber test

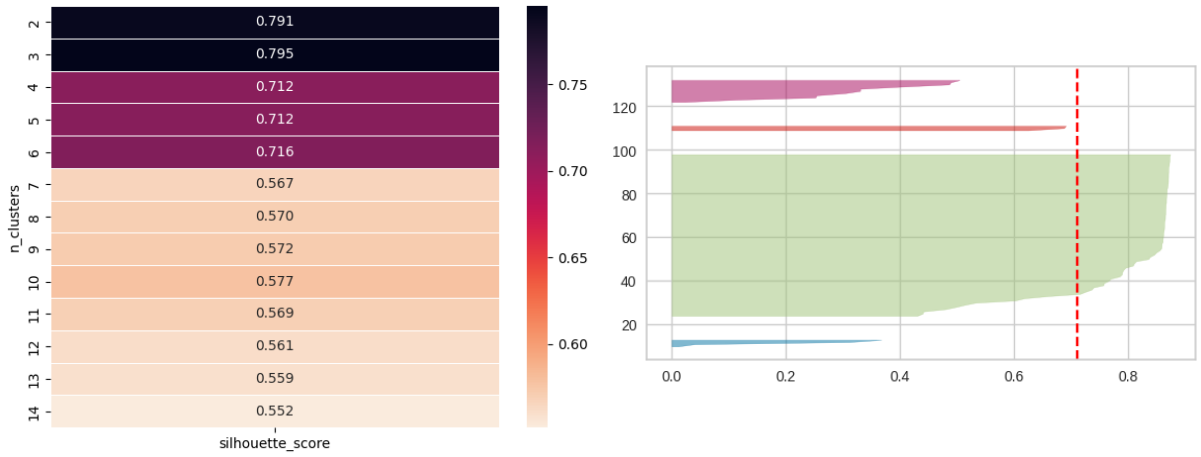
Estimated criterion	Estimated value	Inequality sign	Critical criterion	Critical value	Check results
X^2	81.6434	>	X^2_{cr}	7.8147	Multicollinearity is present
F_{MAV}	16.8229	<	F_{cr}	19.4846	Non-multicollinear
F_{IDS}	39.5230	>			Multicollinear
F_{VUL}	42.7960	>			Multicollinear
$r_{MAV,IDS}$	0.2040	<	r_{cr}	0.2050	Non-multicollinear
$r_{MAV,VUL}$	0.2781	>			Multicollinear
$r_{IDS,VUL}$	0.5702	>			Multicollinear
$t_{MAV,IDS}$	1.9767	<	t_{cr}	1.9867	Statistically significant
$t_{MAV,VUL}$	2.7466	>			Statistically insignificant
$t_{IDS,VUL}$	6.5848	>			Statistically insignificant

Source: *own calculations*

Further testing using Fisher's test, partial correlation, and Student's test revealed that the variable corresponding to the number of malware and viruses distributed through email applications was not multicollinear with the others. As for the number of network attacks factor, in combination with the previous variable, it is not multicollinear ($r_{MAV,IDS} < r_{cr}$, $t_{MAV,IDS} < t_{cr}$). The third variable, which characterizes the number of detected attacks on system vulnerabilities, is collinear with the others ($r_{IDS,VUL} > r_{cr}$, $t_{IDS,VUL} > t_{cr}$, $r_{MAV,VUL} > r_{cr}$, $t_{MAV,VUL} > t_{cr}$). Eliminating the multicollinearity from an array of variables is performed by the principal component method. In our case, its application did not lead to obtaining a data set

that would satisfy all conditions. Therefore, to carry out clustering, it was decided to eliminate the VUL variable and to carry out clustering taking into account only two variables.

Cluster analysis and verification of the cluster consistency using the Silhouette method was conducted using the Python programming language. Since clustering made it possible to obtain uneven sizes of clusters, it became necessary to carry out this procedure in several stages using the example of hierarchical clustering. The results of the first stage are presented in *Figure 4*.

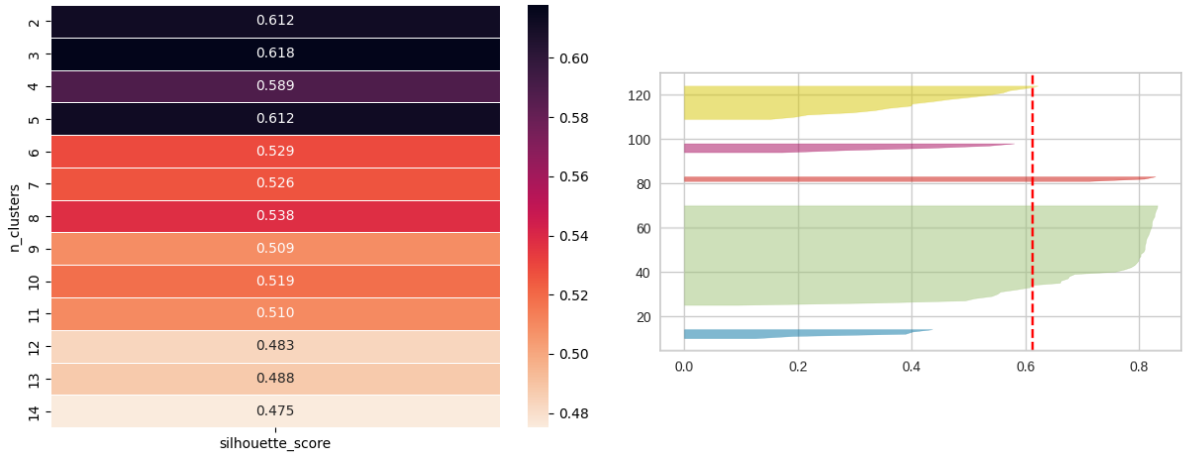


a) b)

Figure. 4. Results of the first clustering stage: a) Silhouette score; b) Silhouette plot
Source: *own calculations*

The highest Silhouette score corresponds to dividing the data into 3 clusters (*Figure 4a*). Under this condition, a share of incorrect classification was also obtained, i.e., some countries (Germany) assigned to the cluster do not belong. It was decided to perform a cluster analysis for 4 groups to reduce the share of incorrect classification. *Figure 4b* confirms the correctness of this division. You can also see that the study identified a cluster that contains 80.65% of all data. It is due to the uneven distribution of initial data, which is caused by the unevenness of cyberattacks on countries. At the same time, the most attacked countries did not fall into this cluster. Therefore, the next clustering stage was carried out for countries in the largest cluster. The second stage results are presented in *Figure 5*.

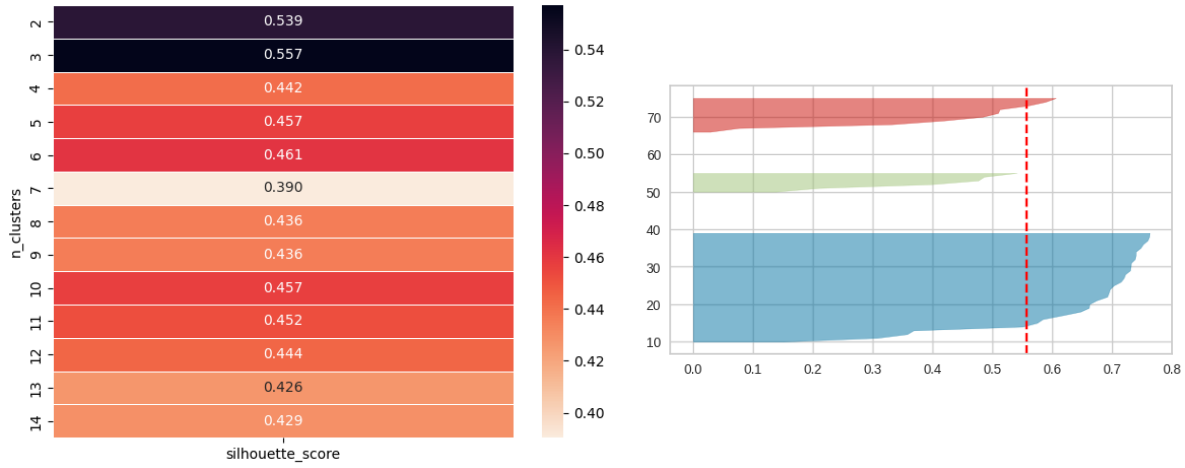
Although the highest Silhouette scores correspond to the three-cluster distribution of the data, a misclassification proportion was also obtained for this situation. This procedure assigned Serbia to another cluster, as evidenced by the negative value of the Silhouette score (-0.1090). Using five clusters allows for avoiding incorrectly classified objects, which is confirmed by *Figure 5b*. It suggests the feasibility of using this type of distribution. Regardless of the number of clusters, it was concluded that one contains 61.33% of the sample data. Thus, there is a need to divide further the sample obtained in the second stage.



a) b)

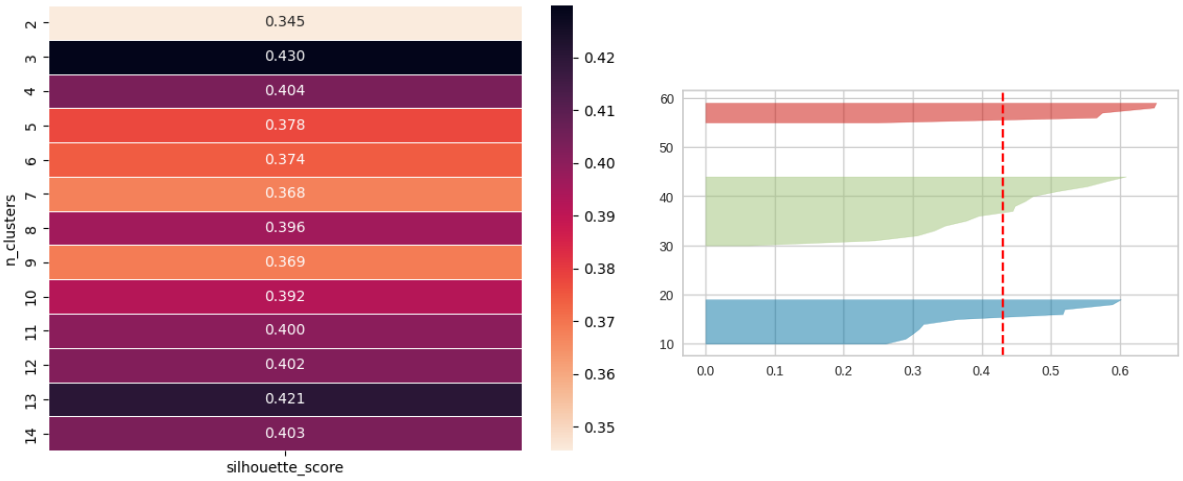
Figure. 5. Results of the second clustering stage: a) Silhouette score; b) Silhouette plot
Source: *own calculations*

The results of the third step of clustering for the countries included in the largest cluster are presented in *Figure 6*. The obtained Silhouette scores are significant for the three-cluster distribution (*Figure 6a*). At the same time, all countries were classified correctly (*Figure 6b*). But at this step, a cluster was also formed, which contains 65.22% of all observations of the sample taken for this step, indicating the need to continue the data clustering procedure for the largest group of countries. The results of the fourth and last stage of clustering are presented in *Figure 7*. The highest value of the Silhouette score corresponds to a three-cluster distribution (*Figure 7a*). Silhouette visualization confirms the correctness of the obtained results with no fraction of misclassification of objects (*Figure 7b*). Although one cluster consists of 50% of the sample, this value corresponds to 16% of the general population, which is acceptable for data analysis. The impracticality of further clustering also confirms the decrease in Silhouette score, which occurs from stage to stage.



a) b)

Figure. 6. Results of the third clustering stage: a) Silhouette score; b) Silhouette plot
Source: *own calculations*



a) b)

Figure. 7. Results of the fourth clustering stage: a) Silhouette score; b) Silhouette plot
Source: *own calculations*

Figure 8 presents a map of countries divided by defined clusters. *Table 4* demonstrates cluster-averaged values for each group of cybercrimes. The type that was removed from the clustering process is also considered here. The countries of clusters 1.3, 1.2 and 1.1 are the most attacked. Countries in groups 4.1, 4.2 and 4.3 are the least attacked.

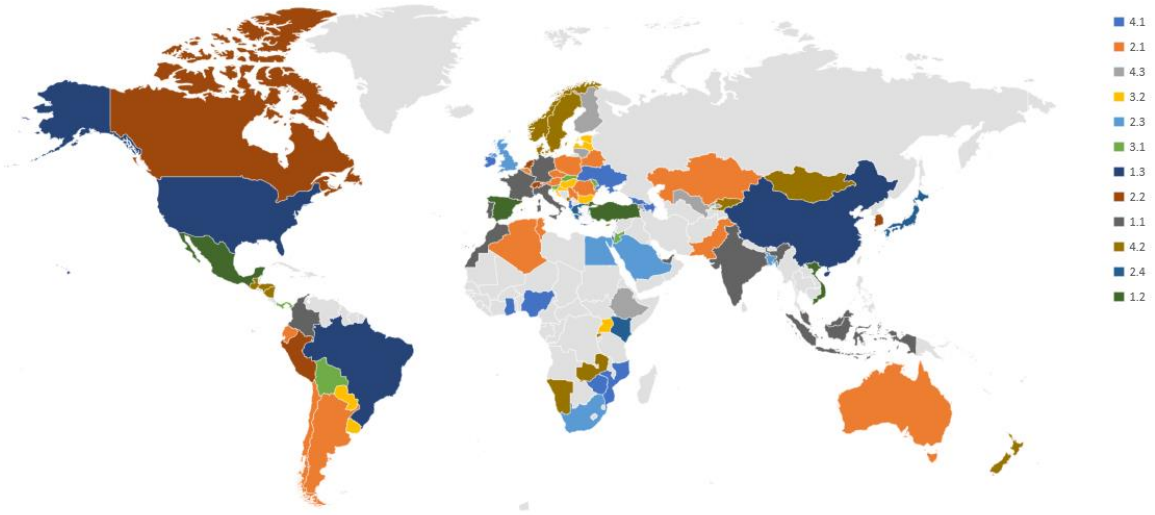


Figure. 8. Map of countries divided into clusters depending on detected cybercrimes
Source: *own calculations*

Table 4. The results of cluster analysis

Cluster	Countries	Average of MAV	Average of IDS	Average of VUL
1.1	France, Germany, India, Italy, Colombia, Indonesia, Iran, Malaysia, Morocco, Portugal, United Arab Emirates	226463.27	3191437.27	54251.55
1.2	Mexico, Vietnam, Turkey, Spain	709347.25	5406983.75	60429.00
1.3	United States, China, Brazil	241351.33	16181266.33	136751.33
2.1	Algeria, Australia, Austria, Poland, Argentina, Belarus, Belgium, Czech Republic, Ecuador, Kazakhstan, Romania, Tunisia, Chile, Pakistan, Serbia, Singapore	51759.31	775175.06	12641.88
2.2	Netherlands, Canada, Switzerland, South Korea, Peru	37788.00	1751951.40	23692.20
2.3	Bangladesh, Egypt, South Africa, Saudi Arabia, United Kingdom	107084.20	1188240.80	22990.40
2.4	Japan, Greece, Kenya	132841.67	256754.33	38421.00
3.1	Panama, Bolivia, Jordan, Moldova, Slovakia, Slovenia	16229.17	346157.33	1560.67
3.2	Bahrain, Bulgaria, Croatia, Estonia, Hungary, Paraguay, Uganda, Uruguay, Tanzania, Latvia	29107.20	167196.00	1904.50
4.1	Albania, Azerbaijan, Georgia, Ghana, Ireland, Zimbabwe, Israel, Mozambique, Nigeria, Ukraine	9014.10	65207.00	1973.40
4.2	Cyprus, Guatemala, Honduras, Kyrgyzstan, Luxembourg, Mongolia, Montenegro, New Zealand, Nicaragua, Norway, Rwanda, Zambia, Denmark, Namibia, Sweden	2873.93	69052.33	1702.13
4.3	Armenia, Ethiopia, Finland, Lithuania, Uzbekistan	6430.40	190978.20	936.40

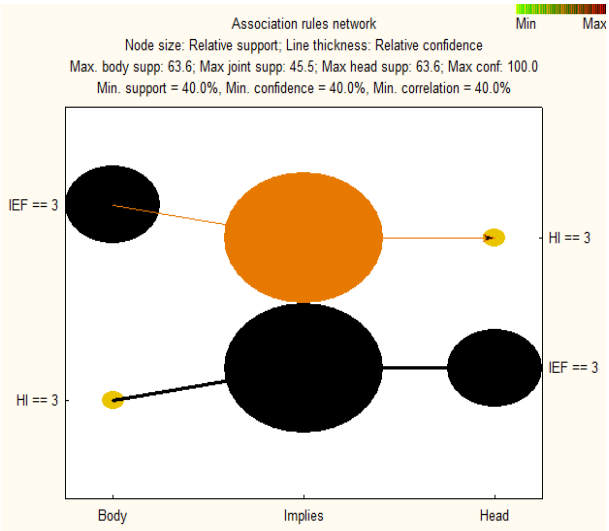
Source: *own calculations*

We will use the statistical data of the Power Index to prove or reject the proposed first hypothesis. It is determined based on 50 factors that combine military, economic, and cultural potential (Wisevoter, 2023). According to this rating, only 25 countries are among the most powerful. At the same time, 12 of them belong to the countries that are the most attacked, i.e., they are the countries of clusters 1.3 (USA, China and Brazil), 1.2 (Spain, Turkey and Vietnam) and 1.1 (France, Germany, Italy, India, Indonesia and Iran). It means that the most attacked countries are also the most powerful globally. At the same time, China (18.83%), the USA (17.05%), Brazil (5.63%), India (5.33%), Germany (5.10%), Vietnam are among the top 10 countries that carry out cyberattacks towards others (4.23%), Thailand (2.51%), Russia (2.46%), Indonesia (2.41%), Netherlands (2.20%) (DavidPur, 2022). Seven countries from this

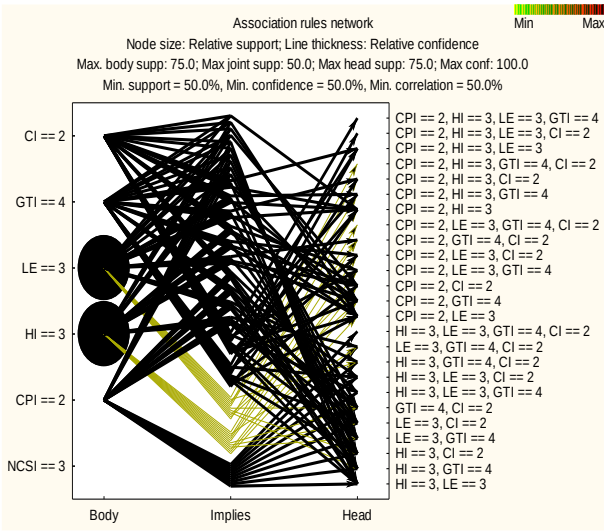
list are countries of clusters 1.1, 1.2 and 1.3. Information on Thailand and Russia is unavailable in the sample selected for this study. As for the Netherlands, this country did not fall into the clusters with the countries that are the biggest cybercrime victims, but it was also not included in the clusters with the least.

Thus, the first hypothesis put forward at the beginning of the study is confirmed for such countries as the USA, China, Brazil, Spain, Vietnam, France, Germany, Italy, India, Indonesia, Iran, and Turkey. On the one hand, they are the most powerful countries in the world and the biggest sources of cyberattacks. On the other hand, they also belong to the clusters of countries that are the biggest victims of cyberattacks. Although many experts deny the existence of cyberwars, the obtained conclusion may indicate the covert and unconcealed cyberwars carried out by powerful countries because they have the largest military potential in the world. In our opinion, the reasons for this are the creation of opposition to such countries and their promotion of the reduction of the influence of others at the world level, harming their economic, social, and political sectors, and the formation of a negative image in the international political arena.

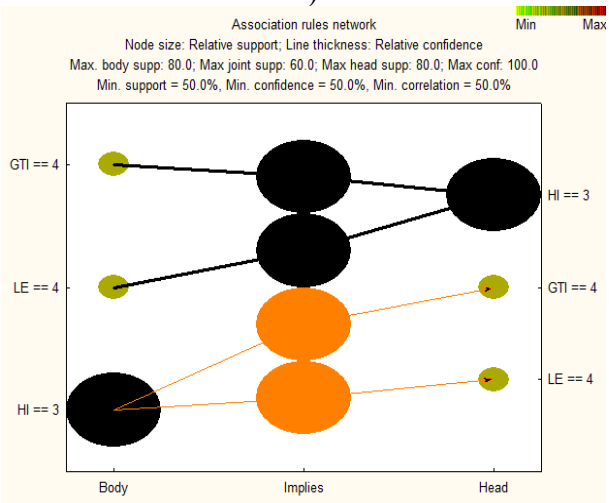
To accept or reject the second hypothesis, analysing the socio-economic profiles of countries' clusters formed depending on the detected cybercrimes is necessary. Thus, an associative analysis was performed using the STATISTICA analytical package. Its results are presented in *Figures 9-10*.



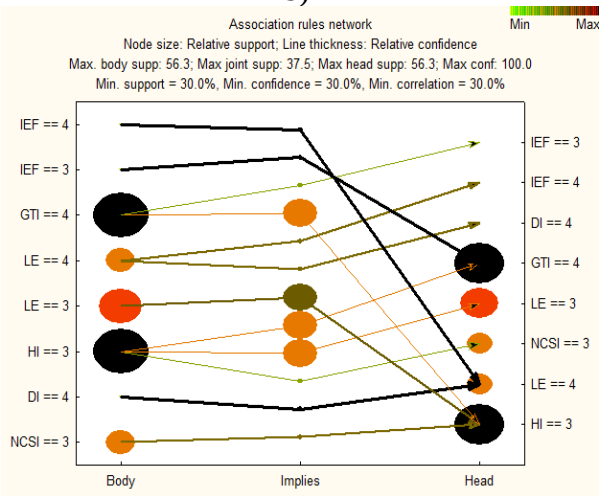
a)



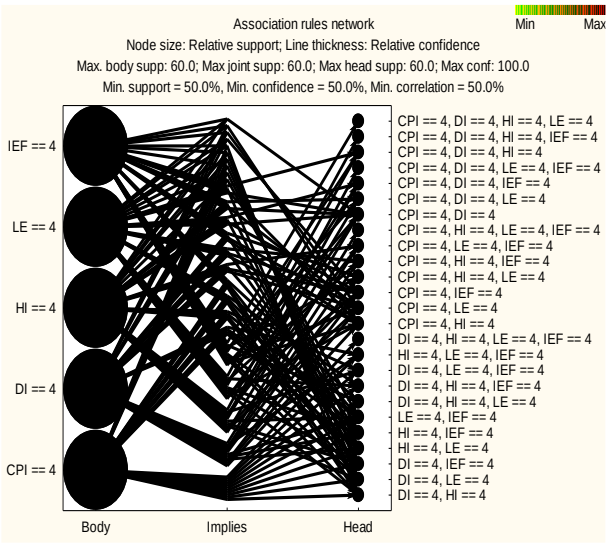
b)



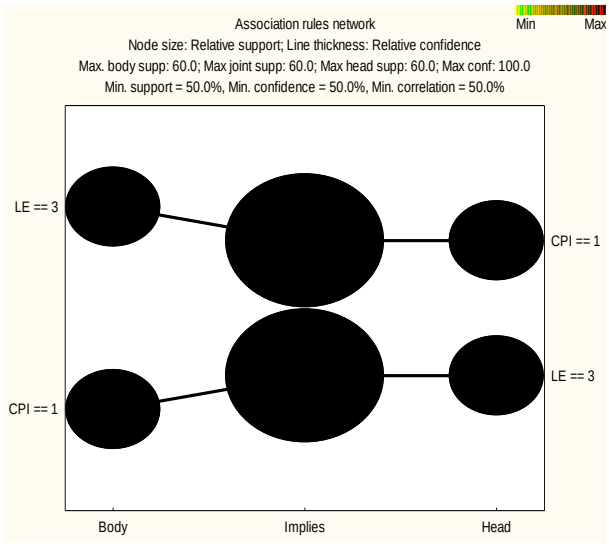
c)



d)

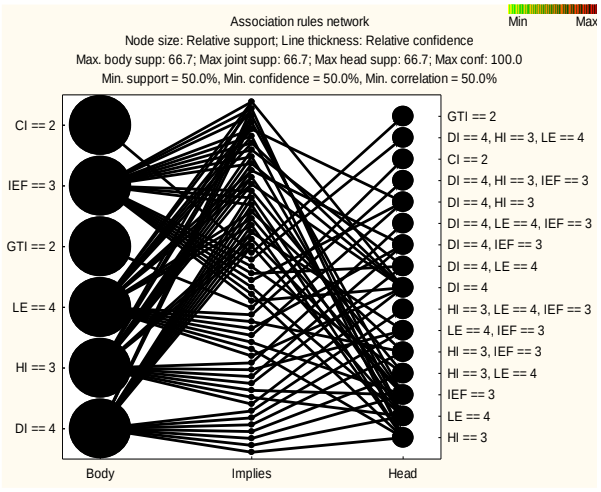


e)

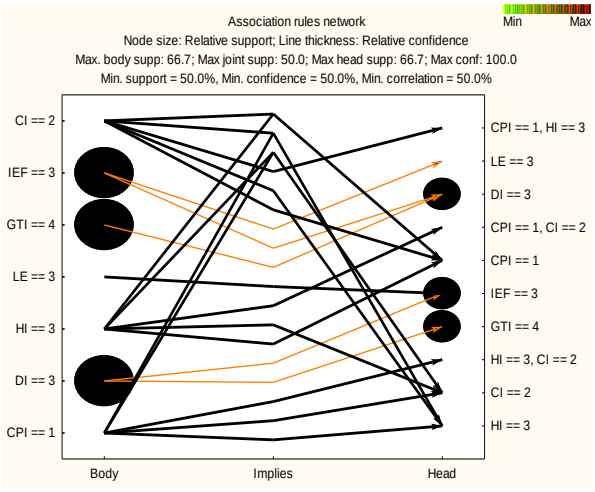


f)

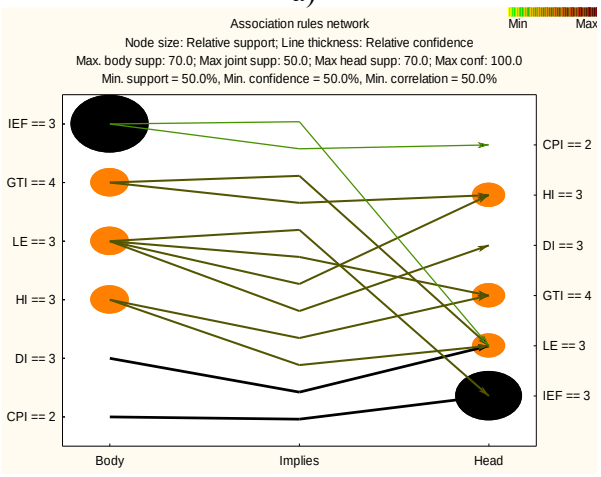
Figure 9. Results of associative analysis for clusters: a) 1.1; b) 1.2; c) 1.3; d) 2.1; e) 2.2; f) 2.3
Source: *own calculations*



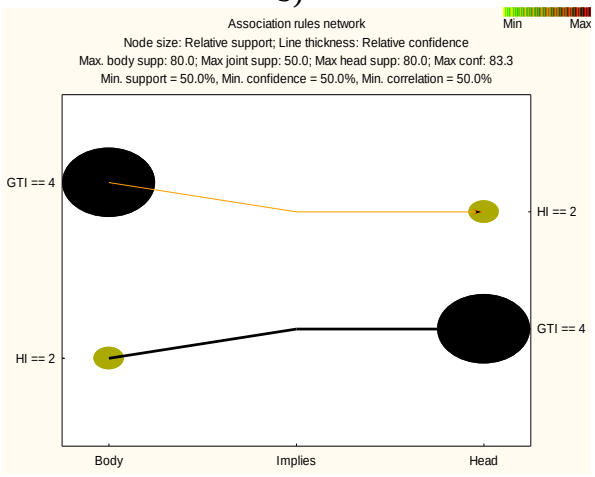
a)



b)



c)



d)

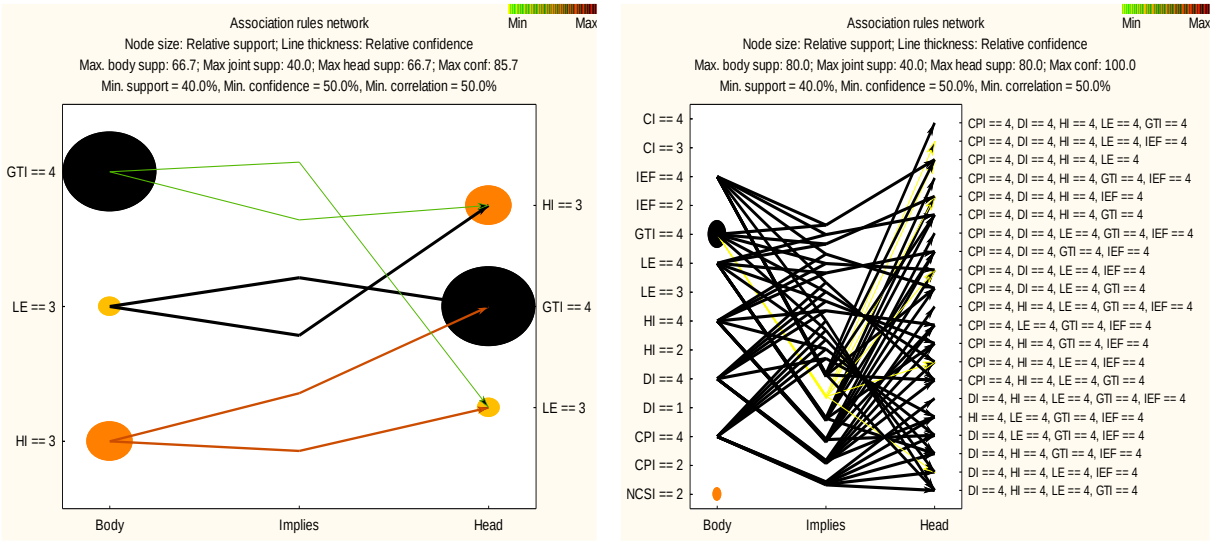


Figure 10. Results of associative analysis for clusters: a) 2.4; b) 3.1; c) 3.2; d) 4.1; e) 4.2; f) 4.3
Source: *own calculations*

Cluster 1.1 contains countries that differ greatly in their socio-economic development. That is why the results of the associative analysis presented in *Figure 9a* show only two standard features by which these countries can belong to this group. These are the Index of Economic Freedom and the Happiness Index. At the same time, the level of economic freedom above the average (IEF = 3) is one of the reasons for happiness at an above-average grade (HI = 3). This relationship is mutually dependent. The mutual support of the observations for this cluster is 45.45%, the confidence level ranges from 71.42% to 100%, and the probability that the countries will be in the same cluster is 0.85. Thus, some countries of cluster 1.1 are characterized by the fact that some of them are enormously powerful, are among the top countries initiating cyberattacks and have a mutually determined relationship between the level of economic freedom and happiness.

Figure 9b demonstrates the results of the associative analysis of cluster 1.2, where countries are the biggest victims of cyberattacks. It has been established that the essential characteristics correspond to indicators such as the National Cyber Security Index, Crime Index, Corruption Perceptions Index, Global Terrorism Index, Happiness Index and Life Expectancy at Birth. The results contain 182 associative rules for which the combined observational support is 50%, the confidence level ranges from 66.67% to 100%, and the probability of finding countries with the selected characteristics is 0.67-1.00. So, the countries of cluster 1.2 are countries with a low level of terrorism, an above average level of development of national cyber security, life expectancy, happiness, corruption and crime in the country. On the one hand, they are characterised by a combination of socio-economic development and the presence of crime and corruption, which can create a specific image for cybercriminals as countries that are the targets of cyber fraud to obtain financial and economic benefits.

The associative analysis of cluster 1.3 containing countries – the biggest victims of cyberattacks – made it possible to identify the essential characteristics of this group, which correspond to the Global Terrorism Index, Happiness Index and Life Expectancy at Birth (*Figure 9c*). At the same time, all three are both cause and effect for each other. For example, a country's low exposure to global terrorism (GTI = 4) is one of the reasons for above-average happiness (HI = 3), which, on the other hand, is the reason for a high life expectancy (LE = 4). Joint support for observations for which both cause and effect are actual is 60%. At the same

time, reliability for all observations ranges from 75% to 100%, and the probability that countries with the selected characteristics will be in the same cluster is relatively high and is equal to 0.87. It means that in most cases, the countries of cluster 1.3 have a low level of terrorism, a high life expectancy and an above-average level of happiness. Other features for these countries vary greatly.

A high level of cyber-attacks also characterizes clusters 2.1, 2.2 and 2.3, but compared to clusters 1.1, 1.2 and 1.3 countries, they are attacked much less (*Table 4*). *Figure 9d* shows the associative rules for cluster 2.1 that identified the essential features corresponding to the National Cyber Security Index, Democracy Index, Happiness Index, Life Expectancy at Birth, Global Terrorism Index, and Index of Economic Freedom. Although the value of the combined support of the observations varies from 31.25% to 37.50%, the reliability level is from 55.56% to 100% and the correlation is from 66.67% to 84.52%. The third of the countries in cluster 2.1 has a high level of democracy, a low level of exposure to global terrorism, above-average levels of happiness and cyber security, high and above-average levels of economic freedom and life expectancy. Unfortunately, expanding the list of features selected for analysis is necessary to form profiles of other countries from this cluster. It can also be assumed that this may be influenced by factors that are exceedingly difficult to detect analytically or by the absence of hidden motives of cybercriminals at all.

Figure 9e shows that a high level of economic freedom, life expectancy, social democracy, happiness and a low level of corruption characterizes the countries of cluster 2.2. At the same time, the level of support is equal to 60% for all associative rules, and the level of reliability and correlation is 100%. That is, 60% of the countries in this cluster belong to countries with an elevated level of socio-economic development, which can become the target of cybercrimes. For countries from group 2.3, the associative rules made it possible to identify such characteristics as the Corruption Perceptions Index and Life Expectancy at Birth (*Figure 9f*). At the same time, a high level of corruption is characteristic of the countries of this cluster. The level of support for this group of countries is 60%, and the level of reliability and correlation is 100%. One should note that the factor of an elevated level of corruption can be an indicator of the formation of a country's image that is attractive to cybercriminals.

Countries from clusters 2.4, 3.1, and 3.2 are also victims of cyberattacks, but compared to the previous groups, they became their targets much less (*Table 4*). The associative rules for group 2.4 are presented in *Figure 10a* and demonstrate such significant features as the Crime Index, Democracy Index, Happiness Index, Life Expectancy at Birth, Global Terrorism Index, and Index of Economic Freedom. It is valid for 66.67% of countries (Greece and Japan) at 100% reliability and correlation levels. One should note that high and above-average levels of economic development, happiness, life expectancy and democratic freedoms characterize this group. Also, there are countries with a rating of "2" (Greece and Kenya) for the impact on the global level of terrorism and crime. It means that the cluster united countries according to polar characteristics - positive socio-economic development and criminal problems.

Figure 10b shows the features of countries in cluster 3.1, classified as having low exposure to global terrorism, above-average happiness levels, democracy, economic freedom, life expectancy, crime, and corruption. The defined rules are fulfilled for 50% of the countries with reliability and correlation between 75% and 100%. The countries of cluster 3.2 are characterized by an above-average level of democracy, economic development, life expectancy and happiness, corruption and a low impact on the global level of terrorism (*Figure 10c*). It is provided with 50% support, 71.43% to 100.00% confidence, and 77.15% to 91.29% correlation. Based on the received considerations for the countries of clusters 3.1 and 3.2, the level of corruption may be a key factor for the commission of cybercrimes, but its influence may not be significant enough.

The clusters that include the countries with the lowest level of cybercrime are clusters 4.1, 4.2 and 4.3 (*Table 4*). *Figure 10d* shows that for cluster 4.1, only two rules were found that characterize the causal relationships between the Global Terrorism Index and the Happiness Index. Moreover, the size of the circle that corresponds to the Global Terrorism Index is large, indicating a high level of support for the cause and effect of this characteristic than for the Happiness Index. Mutual support of observations, in this case, is equal to 50%, the reliability level ranges from 62.5% to 83.3%, and the probability of being in one cluster is equal to 0.72. The obtained indicators are significant. The countries of this cluster have a low level of influence on global terrorism, making them unattractive for mass cyberattacks. *Figure 10e* shows the results of associative analysis for cluster 4.2. It was found that the Global Terrorism Index, Life Expectancy at Birth and Happiness Index are the features of the countries of this group. Mutual support of associations is equal to 40%, with fairly high confidence level values from 60% to 85.71%, as well as probabilities from 0.67 to 0.80. The countries of this cluster, as well as the previous group, are characterized by a low impact on the level of global terrorism (GTI = 4), but a strong connection between the cause of life expectancy and other indicators is also significant (LE = 3 $\Longrightarrow$ GTI = 4; LE = 3 $\Longrightarrow$ HI = 3). The associative analysis for the countries of cluster 4.3 revealed 642 associative rules between the characteristics corresponding to the eight analysed indicators (*Figure 10f*). At the same time, the combined support of associations is equal to 40%, with fairly high confidence level values from 50% to 100%, as well as probabilities from 0.58 to 1.00. So, this group consists of countries that may have a different combination of socio-economic features. It includes countries with high levels of democratic and economic freedoms for the population, happiness, life expectancy, low levels of corruption and exposure to global terrorism. Another group is countries with low levels of democracy, exposure to global terrorism, and below-average levels of cyber security, happiness, and economic freedom. That is, groups of countries that have the listed combinations of socio-economic characteristics in their profile are the least attractive for mass cyberattacks and wars from other countries.

Thus, the identified features of the profiles of the countries' clusters with the highest and lowest levels of cyberattacks can confirm the second hypothesis regarding the indirect influence of socio-economic development on their attractiveness to cyber criminals. It is indicated by the fact that the associative rules found in most cases are peculiar to countries with a high and above-average rating of socio-economic development. For other countries, patterns were not established. The influence of some of them, such as the level of corruption, crime, and influence on global terrorism, was revealed. It indicates factors not identified in the research process, which requires further study.

Conclusion

Today, cybercrime is an integral part of scientific and technological progress, the solution of which requires a lot of effort on the part of world organizations, governments of countries and simply interested persons. It also becomes a convenient tool for manipulating and achieving political, financial, military-strategic, psychological and other goals, both on the part of certain groups of people and state representatives. Mass cybercrime leads to significant financial losses and political, social and economic destabilisation. Therefore, this issue is often put on the agenda of such international organizations as the Economic and Social Council of the United Nations, the Council of Europe, the International Organization for Combating Cyberterrorism "IMPACT", the International Telecommunications Union and the United Nations Office on Drugs and Crime and others. In the framework of such cooperation, a complex of scientific, legal and organizational measures is being developed. It allows for

forming strategies to regulate and protect user behaviour in cyberspace. It is relevant regarding cyberwars carried out by individual countries to reduce the consequences of their aggression towards others. Therefore, the proposed study will be of interest to the analytical units of international organizations to identify potential cybercrime victims and develop special countermeasures and responsibility in cases of targeted cyberattacks that have led to catastrophic consequences.

In the framework of this article, authors hypothesized that countries with a strong Power Index are both initiators of cybercrimes against other countries and victims of cyber aggressions to a greater extent than countries with weak influence at the global level. The cluster analysis and comparison of its results with available statistical data fully confirmed this hypothesis. The most powerful countries, namely the USA, China, Brazil, Spain, Vietnam, France, Germany, Italy, India, Indonesia, Iran, and Turkey, were found to be more susceptible to cybercrime than others. At the same time, they are also sources of active cyberattacks towards others. The conclusions of this study can become the basis for developing appropriate strategies to deter such countries in cases of their active actions. This knowledge will be useful in forming a warning set of measures to monitor the flows of various types of transactions from those countries that are the sources of cyberattacks and belong to critical groups. The accumulation of retrospective data over a longer period and their use to expand the proposed research methodology will allow forming more likely structures of countries - victims of cybercrimes and countries - cyber predators.

Socio-economic profiles of countries, determined by the volume of detected cyberattacks conducted through email applications and the network, were formed based on the associative analysis. Its results identified features peculiar to most countries of the specified groups. Moreover, both their combinations and individual ones were highlighted, which can become a key factor in understanding the motives of cybercrimes on a global scale. The analysis of the profiles of countries that are attacked to a lesser extent showed that an important aspect of the lack of motivation for cybercriminals is the low impact of these countries on global terrorism. It also includes countries with a high level of socio-economic development and less developed countries. The analysis of the clusters of countries - the biggest victims of cybercrimes showed that, according to most characteristics, it included countries with a high and above average rating of socio-economic development, most of which are powerful and those that are the source of massive cyberattacks. In relation to other clusters, the influence of high corruption is important as an indicator of targeted cybercrimes for obtaining financial benefits. The obtained results made it possible to confirm the proposed hypothesis that countries' socio-economic development can indirectly motivate cybercriminals for mass cyberattacks, namely, the level of corruption, crime and influence on global terrorism can affect it. This analysis can help to improve the strategy of combating cybercrime at the level of an individual country and the world as a whole, considering key indicators that influence the motivation of cybercriminals.

Acknowledgement

This work was performed within the framework of state budget research No 0121U109559, No 0121U109553, No 0123U101945.

References

Adeyemo, K. A., Isiavwe, D., Adetula, D., Olamide, O., & Folashade, O. (2020). Mandatory adoption of the Central Bank of Nigeria's cashless and e-payment policy: implications for bank customers. *Banks and Bank Systems, 15*(2), 243-253. https://doi.org/10.21511/bbs.15(2).2020.21

Barabashev, A., Makarov, I., & Zarochintcev, S. (2022). How to shape government policies on high-technology development using the indicative evaluation of risks? *Administratie si Management Public, 38*, 70-89. https://doi.org/10.24818/amp/2022.38-04

Bayram, M., & Akat, M. (2019). Market-Neutral Trading with Fuzzy Inference, a New Method for the Pairs Trading Strategy. *Engineering Economics, 30*(4), 411-421. https://doi.org/10.5755/j01.ee.30.4.14350

Bing, C., & Schectman, J. (2019). *Inside the UAE's secret hacking team of American mercenaries*. Retrieved from: https://www.reuters.com/investigates/special-report/usa-spying-raven/ (31.01.2023).

Bozhenko, V. (2022). Tackling corruption in the health sector. *Health Economics and Management Review, 3*(3), 32-39. https://doi.org/10.21272/hem.2022.3-03

Bozhenko, V. V., Lyeonov, S. V., Polishchuk, E. A., Boyko, A. O., & Artyukhova, N. O. (2022). Identification of determinants of corruption in government: a mar-spline approach. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, 6*, 176-180. https://doi.org/10.33271/nvngu/2022-6/176

Bozhenko, V., Mynenko, S., & Shtefan, A. (2022b). Financial Fraud Detection on Social Networks Based on a Data Mining Approach. *Financial Markets, Institutions and Risks, 6*(4), 119-124. https://doi.org/10.21272/fmir.6(4).119-124.2022

Caballero-Morales, S.-O., Cordero Guridi, J. de J., Alvarez-Tamayo, R. I., & Cuautle-Gutiérrez, L. (2020). EDUCATION 4.0 to support entrepreneurship, social development and education in emerging economies. *International Journal of Entrepreneurial Knowledge, 8*(2), 89-100. https://doi.org/10.37335/ijek.v8i2.119

Chen, Y., Xu, S., Lyulyov, O., & Pimonenko, T. (2023). China's digital economy development: incentives and challenges. *Technological and Economic Development of Economy, 29*(2), 518-538. https://doi.org/10.3846/tede.2022.18018

Ćwiklicki, M., & Wojnarowska, M. (2020). Circular Economy and Industry 4.0: One-Way or Two-way Relationships? *Engineering Economics, 31*(4), 387-397. https://doi.org/10.5755/j01.ee.31.4.24565

DavidPur, N. (2022). *Which Countries are Most Dangerous? Cyber Attack Origin – by Country*. Retrieved from: https://blog.cyberproof.com/blog/which-countries-are-most-dangerous (31.01.2023).

Dečman, M., Stare, J., & Klun, M. (2022). The impact of the COVID-19 crisis on the development of the information society in Slovenia. *Administratie si Management Public, 39*, 77-96. https://doi.org/10.24818/amp/2022.39-05

Deutsche Welle (2022). *Ukrainian websites hacked in 'global attack'*. Retrieved from: https://www.dw.com/en/ukraine-government-websites-hacked-in-global-attack/a-60421475 (31.01.2023).

Dluhopolskyi, O., Pakhnenko, O., Lyeonov, S., Semenog, A., Artyukhova, N., Cholewa-Wiktor, M., & Jastrzębski, W. (2023). Digital financial inclusion: COVID-19 impacts and opportunities. *Sustainability (Switzerland), 15*(3), 2383. https://doi.org/10.3390/su15032383

Economist Intelligence (2023). *Democracy Index*. Retrieved from: https://www.eiu.com/n/campaigns/democracy-index-

2022/?utm_source=google&utm_medium=paid-search&utm_campaign=democracy-index-2022&gclid=CjwKCAjwgqejBhBAEiwAuWHioAEruOQA25JyHg-61MBEiYNJp9hvu3Pf91E_tWO2W0nauZ6on003ORoC6UsQAvD_BwE (31.01.2023).

E-Governance Academy (2023). *National Cyber Security Index*. Retrieved from: https://ncsi.ega.ee/ncsi-index/ (31.01.2023).

Fobel, P., & Kuzior, A. (2019). The future (Industry 4.0) is closer than we think. Will it also be ethical? Paper presented at the *AIP Conference Proceedings, 2186*. https://doi.org/10.1063/1.5137987

Glova, J., Bernatik, W., & Tulai, O. (2020). Determinant Effects of Political and Economic Factors on Country Risk: An Evidence from the EU Countries. *Montenegrin Journal of Economics, 16*(1), 37-53. https://doi.org/10.14254/1800-5845/2020.16-1.3

Gontareva, I., Babenko, V., Kuchmacz, B., & Arefiev, S. (2020). Valuation of information resources in the analysis of cybersecurity entrepreneurship. *Estudios De Economia Aplicada, 38*(4), https://doi.org/10.25115/EEA.V38I4.3984

Gupta, A., & Mishra, M. (2022). Ethical Concerns While Using Artificial Intelligence in Recruitment of Employees. *Business Ethics and Leadership, 6*(2), 6-11. https://doi.org/10.21272/bel.6(2).6-11.2022

Gurbanov, N., Yagublu, N., Akbarli, N., & Niftiyev, I. (2022). Digitalization and the Covid-19-led public crisis management: an evaluation of financial sustainability in the Azerbaijan business sector. *SocioEconomic Challenges, 6*(3), 23-38. https://doi.org/10.21272/sec.6(3).23-38.2022

Institute for Economics and Peace (2022). *Global Terrorism Index 2022*. Retrieved from: https://reliefweb.int/report/world/global-terrorism-index-2022 (31.01.2023).

Kaspersky (2023). *Cyberthreat real-time map*. Retrieved from: https://cybermap.kaspersky.com/ (31.01.2023).

Krebs, B. (2021). *At Least 30,000 U.S. Organizations Newly Hacked Via Holes in Microsoft's Email Software*. Retrieved from: https://krebsonsecurity.com/2021/03/at-least-30000-u-s-organizations-newly-hacked-via-holes-in-microsofts-email-software/ (31.01.2023).

Kumar, N., & Kumar, J. (2019). Efficiency 4.0 for Industry 4.0. *Human Technology, 15*(1), 55-78. https://doi.org/10.17011/ht/urn.201902201608

Kurniawati, E., Kohar, U.H.A., & Pirzada, K. (2022). Change or destroy: the digital transformation of Indonesian MSMES to achieve sustainable economy. *Polish Journal of Management Studies, 26*(2), 248-264. https://doi.org/10.17512/pjms.2022.26.2.15

Kuzior, A., & Kwilinski, A. (2022). Cognitive technologies and artificial intelligence in social perception. *Management Systems in Production Engineering, 30*(2), 109-115. https://doi.org/10.2478/mspe-2022-0014

Kuzmenko, O., Šuleř, P., Lyeonov, S., Judrupa, I., & Boiko, A. (2020). Data mining and bifurcation analysis of the risk of money laundering with the involvement of financial institutions. *Journal of International Studies, 13*(3), 332-339. https://doi.org/10.14254/2071-8330.2020/13-3/22

Lăzăroiu, G., Androniceanu, A., Grecu, I., Grecu, G., & Neguriță, O. (2022). Artificial intelligence-based decision-making algorithms, Internet of Things sensing networks, and sustain-able cyber-physical management systems in big data-driven cognitive manufacturing. *Oeconomia Copernicana, 13*(4), 1047-1080. https://doi.org/10.24136/oc.2022.030

Lucas, G. (2016). *Ethics and Cyber Warfare: The Quest for Responsible Security in the Age of Digital Warfare*. Oxford University Press.

Lyulyov, O., Lyeonov, S., Tiutiunyk, I., & Podgórska, J. (2021). The impact of tax gap on macroeconomic stability: Assessment using panel VEC approach. *Journal of International Studies, 14*(1), 139-152. https://doi.org/10.14254/2071-8330.2021/14-1/10

Mačiulytė-Šniukienė, A., Butkus, M., & Davidavičienė, V. (2022). Development of the model to examine the impact of infrastructure on economic growth and convergence. *Journal of Business Economics and Management, 23*(3), 731-753. https://doi.org/10.3846/jbem.2022.17140

Melnyk, L., Derykolenko, O., Kubatko, O., & Matsenko, O. (2019). Business models of reproduction cycles for digital economy. Paper presented at the *CEUR Workshop Proceedings, 2393*, 269-276. Retrieved from https://www.scopus.com/record/display.uri?eid=2-s2.0-85069504652&origin=resultslist

Melnyk, L., Kubatko, O., Piven, V., Klymenko, K., & Rybina, L. (2021). Digital and economic transformations for sustainable development promotion: A case of OECD countries. *Environmental Economics, 12*(1), 140-148. https://doi.org/10.21511/EE.12(1).2021.12

Millia, H., Adam, P, Muhatlib, A. A., & Tajuddin and Pasrun, Y. P. (2022). The Effect of Inward Foreign Direct Investment and Information and Communication Technology on Economic Growth in Indonesia. *AGRIS on-line Papers in Economics and Informatics, 14*(1), 69-79. https://doi.org/10.7160/aol.2022.140106

Mnohoghitnei, I., Horobeț, A., & Belaşcu, L. (2022). Bitcoin is so Last Decade-How Decentralized Finance (DeFi) could Shape the Digital Economy. *European Journal of Interdisciplinary Studies, 14*(1), 87-99. https://doi.org/10.24818/ejis.2022.01

Numbeo (2023). *Crime Index by Country 2022*. Retrieved from: https://www.numbeo.com/crime/rankings_by_country.jsp?title=2022 (31.01.2023).

Orlov, V., Bukhtiarova, A., Marczuk, M., & Heyenko, M. (2021). International economic and social determinants of the state economic security: A causal analysis. *Problems and Perspectives in Management, 19*(4), 301-310. https://doi.org/10.21511/ppm.19(4).2021.24

Pakhnenko, O., & Kuan, Z. (2023). Ethics of Digital Innovation in Public Administration. *Business Ethics and Leadership, 7*(1), 113-121. https://doi.org/10.21272/bel.7(1).113-121.2023

Pakhnenko, O., Rubanov, P., Girzheva, O., Ivashko, L., Britchenko, I., & Kozachenko, L. (2022). Cryptocurrency: Value formation factors and investment risks. *Journal of Information Technology Management, 14*, 179-200. https://doi.org/10.22059/JITM.2022.88896

Perlroth, N., Scott, M, & Frenkel, S. (2017). *Cyberattack Hits Ukraine Then Spreads Internationally*. Retrieved from: https://www.nytimes.com/2017/06/27/technology/ransomware-hackers.html (31.01.2023).

Remeikienė, R., Ligita, G., Fedajev, A., Raistenskis, E., & Krivins, A. (2022). Links between crime and economic development: EU classification. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 17*(4), 909-938. https://doi.org/10.24136/eq.2022.031

Rousseeuw, P.J. (1987). Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Computational and Applied Mathematics, 20*, 53-65. https://doi.org/10.1016/0377-0427(87)90125-7

Safarov, G., Sadiqova, S., Urazayeva, M., & Abbasova, N (2022). Theoretical and Methodological Aspects of Innovative-Industrial Cluster Development in the Era of

Digitalization. *Marketing and Management of Innovations, 4*, 184-197. https://doi.org/10.21272/mmi.2022.4-17

Șavga, L. (2019). Implementing the Smart Specialization Concept in the Republic of Moldova: Challenges and Initiatives. *Journal of Research on Trade, Management and Economic Development, 6*(2), 6-17.

Șavga, L., & Baran, T. (2022). Boosting the process of smart specialization in the Republic of Moldova. Paper presented in *Contemporary Issues in Economy and Technology* (pp. 187-196).

Shao, X., Wang, D., Li, X., & Shao, H. (2022). Impact of Internet technology on spatial technology heterogeneity: openness or convergence - evidence from provincial data in China. *Transformations in Business & Economics, 21*(2), 193-213.

Shkolnyk, I., Frolov, S., Orlov, V., Datsenko, V., & Kozmenko, Y. (2022). The impact of financial digitalization on ensuring the economic security of a country at war: New measurement vectors. *Investment Management and Financial Innovations, 19*(3), 119-138. https://doi.org/10.21511/imfi.19(3).2022.11

Smith, E.T. (2013). Cyber warfare: a misrepresentation of the true cyber threat. *American Intelligence Journal, 31*(1), 82-85.

Sobczak, A. (2022). Analysis of the Conditions Influencing the Assimilation of the Robotic Process Automation by Enterprises. *Human Technology, 18*(2), 143-190. doi: 10.14254/1795-6889.2022.18-2.4

Statista (2023*). Most commonly reported cyber crime categories worldwide in 2022, by number of individuals affected.* Retrieved from: https://www.statista.com/statistics/184083/commonly-reported-types-of-cyber-crime-global/ (31.01.2023).

Stehel, V., Vochozka, M., Kliestik, T., & Bakes, V. (2019). Economic analysis of implementing VMI model using game theory. *Oeconomia Copernicana, 10*(2), 253-272. https://doi.org/10.24136/oc.2019.013

Straková, J., Talíř, M., & Váchal, J. (2022). Opportunities and threats of digital transformation of business models in SMEs. *Economics and Sociology, 15*(3), 159-171. https://doi.org/10.14254/2071-789X.2022/15-3/9

The Heritage Foundation (2023). *2023 Index of Economic Freedom*. Retrieved from: https://www.heritage.org/index/download (31.01.2023).

The World Bank (2023). *Life expectancy at birth, total (years)*. Retrieved from: https://data.worldbank.org/indicator/SP.DYN.LE00.IN (31.01.2023).

Tiutiunyk, I. V., Zolkover, A. O., Lyeonov, S. V., & Ryabushka, L. B. (2022a). The impact of economic shadowing on social development: challenges for macroeconomic stability. *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, 1*, 183-191. https://doi.org/10.33271/nvngu/2022-1/183

Tiutiunyk, I., Cieśliński, W., Zolkover, A., & Vasa, L. (2022b). Foreign direct investment and shadow economy: One-way effect or multiple-way causality? *Journal of International Studies, 15*(4), 196-212. https://doi.org/10.14254/2071-8330.2022/15-4/12

Tran, L. Q. T., Phan, D. T., Herdon, M., & Kovacs, L. (2022). Assessing the Digital Transformation in Two Banks: Case Study in Hungary. *AGRIS on-line Papers in Economics and Informatics, 14*(2), 121-134. https://doi.org/10.7160/aol.2022.140210

Transparency International (2023). *Corruption_Perceptions_Index*. Retrieved from: https://www.transparency.org/en/cpi/2021?gclid=CjwKCAjw67ajBhAVEiwA2g_jEPyd355cvDdhD7SdWVteYeer5WvV3BZFHMo-Ox6p3vXSGk9wKi4p4BoCRJgQAvD_BwE (31.01.2023).

Tribune (2020). *Major cyber attack by Indian intelligence identified: ISPR*. Retrieved from: https://tribune.com.pk/story/2259193/major-cyber-attack-by-indian-intelligence-identified-ispr (31.01.2023).

Tvaronaviciene, M., & Burinskas, A. (2020). Industry 4.0 significance to competition and the eu competition policy. *Economics & Sociology, 13*(3), 244-258. https://doi.org/10.14254/2071-789X.2020/13-3/15

U.S. Department of Homeland Security (2016). *Joint Statement from the Department of Homeland Security and Office of the Director of National Intelligence on Election Security*. Retrieved from: https://www.dhs.gov/news/2016/10/07/joint-statement-department-homeland-security-and-office-director-national (31.01.2023).

Vasudevan, H. (2022). Management and Leadership in the Klang Valley IT Sector: Conceptual Approach. *Marketing and Management of Innovations, 3*, 56-65. https://doi.org/10.21272/mmi.2022.3-05

Vitvitskiy, S. S., Kurakin, O. N., Pokataev, P. S., Skriabin, O. M., & Sanakoiev, D. B. (2021). Peculiarities of cybercrime investigation in the banking sector of Ukraine: review and analysis. *Banks and Bank Systems, 16*(1), 69-80. https://doi.org/10.21511/bbs.16(1).2021.07

Voo, J., Hemani, I., & Cassidy, D. (2022). *National Cyber Power Index 2022*. Retrieved from: https://www.belfercenter.org/sites/default/files/files/publication/CyberProject_National%20Cyber%20Power%20Index%202022_v3_220922.pdf (31.01.2023).

Voronenko, I., Nehrey, M., Laptieva, A., Babenko, V., & Rohoza, K. (2022). National cybersecurity: Assessment, risks and trends. *International Journal of Embedded Systems, 15*(3), 226-238. https://doi.org/10.1504/IJES.2022.124854

Wang, Q., Chen, Y., Guan, H., Lyulyov, O., & Pimonenko, T. (2022). Technological innovation efficiency in China: Dynamic evaluation and driving factors. *Sustainability (Switzerland), 14*(14). https://doi.org/10.3390/su14148321

Wisevoter (2023). *Most Powerful Countries in the World*. Retrieved from: https://wisevoter.com/country-rankings/most-powerful-countries-in-the-world/ (31.01.2023).

World Happiness Report (2023). *World Happiness Report 2022*. Retrieved from: https://worldhappiness.report/ed/2022/ (31.01.2023).

Yarovenko, H. (2020). Evaluating the threat to national information security. *Problems and Perspectives in Management, 18*(3), 195-210. https://doi.org/10.21511/ppm.18(3).2020.17

Yarovenko, H., & Rogkova, M. (2022). Dynamic and bibliometric analysis of terms identifying the combating financial and cyber fraud system. *Financial Markets, Institutions and Risks, 6*(3), 93-104. https://doi.org/10.21272/fmir.6(3).93-104.2022

Yoshimori, H., & Yoshimori, M. (2022). An Education Gift – Integrated Cognitive and Non-Cognitive Skills – for Future Generations to Grow the Economy in the Digital Phase. *SocioEconomic Challenges, 6*(2), 5-18. https://doi.org/10.21272/sec.6(2).5-18.2022

Yu, Y., Xinxin, W., Ruoxi, L., & Tingting, Y. (2023). The Mediating Role of Human Capital in the Relationship between Education Expenditure and Science and Technology Innovation: Evidence from China. *SocioEconomic Challenges, 7*(1), 129-138. https://doi.org/10.21272/sec.7(1).129-138.2023

Zimaitis, I., Urbonavičius, S., Degutis, M., & Kaduškevičiūtė, V. (2022). Influence of trust and conspiracy beliefs on the disclosure of personal data online. *Journal of Business Economics and Management, 23*(3), 551-568. https://doi.org/10.3846/jbem.2022.16119